National Institute for Health and Care Excellence

Final

Osteoarthritis in over 16s: diagnosis and management

[C] Evidence review for the clinical and cost-effectiveness of exercise for the management of osteoarthritis

NICE guideline NG226

Evidence reviews underpinning recommendations 1.3.1 to 1.3.3 and research recommendations in the NICE guideline

October 2022

Final

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-4740-9

Contents

1 Exercise for the management of osteoarthritis

1.1 Review question

What is the clinical and cost-effectiveness of exercise therapy for the management of osteoarthritis?

1.1.1 Introduction

The benefits of exercise for mental and physical wellbeing are widely accepted. It is known to reduce the risk of major illnesses, such as heart disease, stroke, type 2 diabetes and cancer. Research shows that physical activity can also boost self-esteem, mood, sleep quality and energy, as well as reducing the risk of stress, depression, dementia and Alzheimer's disease. Internationally, recommendations for the use of exercise for osteoarthritis is widespread, however, often exercise programmes are more intensive or delivered for longer than what is provided within the NHS.

Current practice for people with osteoarthritis is to recommend exercise (both aerobic exercise and local joint specific exercises) as a first line core treatment, this may be through provision of information to support home exercises, through face-to-face physiotherapy sessions or within a group exercise programme. There is not a standardised approach to delivering exercise for osteoarthritis on the NHS and approaches are also tailored to patients' needs and preferences. While exercise can seem counterintuitive to people in pain exercises for people with osteoarthritis, it is thought to play a role in maintaining and improving the ability to move and function and to reduce pain in the longer term. This review aims to investigate the effectiveness of supervised and unsupervised exercises including strength, aerobic, flexibility, proprioception and mixed modality exercises have on osteoarthritis.

1.1.2 Summary of the protocol

Table 1: PICO characteristics of review question

Population	Inclusion: • Adults (age ≥16 years) with osteoarthritis affecting any joint Exclusion: • Children (age <16 years) • People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, hemochromatosis, haemophilic arthropathy, diseases of childhood that may predispose to osteoarthritis, and malignancy). • Studies with an unclear population (e,g, proportion of participants with osteoarthritis unclear). • Spinal osteoarthritis
Interventions	Interventions (minimum duration 1 week): • Supervised strength exercise • Supervised aerobic exercise • Other supervised exercise (including flexibility, proprioception)* • Supervised mixed modality exercise (e.g. aerobic and strength exercise combined) • Unsupervised strength exercise

	• Unsupervised aerobic exercise • Other unsupervised exercise (including flexibility, proprioception)* • Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined) *Subgroup analysis if heterogeneity is present within this group
Comparisons	• Each other • Pharmacological treatment*** • No exercise intervention (including either): ○ Exercise versus no treatment* ○ Exercise plus additional treatment versus additional treatment alone** **No treatment defined as either (1) doing nothing or (2) very low intensity intervention such as advice* ***Inclusion of studies where additional treatment is the same in each arm will be assessed on a case by case basis. Studies including high intensity additional treatment may not be included due to the risk that treatment could have an interaction with the intervention of interest and mask the true treatment effect.* ****Pool classes of pharmacological treatment but conduct subgroup analysis if heterogeneity is present*
Outcomes	Stratify by ≤/>3 months (longest time-point in each): • Health-related quality of life [validated patient-reported outcomes, continuous data prioritised] • Physical function [validated patient-reported outcomes, continuous data prioritised] • Pain [validated patient-reported outcomes, continuous data prioritised] • Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flares [validated patient-reported outcomes, continuous data prioritised] • Serious adverse events [dichotomous]
Study design	Systematic reviews of RCTs and RCTs

For full details see the review protocol in Appendix A.

1.1.3 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual. Methods specific to this review question are described in the review protocol in Appendix A and the methods document.

Declarations of interest were recorded according to NICE's conflicts of interest policy.

1.1.4 Effectiveness evidence

1.1.4.1 Included studies

One-hundred and twenty seven RCT studies were included in the review[2, 6, 12, 18, 22, 29, 35, 45, 46, 52, 56, 57, 59, 66, 67, 72, 77, 79-81, 83-85, 89, 91, 100, 101, 106, 107, 111, 114, 116, 119, 122, 129, 134, 145, 148, 153, 157, 158, 173, 177-180, 183, 185, 192, 193, 196, 203, 206, 214-216, 221-225, 227, 228, 233, 234, 237, 242, 248-251, 256, 264-266, 268, 287, 289, 310, 314, 315, 320, 322, 328, 330, 332, 337, 347-350, 355, 369, 374, 376, 377, 379, 380, 382, 384, 386-388, 390, 393, 400, 401, 403, 405, 408, 415, 430, 439, 451, 458, 459, 467, 478, 479, 485, 489, 491, 496-49838, 444]; these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3).

The clinical studies identified included the following comparisons:

- Supervised strength exercise compared to unsupervised strength exercise
- Supervised strength exercise compared to supervised aerobic exercise
- Supervised strength exercise compared to no treatment
- Unsupervised strength exercise compared to unsupervised aerobic exercise
- Unsupervised strength exercise compared to no treatment
- Supervised aerobic exercise compared to no treatment
- Unsupervised aerobic exercise compared to no treatment
- Other supervised exercise compared to supervised strength exercise
- Other supervised exercise compared to unsupervised strength exercise
- Other supervised exercise compared to no treatment
- Other unsupervised exercise compared to unsupervised strength exercise
- Supervised mixed modality exercise compared to supervised strength exercise
- Supervised mixed modality exercise compared to unsupervised strength exercise
- Supervised mixed modality exercise compared to supervised aerobic exercise
- Supervised mixed modality exercise compared to other supervised exercise
- Supervised mixed modality exercise compared to unsupervised mixed modality exercise
- Supervised mixed modality exercise compared to pharmacological treatment
- Supervised mixed modality exercise compared to no treatment
- Unsupervised mixed modality exercise compared to unsupervised strength exercise
- Unsupervised mixed modality exercise compared to other unsupervised exercise
- Unsupervised mixed modality exercise compared to pharmacological treatment
- Unsupervised mixed modality exercise compared to no treatment

A network meta-analysis was not conducted for this review. This was because the categories used for the interventions had considerable variability in what was provided. For example: While exercise may be noted as supervised strength exercise, some exercises may be more intense than others (for example: including more repetitions, longer duration of therapy). For other supervised and unsupervised exercise and mixed modality exercises, the types of exercise offered could vary significantly from others within the same comparison. Therefore, this would introduce heterogeneity into a network meta-analysis that would affect the coherence and make the results difficult to interpret.

See also the study selection flow chart in Appendix C, study evidence tables in Appendix D, forest plots in Appendix E and GRADE tables in Appendix F.

1.1.4.2 Excluded studies

Cochrane reviews were identified but could not be included due to incorrect population (Ashworth 2005[28], Jordan 2010[213]), different interventions (Fransen 2003[139], Fransen 2015[140]), different comparisons (Bartels 2016[32], Fransen 2014[143], Osteras 2017[336], Regnaux 2015[370], Witteveen 2013[487], Zammit 2010[501]). The references were checked any studies that fulfilled the inclusion criteria were included.

See the excluded studies list in Appendix J.

1.1.5 Summary of studies included in the effectiveness evidence

1.1.5.1 Supervised strength exercise compared to unsupervised strength exercise

Table 2: Summary of studies included in the evidence review for the supervised strength exercise compared to unsupervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Kuru colak 2017[251]	**Supervised strength exercise** (n=39) Low-intensity therapeutic isometric and isotonic exercises for major muscle groups in both lower extremities and simple balance exercises Group or individual: Individual session **Unsupervised strength exercise** (n=39) Exercises taught under the supervision of a physiotherapist in one session, followed up by instruction to perform the exercises at home at least three times a week Group or individual: Individual session **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (range): 60 (49-84) N = 78 Definition: Kellgren Lawrence grade 2-3 knee osteoarthritis Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months	
Nambi 2020[315]	**Supervised strength exercise** (n=20)	**Knee osteoarthritis** Mean age (SD): 22.5 (1.5) years	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	The training knee was kept at a 90 degrees flexed position, and the dynamometer axis was aligned with the centre of the lateral femoral condyle. The knee was tested from 0 degrees to 120 degrees of flexion, where 0 degrees was considered full extension. Training was performed on 5 days a week for 4 weeks. **Other supervised exercise** (n=20) Sensory motor training which was given in 3 stages; static, dynamic and functional. All exercises were performed 5 times in 1 set, for 3 sets, with a sufficient with 5 minutes rest period between sets for 4 weeks. **Unsupervised strength exercise** (n=20) Home-based exercises performed with 10-15 repetitions/day, 5 days a week for 4 weeks. Stretching was focused on each muscle group for 3 repetitions of 15 s per muscle group.	N = 60 Definition: Chronic osteoarthritis after ACL injury (secondary osteoarthritis) Severity: Not stated/unclear Duration of injury (SD): 5.4 (0.4) months. Presence of multimorbidities: Not stated/unclear		

Study	Intervention and comparison	Population	Outcomes	Comments
	Concurrent medication/care: No additional information.			

1.1.5.2 Supervised strength exercise compared to supervised aerobic exercise

Table 3: Summary of studies included in the evidence review for the supervised strength exercise compared to supervised aerobic exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Beckwee 2017[38]	**Supervised strength exercise** (n=19) 45 minute strength training sessions focussed on the lower limb delivered as 54 sessions over 18 weeks (18 supervised, the remainder unsupervised) Group or individual: Individual session **Supervised aerobic exercise** (n=18) Walking-based exercise for the same duration. Group or individual: Individual session **Concomitant therapy:** No additional information.	**Knee osteoarthritis** Mean age (SD): 61.8 (8.88) years N = 37 Definition: People who fulfilled the criteria defined by the American College of Rheumatology for knee osteoarthritis. Severity: Kellgren Lawrence grade 1-4, median grade 1-2. Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at >3 months	
Bieler 2017[52]	**Supervised strength exercise** (n=50) Machine based strength training with three resistances exercises (leg press, seated knee extension, hip extension)	**Hip osteoarthritis** Mean age (SD): 69.6 (6.1) years N = 152 Definition:	Pain at >3 months Physical function at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Individual session **Supervised aerobic exercise** (n=50) Nordic walking with progressive intensity for 1 hour three times weekly Group or individual: Group session A third group (n=52) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** Individual counselling, a one hour patient education session on the important of exercise and some telephone assisted counselling to improve adherence	Clinical hip osteoarthritis according to the American College of Rheumatology Severity: Kellgren Lawrence grade 2.1 (1.5) Duration of symptoms (mean [SD]): 6.1 (6.3 years) Presence of multimorbidities: High morbidity score		
Bokaeian 2021[56]	**Other supervised exercise** (n=22) YogaMT group who practiced MT gait for 20 minutes and performed three repetitions of goddess squat and warrior lunge exercises with a minimum rest interval of 40s, supervised by a physiotherapist. For the MT gait, they were trained to walk with slight hip internal rotation and knee flexion (about 20	**Knee osteoarthritis** Age (mean, SD): 56.1 (5.0) N = 59 Definition: People with unilateral or bilateral knee osteoarthritis diagnosed radiographically and by symptoms	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	degrees) at their selected speed on the treadmill. Duration 4 weeks. **Supervised strength exercise** (n=19) The KMS group that consisted of three 10-repetition sets of resistive knee extension/flexion exercises with 2-min interval, using the quadriceps chair. The maximum load that each participant could lift to complete 10-repetition maximum without pain was determined to adjust the amount of load for each exercise weekly. Exercise was performed under supervision of a physiotherapist. Duration 4 weeks. **Supervised aerobic exercise** (n=18) The TMW group that practiced walking on a treadmill at their self-selected speed for 20 minutes, with no gait modification, under the supervision of the physiotherapist. Duration 4 weeks. **Concomitant therapy:** People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also	Severity: Kellgren Lawrence grades 2-3 Duration of symptoms: history of pain for >one month Presence of multimorbidities: Not applicable		

Study	Intervention and comparison	Population	Outcomes	Comments
	instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification was also offered.			
Samut 2015[390]	**Supervised strength exercise** (n=15) Isokinetic exercise performed 3 days/week for 6 weeks using the Biodex isokinetic system. 5 concentric fexion and extension at angular belocities of 60-180 degrees/s with one set of contractions increased to 6 sets by the end of the study. Group or individual: Individual **Supervised aerobic exercise** (n=14) Aerobic exercise performed 3 days/week for 6 weeks on a treadmill. Exercise intensity was adjusted for 65-70% of age-related heart rate for the first 4 weeks and 70-75% for the next 2 weeks. Group or individual: Individual **No treatment** (n=13) **Concomitant therapy:**	**Knee osteoarthritis** Mean age (SD): 60.4 (7.8) years N = 42 Definition: Kellgren Lawrence grade 2-3 knee osteoarthritis fulfilling the American College of Rheumatology criteria Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: No stated Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Each person was allowed to take paracetamol whenever needed			

1.1.5.3 Supervised strength exercise compared to pharmacological treatment

Table 4: Summary of studies included in the evidence review for the supervised strength exercise compared to pharmacological treatments comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Chao, 2020[80]	**Supervised strength exercise** (n=105) Systematic exercise rehabilitation program mainly including lower limb static, dynamic and flexibility exercises; exercises targeting the gluteus muscles; and core strength training for 20 minutes per day. Duration 12 weeks. Group or individual : Individual session Type of exercise: Not applicable **Pharmacological treatment** (n=80) Administration of NSAIDs and COX-2 inhibitors. In this trial, naproxen and diclofenac were administrated to patients, respectively (27 had diclofenac, 28 had naproxen, 19 had celecoxib). All people received	Knee osteoarthritis Mean age (SD): 56.3 (10.1) years N = 185 Definition: Diagnosis of knee osteoarthritis, Kellgren Lawrence grades I to III with obvious symptoms Severity: Kellgren Lawrence grade I to III Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	the same drug dosage.. Duration 12 weeks. Class of medicine: Oral treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable **Concomitant therapy:** No additional information			

1.1.5.3 Supervised strength exercise compared to no treatment

Table 4: Summary of studies included in the evidence review for the supervised strength exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Anon 2016[119]	**Supervised strength exercise** (n=19) Exercise programme to strengthen quadriceps and hamstrings. Five sessions per week for 15 days, follow by a home regimen Group or individual: Individual session **No treatment** (n=16)	**Knee osteoarthritis** Mean age (SD): 51.1 (6.0) years N = 35 Definition: ACR grade 1-3 osteoarthritis Severity: ACR grade 1-3 osteoarthritis Duration of symptoms (mean [SD]): 33.9 (36.9) months	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No exercise treatment **Concomitant therapy:** 30 minutes along with TENS (pulse duration of 150msec, frequency of 120Hz, amplitude of 50mA)	Presence of multimorbidities: Not stated/unclear		
Anwer 2014[22]	**Supervised strength exercise** (n=21) Exercises including isometric quadriceps exercise, straight leg raises, isometric hip adduction exercise, five times per week for five weeks Group or individual: Individual session **No treatment** (n=21) Usual activity **Concomitant therapy:** Ultrasound therapy as per the patient's requirement with 1.5 watts/cm² for 7 minutes in continuous mode at the tender point around the knee joint prior to exercise	**Knee osteoarthritis** Mean age (SD): 55.5 (7.3) years N = 42 Definition: Diagnosis as per the American College of Rheumatology and radiological evidence of grade 3 or less on the Kellgren Lawrence scale Severity: Kellgren Lawrence grade 3 or less Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	
Bautch 1997[35]	**Supervised strength exercise** (n=17) Exercises including low intensity walking, and range of motion exercises of the trunk and upper and lower extremities, for 1 hour, three times per week	**Knee osteoarthritis** Mean age (SD): 67.8 (3.0) years N = 34	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Group session **No treatment** (n=17) No additional information **Concomitant therapy:** A weekly educational program with content related to health, exercise and arthritis	Definition: The American College of Rheumatology clinical and radiographic criteria for primary osteoarthritis of the knee Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear		
Borjesson 1996[57]	**Supervised strength exercise** (n=34) Physiotherapy aiming to increase the strength and range of motion of the knee, and strength of the whole leg, three times per week for 5 weeks. Home exercise was also completed twice per week Group or individual: Individual session **No treatment** (n=34) No intervention **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 64 (4.5) years N = 68 Definition: Medial knee osteoarthrosis grade 1-3 according to the classification based on weight-bearing radiographs Severity: Ahlback Osteoarthrosis grade 1-3, median grade 2 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months	
Bruce-Brand 2012[67]	**Supervised strength exercise** (n=14)	**Knee osteoarthritis**	Quality of life at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Resistance training for three home-based session per week for 6 weeks, two per week were supervised by a specialist. Group or individual: Group session **No treatment** (n=13) No intervention A third group (n=14) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review'. **Concomitant therapy:** Standard care was available to all including osteoarthritis education, weight loss, pharmacologic therapy and physical therapy	Mean age (SD): 64 (5.4) years N = 41 Definition: Symptomatic moderate to severe knee osteoarthritis confirmed radiographically as Kellgren Lawrence grade 3-4 Severity: Kellgren Lawrence grade 3-4 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at >3 months Physical function at >3 months	
Foley 2003[134]	**Other supervised exercise** (n=35) Hydrotherapy, including strengthening exercises. One set of 10 repetitions increased to three sets of 10-15, plus weighted gaiters, for 12 weeks Group or individual: Individual session Type of exercise: Hydrotherapy	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 70.9 (8.8) years N = 105 Definition: Radiological diagnosis of osteoarthritis of the hip, knee or both Severity: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Supervised strength exercise (n=35) Stationary cycling warm up, followed by strengthening exercise starting at the 10-repetition maximum or just below Group or individual: Individual session **No treatment** (n=35) Fortnightly telephone calls to record any changes in their condition, drug use, or injuries and free exercise treatment at the hospital at the end of the study period **Concomitant therapy:** Not stated/unclear	Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear		
Henriksen 2014[177] Subsidiary papers: Bartholdy 2016[34]	**Supervised strength exercise** (n=31) Bicycle ergometer warm up, followed by a circuit training program focusing on strength and coordination exercises of the trunk, hips and knees, three times per week for 12 weeks Group or individual: Group session **No treatment** (n=29) No attention control	**Knee osteoarthritis** Mean age (SD): 63.7 (8.2) years N = 60 Definition: Clinical diagnosis of tibiofemoral osteoarthritis confirmed by radiography Severity: Not stated/unclear Duration of symptoms: Not stated/unclear Presence of multimorbidities:	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: Not stated/unclear	Not stated/unclear		
Hermann 2016[178]	**Supervised strength exercise** (n=40) Progressive explosive-type resistance training programme, including a stationary bike warm-up, followed by four resistance training exercises, for 1 hour twice a week for 10 weeks Group or individual: Group session **No treatment** (n=40) Standardised preoperative information only (no attention control) **Concomitant therapy:** Not stated/unclear	**Hip osteoarthritis** Mean age (SD): 70.4 (7.6) years N = 80 Definition: People with primary hip osteoarthritis scheduled for total hip arthroplasty Severity: Not stated explicitly. On the waiting list for surgery Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	
Huang 2003[192]	**Supervised strength exercise** (n=99) Three groups: isokinetic exercise, isotonic exercise, and isometric exercise, three time per week for 8 weeks Group or individual: Individual session **No treatment** (n=33) No treatment	**Knee osteoarthritis** Mean age (SD): 62 (4.5) years N = 132 Definition: Moderate bilateral knee osteoarthritis (Altman grade II) Severity: Altman grade II	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: The people in all groups also received 20 minutes of hot packs and passive range motion exercise by an electric stationary bike (20 cycles per minute) for 5 minutes to both knees before exercise (unclear as to whether this applied to the control group)	Duration of symptoms (range): 4 months - 9 years Presence of multimorbidities: Not stated/unclear		
Huang 2005[193]	**Supervised strength exercise** (n=35) Isokinetic muscular strengthening exercises completed over 8 weeks Group or individual: Individual session **No treatment** (n=35) No treatment A third and fourth group (n=79) were reported but not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 65.0 (6.4) years N = 149 Definition: Bilateral moderate knee osteoarthritis (Altman grade 2) Severity: Altman grade II Duration of symptoms (range): 5 months – 12 years Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months	In Forest plots this study is referred to as Huang 2005A
Huang 2005[191]	**Supervised strength exercise** (n=30)	**Knee osteoarthritis** Mean age (SD): 62.0 (8.4) years N = 120	Pain at ≤3 months and >3 months	In Forest plots this study is referred to as Huang 2005B

Study	Intervention and comparison	Population	Outcomes	Comments
	Isokinetic muscular strengthening exercises completed over 8 weeks Group or individual: Individual session **No treatment** (n=30) No treatment A third and fourth group (n=60) were reported but not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** No additional information	Definition: Bilateral moderate knee osteoarthritis (Altman grade 2) with periarticular soft tissue pain, as identified by painful sensations during palpation or passive stretching of the arthritis knee under orthopedic examination. Confirmed by radiography. Severity: Altman grade II Duration of symptoms (range): 6 months – 11 years Presence of multimorbidities: Not stated/unclear		
Imoto 2012[203]	**Supervised strength exercise** (n=50) Muscle strengthening group activities based on the 10 maximum repetitions test, including stationary bike warmup, ischiotibial stretching and knee extension exercises. Twice a week for 30-40 minutes, for 8 weeks. Group or individual: Group session **No treatment** (n=50) An explanation about a manual after initial evaluation. The manual consisted of a	**Knee osteoarthritis** Mean age (SD): 60.1 (8.5) years N = 100 Definition: Knee osteoarthritis according to the criteria of the American College of Rheumatology Severity: Kellgren Lawrence grade 2-4, median grade 2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	description of knee osteoarthritis, as well as the possible signs and symptoms presented by the patients, and pointed them in the direction of a better way of dealing with the functional difficulties. **Concomitant therapy:** Not stated/unclear			
Jan 2008[206]	**Supervised strength exercise** (n=68) High resistance or low resistance exercise using the EN-Dynamic Track leg press machine, three training sessions per week for 8 weeks Group or individual: Individual session **No treatment** (n=30) No treatment control **Concomitant therapy:** People were not allowed to take non-steroidal anti-inflammatory medication during the study	**Knee osteoarthritis** Mean age (SD): 62.6 (6.7) years N = 98 Definition: Bilateral knee pain that fulfilled the American College of Rheumatology criteria for knee osteoarthritis Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 3.2 (2.7) years Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	
Jorge 2015[214]	**Supervised strength exercise** (n=29) Progressive resistance exercise programme including knee extension/flexion and hip abduction/ adduction, performed twice per week for 12 weeks	**Knee osteoarthritis** Mean age (SD): 60.8 (7.0) years N = 60 Definition: Unilateral or bilateral osteoarthritis of the	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Individual session **No treatment** (n=31) No exercise control **Concomitant therapy:** When pain exceeded a 7 on the visual analogue scale, the subject could take 50mg of diclofenac every 8 hours.	knee, based on the classification criteria of the American College of Rheumatology Severity: Radiographic grade 1-2 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear		
Kang 2019[221]	**Supervised strength exercise** (n=15) Finger exercise programme, to maintain or increase the flexibility of the MCP, PIP and DIP joints, to increase opponens pollicis strength and grip strength, and to strengthen the extensor and abductor pollicis muscles Group or individual: Individual session **No treatment** (n=14) No additional treatment **Concomitant therapy:** Both groups received dip-wrap paraffin bath therapy sessions. Subjects dipped the affect hand in, removed the hand, and waited for the paraffin to harden and become opaque. They then	**Hand osteoarthritis** Mean age (SD): 47.3 (4.4) years N = 29 Definition: Hand osteoarthritis as suggested by the American College of Rheumatology Severity: Not stated Duration of symptoms (mean [SD]): 3.5 (1.1) years Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	re-dipped the hand up to 10 times. When the last layer hardened, the hand was covered with a towel for 20 minutes.			
Kigozi 2018[234]	**Supervised strength exercise** (n=176) Progressive lower limb exercise programme with 6-8 sessions over 12 weeks. Participants received a print-out of a specific exercise prescription individualised for them based on their progress Group or individual: Individual session **No treatment** (n=175) Usual care only A third group (n=163) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** All participants received a booklet providing information about benefits of exercise and physical activity and a home exercise program. Usual physical therapy care included advice and lower-limb exercise provided in up to four individual	**Knee osteoarthritis** Mean age: 63 years N = 514 Definition: Current knee pain and/or stiffness in one or both knees who met the criteria recommended by the National Institute for Health and Care Excellence guidelines for a clinical diagnosis of knee osteoarthritis Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	one-to-one treatment sessions over 12 weeks, in line with usual physical therapy practice in the National Health Service			
Kuptniratsaikul 2002[250]	**Supervised strength exercise** (n=193) An exercise class, emphasising quadriceps muscle strengthening, for two sessions per week lasting 1 hour, for 8 weeks Group or individual: Not stated/unclear **No treatment** (n=199) No exercise **Concomitant therapy:** People were allowed to continue their usual medical treatments	**Knee osteoarthritis** Mean age (SD): 67.8 (5.9) years N = 392 Definition: Osteoarthritis of the knee, Kellgren Lawrence grade 2-3 Severity: Median Mild Duration of symptoms (mean [SD]): 45.3 (40.2) months Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	
Lin 2009[266]	**Other supervised exercise** (n=36) Computer game foot-stepping exercises predominantly involving knee movement in a sitting position with a 150-250N force applied to the foot. Training for 20 minutes for each leg 3 sessions per week for 8 weeks. Group or individual: Individual session Type of exercise: Proprioception	**Knee osteoarthritis** Mean age (SD): 62.5 (7.5) years N = 108 Definition: Osteoarthritis diagnosed by an orthopaedic surgeon based on the clinical history, radiographic imaging and physical assessment Severity: Radiographic median grade 3	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Supervised strength exercise (n=36) Quadriceps exercises completed with dynamometer cables and weights to increase resistance. 3 sessions per week with 4 sets (6 repetitions per set) for 8 weeks. Group or individual: Individual session **No treatment** (n=36) **Concomitant therapy:** All people were asked to cease any exercise activity outside of the exercise training	Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Nahayatbin 2018[314]	**Other supervised exercise** (n=16) 10 minutes of Yang style Tai Chi with 5 minutes of warm up and cool down. Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=16) Closed chain kinetic exercises with10 minutes of exercise and 5 minutes of warm up and cool down. Group or individual: Group session	**Knee osteoarthritis** Mean age (SD): 55.89 (5.97) years N = 48 Definition: Knee osteoarthritis with grade 2-3 changes based on the Kellgren Lawrence classification Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=16) **Concomitant therapy:** All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound			
Nejati 2015[320]	**Supervised strength exercise** (n=28) Strengthening and stretching exercises for the muscles of the knee performed daily with cuff weights for resistance. Completed over 3 months Group or individual: Individual session **No treatment** (n=28) **Concomitant therapy:** In both groups people received acupuncture during 10 sessions, twice per week, physical modalities during 1- sessions, three times a week (including TENS, ultrasound and infrared) and could receive diclofenac 100mg once daily for pain. All people were recommended to use 1500mg glucosamine and 800mg chondroitin.	**Knee osteoarthritis** Mean age (SD): 61.3 (9.2) years N = 56 Definition: Knee osteoarthritis according to the American College of Rheumatology criteria with radiographic Kellgren Lawrence grade 2-4 changes Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	
Nery, 2021[322]	**Supervised strength exercise** (n=30)	**Hand osteoarthritis**	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Participants completed a programme of progressive resistance exercises in two sessions per week over twelve weeks. Group or individual: Individual session 3. Type of exercise: Not applicable **No treatment** (n=30) No additional treatment. Duration 12 weeks **Concomitant therapy:** Both groups had a single education session to receive information about the illness before the randomisation. This briefing included information about the disease and the impairment it caused, treatment and guidelines for joint protection and energy conservation. The participants did not receive any extra material. Both groups were instructed to continue medication without change and orthoses were not allowed during the study.	Mean age (SD): 66.8 (9.1) years N = 60 Definition: Hand osteoarthritis under the American College of Rheumatology classification criteria Severity: Kellgren Lawrence grades I-IV, median grade III. Duration of symptoms (SD): 7.1 (5.1) years. Presence of multimorbidities: Not stated/unclear		
Oliveira 2012[332]	**Supervised strength exercise** (n=50) Strength exercise performed twice a week over 8 weeks Group or individual: Individual session	**Knee osteoarthritis** Mean age (SD): 60.1 (8.5) years N = 100	Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=50) **Concomitant therapy:** Both groups received a manual with instructions to prevent knee overload during daily activities and instructions about the use of knee ice packs for pain with inflammation, and warm dressing for pain with no inflammatory signs. In addition, people in both groups were already prescribed medication.	Definition: Knee osteoarthritis diagnosed according to the American College of Rheumatology criteria Severity: Kellgren Lawrence grade 2-4, median grade 2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Park 2021[347]	**Supervised strength exercise** (n=27) Isometric exercise group. Eight types of isometric movements were performed during the impulse phase as per the instructor's direction. Duration 8 weeks. Group or individual: Individual session Type of exercise: Not applicable **No treatment** (n=27) No treatment control. Duration 8 weeks. A third group (n=27) did not fulfil the inclusion criteria in the protocol for this review and so was not included in the analysis.	**Knee osteoarthritis** Mean age (SD): 66.9 (4.2) years N = 81 Definition: Degenerative knee osteoarthritis diagnosed by bilateral radiographic examination Severity: Kellgren Lawrence grades I or II Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: No additional information.			
Pazit 2018[349]	**Supervised mixed modality exercise** (n=10) High speed resistance training with balance training for 8 weeks. Group or individual: Group session Type of exercise: Strength and balance **Supervised strength exercise** (n=10) High speed resistance training for 8 weeks. Group or individual: Group session **No treatment** (n=10) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 67.68 (6.68) years N = 28 Definition: Knee osteoarthritis based on the presence of clinical symptoms of knee osteoarthritis as defined by the American College of Rheumatology criteria Severity: Not stated Duration of symptoms: At least 6 months Presence of multimorbidities: High morbidity score	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Rezasoltani 2020[376]	**Supervised strength exercise** (n=16) Aquatic cycling exercise, 3 sessions per week for 4 weeks totalling 12 sessions guided by a physiotherapist certified in aquatic physiotherapy.	**Knee osteoarthritis** Mean age (SD): 51.0 (2.93) years N = 30	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Individual session Type of exercise: Hydrotherapy **No treatment** (n=16) No additional treatment. Duration 4 weeks. **Concomitant therapy:** People were instructed to use paracetamol if needed and to follow lifestyle recommendations to use their knees more appropriately.	Definition: Knee osteoarthritis with knee pain for at least 3 months Severity: Not stated/unclear Duration of symptoms: At least 3 months. Presence of multimorbidities: Not stated/unclear		
Rosedale 2014[384]	**Supervised strength exercise** (n=120) Specific strength based exercises with advice on aerobic exercise. Exercises prescribed as 10 repetitions every 2 to 3 hours for 3 months Group or individual: Individual session **No treatment** (n=60) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 65.3 (10.4) years N = 180 Definition: People with knee pain for greater than 4 months and radiologically confirmed diagnosis of knee osteoarthritis Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: High morbidity score	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Salli 2010[388]	**Supervised strength exercise** (n=47) Concentric-eccentric type isokinetic exercises or isometric exercises performed 3 days a week for 8 weeks Group or individual: Group session **No treatment** (n=24) **Concomitant therapy:** People in all groups received 500mg paracetamol tablets as required, up to 3 grams per day	**Knee osteoarthritis** Mean age (SD): 57.06 (7.31) years N = 71 Definition: People with clinically and radiologically diagnosed osteoarthritis in both knees according to the American College of Rheumatology criteria Severity: Kellgren Lawrence grade 1-2 Duration of symptoms: Not stated Presence of multimorbidities: Low morbidity score	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	
Samut 2015[390]	**Supervised strength exercise** (n=15) Isokinetic exercise performed 3 days/week for 6 weeks using the Biodex isokinetic system. 5 concentric flexion and extension at angular velocities of 60-180 degrees/s with one set of contractions increased to 6 sets by the end of the study. Group or individual: Individual	**Knee osteoarthritis** Mean age (SD): 60.4 (7.8) years N = 42 Definition: Kellgren Lawrence grade 2-3 knee osteoarthritis fulfilling the American College of Rheumatology criteria Severity: Kellgren Lawrence grade 2-3	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Supervised aerobic exercise (n=14) Aerobic exercise performed 3 days/week for 6 weeks on a treadmill. Exercise intensity was adjusted for 65-70% of age-related heart rate for the first 4 weeks and 70-75% for the next 2 weeks. Group or individual: Individual **No treatment** (n=13) **Concomitant therapy:** Each person was allowed to take paracetamol whenever needed	Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear		
Sayers 2012[393]	**Supervised strength exercise** (n=30) Power therapy (including low intensity and explosive high intensity) performed 3 times a week for 12 weeks Group or individual: Group session **No treatment** (n=15) Stretches only **Concomitant therapy:** The exercises started with 12 stretches including the back, trunk and lower extremity	**Knee osteoarthritis** Mean age (SD): 67.1 (7.3) years N = 45 Definition: Knee osteoarthritis according to the American College of Rheumatology clinical classification Severity: Kellgren Lawrence mean grade 1-2 Duration of symptoms: Not stated	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	stretches. Each stretch was initiated and held for 30 seconds by a trained physical therapy study research assistant. Following the stretching protocol, a 5 minute warm up on a cycle ergometer was performed before starting the exercise	Presence of multimorbidities: Not stated/unclear		
Thorstensson 2005[444]	**Supervised strength exercise** (n=30) Physical therapist led exercise program (one-hour sessions, twice a week for 6 weeks). Group or individual: Group session **No treatment** (n=31) No change to usual care until the end of the follow up period. **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 56.1 (6.1) years N = 61 Definition: Radiographic diagnosis with a Kellgren and Lawrence grade of 3 or more. Severity: Kellgren Lawrence grade at least 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	
Williamson 2007[485]	**Supervised strength exercise** (n=60) Strength exercise circuit conducted once a week for 6 weeks Group or individual: Group session **No treatment** (n=61) Exercise and advice leaflet only	**Knee osteoarthritis** Mean age (SD): 70.6 (9.0) years N = 181 Definition: People listed for knee arthroplasty with osteoarthritis of the knee	Pain at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	A third group (n=60) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** No additional information	Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Wortley 2013[489]	**Other supervised exercise** (n=15) 1 hour group training session twice a week based on the 12 basic movements adapted from the Yang Style Tai Ji for 10 weeks Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=15) Resistance training program consisting of two 1-hour sessions per week using ankle cuff weights for resistance. Completed over 10 weeks Group or individual: Group session **No treatment** (n=9) **Concomitant therapy:**	**Knee osteoarthritis** Mean age (SD): 69.2 (6.0) years N = 39 Definition: The Classification Criteria for Knee OA of the American College of Rheumatology and bilateral knee x-rays Severity: Kellgren Lawrence median grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Participants were asked not to alter their regular physical activity or pain medications			

1.1.5.4 Unsupervised strength exercise compared to unsupervised aerobic exercise

Table 5: Summary of studies included in the evidence review for the unsupervised strength exercise compared to unsupervised aerobic exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Evcik 2002[122]	**Unsupervised strength exercise** (n=27) Home-based exercise program including isometric and isotonic quadriceps exercises. Each exercise was done ten times, twice a day at home Group or individual: Individual session **Unsupervised aerobic exercise** (n=28) Regular walking program, started at 10 minutes, three times per week. Gradually increased up to half an hour Group or individual: Individual session **No treatment** (n=26) Continued normal daily activities, no extra exercise or regular walking programs	**Knee osteoarthritis** Mean age (SD): 56.4 (6.5) years N = 81 Definition: Knee osteoarthritis with Kellgren and Lawrence grade 1-3 changes Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 8.0 (3.3) years Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: All people were allowed to take analgesic drugs (paracetamol)			

1.1.5.5 Unsupervised strength exercise compared to no treatment

Table 6: Summary of studies included in the evidence review for the unsupervised strength exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Bennell 2010[45] Subsidiary papers: Bennell 2007[44]	**Unsupervised strength exercise** (n=45) Home-based hip strengthening exercises with ankle weights or elastic bands, five times per week for 12 weeks. Participants also attended a physiotherapy clinic Group or individual: Individual session **No treatment** (n=44) No additional exercise treatment **Concomitant therapy:** Participants were asked to refrain from seeking other forms of treatment, although analgesia and non-steroidal anti-inflammatory drugs were permitted as required	**Hip osteoarthritis** Mean age (SD): 64.6 (8.4) years N = 89 Definition: The American College of Rheumatology classification criteria with radiographic verification Severity: Kellgren and Lawrence grades 2-4 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Chang 2012[79]	**Unsupervised strength exercise** (n=30)	**Knee osteoarthritis** Mean age (SD): 67.4 (8.9) years	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Elastic band leg press exercises, 10 repetitions per set, 3 set per day, 2 days per week. Advancement to higher intensity every 2 weeks Group or individual: Individual session **No treatment** (n=30) Conventional modality treatments only **Concomitant therapy:** Conventional modality treatments available to everyone included shortwave diathermy, hot packs, transcutaneous electrical nerve stimulation, interferential current and "so on".	N = 60 Definition: Diagnosis of knee osteoarthritis no less than Kellgren-Lawrence grade 3 Severity: Kellgren and Lawrence grades 2-3 (majority grade 3) Duration of symptoms (mean [SD]): 9.0 (7.6) months Presence of multimorbidities: Not stated/unclear		
Chen 2019[81]	**Unsupervised strength exercise** (n=84) Home-based strengthening exercise over 12 weeks with 4 weeks of physiotherapy training in exercise and health education. Exercise prescription was 30-40 minutes per day, at least 3 days per week Group or individual: Group session **No treatment** (n=87)	**Knee osteoarthritis** Mean age: 68.9 years N = 171 Definition: Previously diagnosed with knee osteoarthritis with knee pain Severity: Not stated Duration of symptoms (mean): 6.4 years	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Health education without any reference to exercise **Concomitant therapy:** Health education was available to both groups (with the control group not receiving any education regarding exercise)	Presence of multimorbidities: Not stated/unclear		
Dziedzic 2015[114] Subsidiary papers: Dziedzic 2011[115] Oppong 2014[333]	**Unsupervised strength exercise** (n=65) Strengthening exercises with stretches with 10 repetitions of each exercise daily for 12 months **No treatment** (n=65) Leaflet only Two other groups were reported in the study (n=127) but were not included in the analysis as they did not fulfil the inclusion criteria **Concomitant therapy:** Leaflet and advice - all participants were given standardised written information on self-management approaches for hand osteoarthritis including general information only looking after hand joints, and using analgesia.	Hand osteoarthritis Mean age (SD): 65.8 (9.1) N = 257 Definition: People meeting the criteria for hand osteoarthritis according to the American College of Rheumatology criteria or had unilateral or bilateral thumb base osteoarthritis Severity: Not stated Duration of symptoms: Not stated Multimorbidity: Not stated/unclear	Quality of life at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Evcik 2002[122]	**Unsupervised strength exercise** (n=27) Home-based exercise program including isometric and isotonic quadriceps exercises. Each exercise was done ten times, twice a day at home Group or individual: Individual session **Unsupervised aerobic exercise** (n=28) Regular walking program, started at 10 minutes, three times per week. Gradually increased up to half an hour Group or individual: Individual session **No treatment** (n=26) Continued normal daily activities, no extra exercise or regular walking programs **Concomitant therapy:** All people were allowed to take analgesic drugs (paracetamol)	**Knee osteoarthritis** Mean age (SD): 56.4 (6.5) years N = 81 Definition: Knee osteoarthritis with Kellgren and Lawrence grade 1-3 changes Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 8.0 (3.3) years Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	
Hennig 2015[173]	**Unsupervised strength exercise** (n=40) Home-based hand exercise programme aimed at maximising the stable and pain-free functional range of motion of the finger joints, increasing	**Hand osteoarthritis** Mean age (SD): 60.8 (7.0) years N = 80 Definition:	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	grip strength, maintaining joint stability and preventing or delaying fixed deformities Group or individual: Individual session **No treatment** (n=40) Leaflet only **Concomitant therapy:** All participants received a leaflet containing information about hand osteoarthritis and ergonomic principles	Hand osteoarthritis diagnosed by rheumatologists or orthopaedic surgeons according to the American College of Rheumatology criteria Severity: Not stated Duration of symptoms (median [range]): 10.0 (0-40) years Presence of multimorbidities: High morbidity score (32 people had comorbidities)		
Juhakoski 2011[216]	**Unsupervised strength exercise** (n=60) Home exercise programme of hip strengthening exercises, conducted over around 30-35 minutes where exercises were made with the maximal effort to achieve the highest possible movement velocity, 12 sessions once per week Group or individual: Individual session **No treatment** (n=58) No exercise treatment **Concomitant therapy:**	**Hip osteoarthritis** Mean age (SD): 66.6 (6.5) years N = 120 Definition: People with unilateral or bilateral hip osteoarthritis fulfilling the clinical and radiological criteria of the American College of Rheumatology Severity: Radiological grade 1-4, median grade 2 Duration of symptoms: Not stated	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	All people received an hour-long instruction session regarding the basic principles of non-operative treatment for hip osteoarthritis. All people received GP standard care.	Presence of multimorbidities: High morbidity score (no chronic disease = 49, 1 chronic disease = 53, 2 or more chronic diseases = 16)		
Karadag 2019[222]	**Unsupervised strength exercise** (n=36) Combination of exercise after heat application and exercise only group. Practiced twice a day for 5 days a week for 4 weeks. Group or individual: Individual session Type of exercise: Not applicable **No treatment** (n=36) Combination of the heat pack and control group. The heat pack group received two hot-packs to be applied to both knees and were recommended to use them for 20 minutes, twice a day for 5 days a week. **Concomitant therapy:** No additional information.	**Knee osteoarthritis** Mean age (SD): 57.9 (10.8) years N = 62 Definition: People with a diagnosis of knee osteoarthritis according to the American College of Rheumatology criteria Severity: Stage 2-4, median stage 3 Duration of symptoms (SD): 32.4 (6.6) years Presence of multimorbidities: High morbidity score (People with any other chronic disease: 39. People without chronic disease: 23.).	Pain at ≤3 months Physical function at ≤3 months	
Lim 2008[264]	**Unsupervised strength exercise** (n=53)	**Knee osteoarthritis** Mean age (SD): 64.6 (8.6) years N = 107	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Quadriceps muscle strength exercise completed 5 days a week for 12 weeks Group or individual: Individual session **No treatment** (n=54) **Concomitant therapy:** No additional information	Definition: Tibiofemoral joint osteoarthritis in at least 1 knee fulfilling the American College of Rheumatology criteria Severity: Median Kellgren Lawrence grade 3 Duration of symptoms (mean [SD]): 6.7 (6.5) years Presence of multimorbidities: Not stated/unclear		
O'Reilly 1999[328]	**Unsupervised strength exercise** (n=113) Graded exercise program increased to a maximum of 20 repetitions on each leg for 6 months Group or individual: Individual **No treatment** (n=78) **Concomitant therapy:** General verbal advice concerning knee pain and knee osteoarthritis with advice on the importance of losing weight or not becoming overweight, wearing training shoes/air filled soles and maintaining fitness by walking or swimming was given to all participants	**Knee osteoarthritis** Mean age (SD): 62.05 (9.87) years N = 191 Definition: People with knee pain who responded affirmatively to both parts of the following questions "Have you ever had pain in or around the knee on most days for at least a month? If so, have you experienced any pain during the last year" who were then further assessed Severity: Not stated Duration of symptoms: At least 1 year Presence of multimorbidities: Not stated/unclear	Quality of life at >3 months Pain at >3 months Physical function at >3 months Psychological distress at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Osteras 2014[337]	**Unsupervised strength exercise** (n=65) Exercise program focussing on strength performed 3 times weekly as 1 set of 10 repetitions in weeks 1-2 and 15 repetitions in weeks 3-12. Group or individual: Individual **No treatment** (n=65) **Concomitant therapy:** Usual care included visits from general practitioners only, and very infrequently a referral to a consultation with an occupational therapist in secondary care	**Hand osteoarthritis** Mean age (SD): 66 (8.6) years N = 130 Definition: Hand osteoarthritis meeting the American College of Rheumatology criteria for features of hand osteoarthritis or uni/bilateral osteoarthritis of the first carpometacarpal joint, and a Functional Index for Hand Osteoarthritis score of at least 5 Severity: Not stated Duration of symptoms (year [SD]): 11.5 (8.1) years Presence of multimorbidities: Low morbidity score	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months Serious adverse events at >3 months	
Ravaud 2004[369]	**Unsupervised strength exercise** (n=735) Exercise taught by verbal instruction and a videotape presentation with a 30 minute program increasing from 10 repetitions up to a maximum of 30, 4 times a week for 6 months Group or individual: Individual session **No treatment** (n=760)	**Knee osteoarthritis** Mean age (SD): 66.78 (10.39) years N = 2957 Definition: People who met the clinical and radiographic American College of Rheumatology criteria for osteoarthritis of the knee or hip	Pain at >3 months Physical function at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	A third group (n=1462) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** All people received the non-steroidal anti-inflammatory drug rofecoxib. The drug was administered once daily at 12.5mg during the first month and thereafter at 25mg if necessary. People were permitted to take paracetamol if necessary.	Severity: Kellgren and Lawrence grade 2-4, median grade 3 Duration of symptoms (mean [SD]) 69.5 (75.5) months Presence of multimorbidities: Not stated/unclear		

1.1.5.6 Supervised aerobic exercise compared to no treatment

Table 7: Summary of studies included in the evidence review for the supervised aerobic exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Brosseau 2012[66]	**Supervised aerobic exercise** (n=79) Walking with an aim of achieving an intensity of 50-75% maximum heart rate. Consisted of a progressive aerobic phase and a maintenance phase. Three weekly sessions for 12 months Group or individual: Group session	**Knee osteoarthritis** Mean age (SD): 63.4 (8.6) years N = 222 Definition: The American College of Rheumatology clinical and radiographic/magnetic resonance imaging criteria	Quality of life at >3 months Pain at >3 months Physical function at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=74) No walking intervention A third group (n=69) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** All participants received an educational pamphlet describing the benefits of walking, and a pedometer	Severity: Mild to moderate Duration of symptoms (mean [SD]): 10.3 (9.26) years Presence of multimorbidities: Not stated/unclear		
Christensen 2015[89]	**Supervised aerobic exercise** (n=64) Exercise programme including a warm up phase, a circuit training phase and a cool down phase, three times per week for 52 weeks Group or individual: Not stated/unclear **No treatment** (n=64) No attention control A third group (n=64) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review.	**Knee osteoarthritis** Mean age (SD): 62.4 (6.4) years N = 192 Definition: Knee osteoarthritis confirmed by clinical symptoms, including pain, and on standing radiographs in at least 1 joint compartment Severity: Kellgren Lawrence grade I-IV, median grade III Duration of symptoms (median [IQR]): Control = 8.0 (4.5-13.0), diet = 8.0 (3.8-	Quality of life at >3 months Pain at >3 months Physical function at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: All participants had a 12-week period prior to the studies where they had intensive weight loss before being assigned to the groups	10.0), exercise = 9.5 (4.8-15.0) Presence of multimorbidities: Not stated/unclear		
Samut 2015[390]	**Supervised strength exercise** (n=15) Isokinetic exercise performed 3 days/week for 6 weeks using the Biodex isokinetic system. 5 concentric flexion and extension at angular velocities of 60-180 degrees/s with one set of contractions increased to 6 sets by the end of the study. Group or individual: Individual **Supervised aerobic exercise** (n=14) Aerobic exercise performed 3 days/week for 6 weeks on a treadmill. Exercise intensity was adjusted for 65-70% of age-related heart rate for the first 4 weeks and 70-75% for the next 2 weeks. Group or individual: Individual **No treatment** (n=13)	**Knee osteoarthritis** Mean age (SD): 60.4 (7.8) years N = 42 Definition: Kellgren Lawrence grade 2-3 knee osteoarthritis fulfilling the American College of Rheumatology criteria Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: Each person was allowed to take paracetamol whenever needed			
Salacinski 2012[387]	**Supervised aerobic exercise** (n=19) Facility based cycling exercises conducted as 2 supervised group sessions per week for 12 weeks Group or individual: Group session **No treatment** (n=18) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 57.7 (9.8) years N = 37 Definition: Mild-to-moderate osteoarthritis of the knee with grades 1-3 Kellgren Lawrence changes on radiography Severity: Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

1.1.5.6 Supervised aerobic exercise compared to other supervised exercise

Table 8: Summary of studies included in the evidence review for the supervised aerobic exercise compared to no other supervised exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Bokaeian 2021[56]	**Other supervised exercise** (n=22) YogaMT group who practiced MT gait for 20 minutes and performed three repetitions of	**Knee osteoarthritis** Mean age (SD): 56.1 (5.0) years N = 59	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	goddess squat and warrior lunge exercises with a minimum rest interval of 40s, supervised by a physiotherapist. For the MT gait, they were trained to walk with slight hip internal rotation and knee flexion (about 20 degrees) at their selected speed on the treadmill. Duration 4 weeks. Group or individual: Individual session Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) **Supervised strength exercise** (n=19) The KMS group that consisted of three 10-repetition sets of resistive knee extension/flexion exercises with 2-min interval, using the quadriceps chair. Exercise was performed under supervision of a physiotherapist. Duration 4 weeks Group or individual : Individual session Type of exercise: Not applicable **Supervised aerobic exercise** (n=18) The TMW group that practiced walking on a treadmill at their self-selected speed for 20 minutes, with no gait	Definition: People with unilateral or bilateral knee osteoarthritis diagnosed radiographically and by symptoms Severity: Kellgren Lawrence grades 2-3 Duration of symptoms: history of pain for greater than one month Presence of multimorbidities: Not applicable		

Study	Intervention and comparison	Population	Outcomes	Comments
	modification, under the supervision of the physiotherapist. Duration 4 weeks. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification was also offered.			

1.1.5.7 Unsupervised aerobic exercise compared to no treatment

Table 9: Summary of studies included in the evidence review for the unsupervised aerobic exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Bossen 2013[59] Subsidiary papers: Bossen 2013[58]	**Unsupervised aerobic exercise** (n=100) Joint2move website-based exercise programme using exercise activities that a person enjoys (e.g. cycling) and making goals. The website also provided information about	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 62.0 (5.7) years N = 199 Definition:	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical functioning at ≤3 months and >3 months Psychological distress at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	osteoarthritis and lifestyle choices Group or individual: Individual session **No treatment** (n=99) Waiting list control **Concomitant therapy:** No additional information	Self-reported knee and/or hip osteoarthritis - defined by if they had a painful knee or hip joint and if a doctor or other health care provider had ever told them this was a result of osteoarthritis Severity: Not stated Duration of symptoms: <1 to >7 years - median >3-7 years Presence of multimorbidities: Low morbidity score (Majority had no comorbidities (125) while 35 had one comorbidity and 39 had two comorbidities)		
Evcik 2002[122]	**Unsupervised strength exercise** (n=27) Home-based exercise program including isometric and isotonic quadriceps exercises. Each exercise was done ten times, twice a day at home Group or individual: Individual session **Unsupervised aerobic exercise** (n=28) Regular walking program, started at 10 minutes, three	**Knee osteoarthritis** Mean age (SD): 56.4 (6.5) years N = 81 Definition: Knee osteoarthritis with Kellgren and Lawrence grade 1-3 changes Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]):	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	times per week. Gradually increased up to half an hour Group or individual: Individual session **No treatment** (n=26) Continued normal daily activities, no extra exercise or regular walking programs **Concomitant therapy:** All people were allowed to take analgesic drugs (paracetamol)	8.0 (3.3) years Presence of multimorbidities: Not stated/unclear		
Shahine 2020[405]	**Unsupervised aerobic exercise** (n=33) Routine care, educational sessions about pedometer self monitoring, aerobic weekly step count goals and weekly telephone follow up. People were given an individualised step count goal every week to gradually increase by 10% of baseline steps/day for weeks 2-12. Duration 12 weeks Group or individual: Individual session Type of exercise: Not applicable **No treatment** (n=33) Usual routine care only. Duration 12 weeks	**Knee osteoarthritis** Mean age (mean, SD): 66.2 (5.5) years N = 66 Definition: Diagnosed with knee osteoarthritis for at least 1 year Severity: Not stated/unclear Duration of symptoms: At least one year. Between 1 year and 10+ years, median 5-10 years. Presence of multimorbidities: High morbidity score (No comorbidities = 19. One comorbidity = 30. Two comorbidities = 14. More than 3 = 3.).	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: Everyone received a booklet as a disease guide and a pedometer.			

1.1.5.8 Other supervised exercise compared to supervised strength exercise

Table 10: Summary of studies included in the evidence review for the other supervised exercise compared to supervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Avelar 2011[29]	**Other supervised exercise** (n=11) Whole body vibration while performing squat exercises. Squat exercise progressed in time and repetitions over the study. Group or individual: Individual session Type of exercise: Whole body vibration and strengthening **Supervised strength exercise** (n=10) Squat exercises only, with progressive increasing time and repetitions over the study Group or individual: Individual sessions **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 73.1 (5.0) years N = 21 Definition: Clinical and radiographic criteria of the American College of Rheumatology with a classification of Kellgren and Lawrence grade 1-4 Severity: Kellgren Lawrence grade 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	As whole body vibration was delivered at the same time as exercise the committee agreed this should be classed as other supervised exercise

Study	Intervention and comparison	Population	Outcomes	Comments
Bennell 2014[46]	**Other supervised exercise** (n=50) Neuromuscular exercise to improve position of the trunk and lower limb joint while dynamically and functionally strengthening the lower limb, 14 times over 12 weeks. Home exercise was also performed 3 times per week Group or individual: Individual session Type of exercise: Neuromodulatory **Supervised strength exercise** (n=50) Strengthening exercises focusing on the quadriceps with non-weight bearing exercises, 14 times over 12 weeks. Home exercise was also performed 3 times per week Group or individual: Individual session **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 62.5 (7.4) years N = 100 Definition: Knee pain with radiographic medial tibiofemoral joint osteoarthritis Severity: Kellgren Lawrence grade no less than 2 Duration of symptoms (median [IQR]): Neuromuscular = 60.0 (96.0) months, strength = 84.0 (96.3) months Presence of multimorbidities: Not stated/unclear	Quality of life at >3 months Pain at >3 months Physical function at >3 months Serious adverse events at >3 months	
Bokaeian 2021[56]	**Other supervised exercise** (n=22) YogaMT group who practiced MT gait for 20 minutes and performed three repetitions of goddess squat and warrior	**Knee osteoarthritis** Mean age (SD): 56.1 (5.0) N = 59 Definition: People with unilateral or bilateral knee	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	lunge exercises with a minimum rest interval of 40s, supervised by a physiotherapist. For the MT gait, they were trained to walk with slight hip internal rotation and knee flexion (about 20 degrees) at their selected speed on the treadmill. Duration 4 weeks. **Supervised strength exercise** (n=19) The KMS group that consisted of three 10-repetition sets of resistive knee extension/flexion exercises with 2-min interval, using the quadriceps chair. The maximum load that each participant could lift to complete 10-repetition maximum without pain was determined to adjust the amount of load for each exercise weekly. Exercise was performed under supervision of a physiotherapist. Duration 4 weeks. **Supervised aerobic exercise** (n=18) The TMW group that practiced walking on a treadmill at their self-selected speed for 20 minutes, with no gait modification, under the supervision of the	osteoarthritis diagnosed radiographically and by symptoms Severity: Kellgren Lawrence grades 2-3 Duration of symptoms: history of pain for greater than one month Presence of multimorbidities: Not applicable		

Study	Intervention and comparison	Population	Outcomes	Comments
	physiotherapist. Duration 4 weeks. **Concomitant therapy:** People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification was also offered.			
Ebnezar 2011[116] Subsidiary papers: Ebnezar 2012[117] Ebnezar 2012[118]	**Other supervised exercise** (n=125) Integrated yoga including shithilikaranavyayama (loosening and strengthening), asanas, relaxation techniques, pranayama, meditation and didactic lectures Group or individual: Group session Type of exercise: Yoga **Supervised strength exercise** (n=125) Strengthening exercises focusing on the quadriceps with non-weight bearing exercises, 14 times over 12 weeks. Home exercise was also performed 3 times per week Group or individual:	**Knee osteoarthritis** Mean age (SD): 59.5 (9.5) years N = 250 Definition: The American College of Rheumatology criteria for the diagnosis of osteoarthritis of the knee Severity: Kellgren Lawrence grade 2-4 Duration of symptoms (median): 1-2 years Presence of multimorbidities: High morbidity score (38 had diabetes, 49 had hypertension, 171 were overweight /obese, 145 had	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Individual session **Concomitant therapy**: No additional information	osteoporosis and 56 had other diseases)		
Foley 2003[134]	**Other supervised exercise** (n=35) Hydrotherapy, including strengthening exercises. One set of 10 repetitions increased to three sets of 10-15, plus weighted gaiters, for 12 weeks Group or individual: Individual session Type of exercise: Hydrotherapy **Supervised strength exercise** (n=35) Stationary cycling warm up, followed by strengthening exercise starting at the 10-repetition maximum or just below Group or individual: Individual session **Concomitant therapy:** No additional information	**Hip and/or knee osteoarthritis** Mean age (SD): 70.9 (8.8) years N = 105 Definition: Radiological diagnosis of osteoarthritis of the hip or knee, or both Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	
Gill 2009[153]	**Other supervised exercise** (n=42) Water-based exercise of moderate intensity, including walking and active range of	**Hip or knee osteoarthritis** Mean age (SD): 70.4 (9.8) years N = 82	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	motion exercise and stretching, twice per week for 6 weeks Group or individual: Group session Type of exercise: Hydrotherapy **Supervised strength exercise** (n=40) Land-based exercises of moderate intensity, including forward, sideways and backward walking, stationary exercise bike, resistance exercises and stretching, twice a week for 6 weeks Group or individual: Group session **Concomitant therapy:** All participants were asked to complete 30 minutes of exercise at home, 3 times a week, including walking, riding a stationary bike, or performing exercises similar to those completed in the supervised classes. After the intervention, participants continued to exercise at home until follow-up. All people received a home visit and environmental assessment from an occupational therapist	Definition: People on the waiting list for joint replacement surgery or the hip or knee Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Physical function at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Gomiero 2018[157]	**Other supervised exercise** (n=32) Sensory-motor training, including walking in different directions, crossing steps whilst walking, changing direction, walking on different surfaces, maintaining posture using a balance board and using a mini-trampoline, twice a week for 16 weeks Group or individual: Group session Type of exercise: Neuromodulatory **Supervised strength exercise** (n=32) Strengthening exercises including isometric exercises for quadriceps, stretching for lower limbs, and use of ankle weights, twice a week for 16 weeks Group or individual: Group session Concomitant therapy: Both groups had concomitant interventions such as informative talks. They also received an educational program on knee osteoarthritis, which allowed the people to	**Knee osteoarthritis** Mean age (SD): 61.7 (6.6) years N = 64 Definition: Knee osteoarthritis of the American College of Rheumatology with radiographic confirmation Severity: Kellgren and Lawrence grade 1-4, median grade 2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at >3 months Pain at >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	clarify their doubts and concerns about the disease			
Khruakhorn 2021[233]	**Other supervised exercise** (n=17) Hydrotherapy with progressive strengthening exercises. Both groups attended the exercise classes for 45-60 minutes, three times per week for 6 weeks. Group or individual: Individual session Type of exercise: Hydrotherapy **Supervised strength exercise** (n=17) Land based therapy with progressive strengthening exercises. Both groups attended the exercise classes for 45-60 minutes, three times per week for 6 weeks. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** No additional information.	**Knee osteoarthritis** Mean age (SD): 61.4 (8.4) years N = 34 Definition: Osteoarthritis of the knee diagnosed with radiography Severity: Kellgren-Lawrence grade 2-3 Duration of symptoms: Not stated/unclear. Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months and >3 months Pain at ≤3 months Physical function at ≤3 months	
Lin 2009[266]	**Other supervised exercise** (n=36) Computer game foot-stepping exercises predominantly involving knee movement in a sitting position with a 150-250N force applied to the foot.	**Knee osteoarthritis** Mean age (SD): 62.5 (7.5) years N = 108 Definition: Osteoarthritis diagnosed by an orthopedic	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Training for 20 minutes for each leg 3 sessions per week for 8 weeks. Group or individual: Individual session Type of exercise: Proprioception **Supervised strength exercise** (n=36) Quadriceps exercises completed with dynamometer cables and weights to increase resistance. 3 sessions per week with 4 sets (6 repetitions per set) for 8 weeks. Group or individual: Individual session **No treatment** (n=36) **Concomitant therapy:** All people were asked to cease any exercise activity outside of the exercise training	surgeon based on the clinical history, radiographic imaging and physical assessment Severity: Radiographic median grade 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Mccaffrey 2019[287]	**Other supervised exercise** (n=9) Chair yoga program twice-weekly 50-minute sessions for 8 weeks, a total of 16 sessions, led by a certified yoga instructor. Group or individual: Individual session	**Knee osteoarthritis** Mean age (SD): 78.5 (2.4) years N = 18 Definition: Reported pain associated with lower extremity osteoarthritis (hip, knee or other lower	Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Yoga). **Supervised strength exercise** (n=9) Chair exercise for older adults. An exercise program adapted from the standing Go4Life program designed for older adults to increase muscle strength, range of motion and activities of daily living. 50-minute program twice weekly for 8 weeks, for a total of 16 sessions. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** No additional information	extremities) verified by a nurse practitioner Severity: Not stated/unclear Duration of symptoms: Not stated/unclear. Presence of multimorbidities: Not stated/unclear		
Nahayatbin 2018[314]	**Other supervised exercise** (n=16) 10 minutes of Yang style Tai Chi with 5 minutes of warm up and cool down. Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=16)	**Knee osteoarthritis** Mean age (SD): 55.89 (5.97) years N = 48 Definition: Knee osteoarthritis with grade 2-3 changes based on the Kellgren Lawrence classification Severity: Kellgren Lawrence grade 2-3	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Closed chain kinetic exercises with10 minutes of exercise and 5 minutes of warm up and cool down. Group or individual: Group session **No treatment** (n=16) **Concomitant therapy:** All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound	Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Nambi, 2020[315]	**Supervised strength exercise** (n=20) The training knee was kept at a 90 degrees flexed position, and the dynamometer axis was aligned with the centre of the lateral femoral condyle. The knee was tested from 0 degrees to 120 degrees of flexion, where 0 degrees was considered full extension. Training was performed on 5 days a week for 4 weeks. Group or individual: Individual session Type of exercise: Not stated / Unclear	**Knee osteoarthritis** Mean age (SD): 22.5 (1.5) years N = 60 Definition: Chronic osteoarthritis after ACL injury (secondary osteoarthritis) Severity: Not stated/unclear Duration of injury (SD): 5.4 (0.4) months. Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Other supervised exercise (n=20) Sensory motor training which was given in 3 stages; static, dynamic and functional. All exercises were performed 5 times in 1 set, for 3 sets, with a sufficient with 5 minutes rest period between sets for 4 weeks. Group or individual: Individual session Type of exercise: Neuromodulatory **Unsupervised strength exercise** (n=20) Home-based exercises performed with 10-15 repetitions/day, 5 days a week for 4 weeks. Stretching was focused on each muscle group for 3 repetitions of 15 s per muscle group. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** No additional information.			
Ojoawo 2016[330]	Other supervised exercise (n=25)	Knee osteoarthritis Mean age (SD): 68.89 (10.28) years	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Proprioceptive exercise completed for 6 weeks Group or individual: Not stated/unclear Type of exercise: Proprioceptive Supervised strength exercise (n=25) Isometric quadriceps strengthening exercise with standard weights hung for resistance completed over 6 weeks Group or individual: Not stated/unclear **Concomitant therapy:** Infrared radiation therapy was applied with a methyl salicylate ointment for 20 minutes twice a week for 6 weeks.	N = 50 Definition: Diagnosis of knee osteoarthritis (in people with symptomatic and radiologic evidence) with symptoms of pain, stiffness and functional difficulty of no less than 6 weeks duration Severity: Not stated Duration of symptoms: At least 6 weeks Presence of multimorbidities: Not stated/unclear		
Rogers 2012[380]	**Supervised mixed modality exercise** (n=11) Kinaesthesia, balance and agility exercise training with resistance exercise training. Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory and strength **Other supervised exercise**	**Knee osteoarthritis** Mean age (SD): 70.4 (9.8) years N = 44 Definition: met the American College of Rheumatology diagnostic criteria for unilateral or bilateral symptomatic knee osteoarthritis as confirmed by the person's physician	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	(n=11) Kinaesthesia, balance and agility exercise training Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory **Supervised strength exercise** (n=11) Resistance exercise training only for 8 week three times a week for 30-40 minutes. Group or individual: Not stated/unclear A fourth group (n=11) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review **Concomitant therapy:** No additional information	Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Wortley 2013[489]	**Other supervised exercise** (n=15) 1 hour group training session twice a week based on the 12 basic movements adapted from the Yang Style Tai Ji for 10 weeks	**Knee osteoarthritis** Mean age (SD): 69.2 (6.0) years N = Definition: The Classification Criteria for Knee OA of the American College of	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=15) Resistance training program consisting of two 1 hour sessions per week using ankle cuff weights for resistance. Completed over 10 weeks Group or individual: Group session **No treatment** (n=9) **Concomitant therapy:** Participants were asked not to alter their regular physical activity or pain medications	Rheumatology and bilateral knee x-rays Severity: Kellgren Lawrence median grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		

1.1.5.9 Other supervised exercise compared to unsupervised strength exercise

Table 11: Summary of studies included in the evidence review for the other supervised exercise compared to unsupervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Kars Fertelli 2018[225]	**Other supervised exercise** (n=60) Aquatic exercise programme including a warmup, basic exercises and cool down. Intensity was gradually increased, and foam boards and	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 55.6 (7.8) years N = 120 Definition:	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	balls were used as aids. Three days a week for 8 weeks Group or individua): Group session Type of exercise: Hydrotherapy **Unsupervised strength exercise** (n=60) People were informed about how to do exercises that should be done by people with osteoarthritis and told to do exercises at home Group or individual: Individual session **Concomitant therapy:** No additional information	Knee or hip osteoarthritis as diagnosed by the American College of Rheumatology criteria insert the method of assessment Severity: Median grade 2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Kuptniratsaikul 2019[249]	**Other supervised exercise** (n=40) Underwater treadmill exercise with moderate intensity for 30 minutes three times a week for 4 weeks Group or individua): Individual session Type of exercise: Hydrotherapy **Unsupervised strength exercise** (n=40)	**Knee or hip osteoarthritis** Mean age (SD): 61.9 (6.7) years N = 80 Definition: Knee osteoarthritis with mild to moderate knee pain Severity: Not stated Duration of symptoms (median [range}): Home	Pain at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Quadriceps exercises repeated at 10-20 repetitions per set daily for 4 weeks Group or individual: Individual session **Concomitant therapy:** All participants received a leaflet advising them on how to use their knee joints in daily practice (i.e. warm compression for pain relief, regular isometric quadriceps exercise, and avoid bending the knee more than 90 degrees)	exercise: 4.0 (0.2, 20.0), UTM: 3.0 (0.1, 30.0) years Presence of multimorbidities: Not stated/unclear		
Lim 2010[265]	**Supervised mixed modality exercise** (n=26) Land based exercise program with generalised conditioning and knee-specific exercises conducted over 8 weeks in 40 minute sessions Group or individual: Individual session Type of exercise: Strength, aerobic, stretching/range of motion **Other supervised exercise** (n=25) Aquatic exercise with 3x40 minute sessions per week for 8 weeks Group or individual:	**Knee osteoarthritis** Mean age (SD): 65.0 (8.7) years N = 79 Definition: Definitive medial tibiofemoral osteophytes on x-ray with joint space narrowing greater in the medial tibiofemoral compartment compared to the lateral compartment Severity: Median radiographic severity – Moderate Duration of symptoms: At least 6 months Presence of multimorbidities:	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Individual session Type of exercise: Hydrotherapy **Unsupervised strength exercise** (n=24) Home-based exercise instructions including Q-sets exercise for strengthening of quadriceps muscles and a partial squatting along with behavioural correction of daily activities Group or individual: Individual session **Concomitant therapy:** No additional information	Not stated/unclear		
Nambi 2020[315]	**Supervised strength exercise** (n=20) The training knee was kept at a 90 degrees flexed position, and the dynamometer axis was aligned with the centre of the lateral femoral condyle. The knee was tested from 0 degrees to 120 degrees of flexion, where 0 degrees was considered full extension. Training was performed on 5 days a week for 4 weeks. **Other supervised exercise** (n=20)	**Knee osteoarthritis** Mean age (SD): 22.5 (1.5) years N = 60 Definition: Chronic osteoarthritis after ACL injury (secondary osteoarthritis) Severity: Not stated/unclear Duration of injury (SD): 5.4 (0.4) months. Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Sensory motor training which was given in 3 stages; static, dynamic and functional. All exercises were performed 5 times in 1 set, for 3 sets, with a sufficient with 5 minutes rest period between sets for 4 weeks. **Unsupervised strength exercise** (n=20) Home-based exercises performed with 10-15 repetitions/day, 5 days a week for 4 weeks. Stretching was focused on each muscle group for 3 repetitions of 15 s per muscle group. **Concomitant therapy:** No additional information.			

1.1.5.10 Other supervised exercise compared to no treatment

Table 12: Summary of studies included in the evidence review for the other supervised exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
An 2008[18]	**Other supervised exercise** (n=14) Baduanjin in eight sections repeated 20 times, delivered by a senior instructor, five times per week Group or individual: Group session	**Knee osteoarthritis** Mean age (SD): 65.0 (7.5) years N = 28 Definition: The clinical criteria for the classification of idiopathic	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Type of exercise: Mind-body **No treatment** (n=14) No treatment **Concomitant therapy:** No change in medication for arthritis was permitted during the trial	osteoarthritis of the knee developed by the American College of Rheumatology Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Cheung 2014[85]	**Other supervised exercise** (n=18) Hatha yoga classes, 60 minutes once per week for 8 weeks plus home practice for 30 minutes, four times per week. Sessions included ansanas (poses), pranas (breathing), and meditation Group or individual: Group session Type of exercise: Mind-body (yoga) **No treatment** (n=18) Waiting list control **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (95% CIs): yoga = 71.0 (69.3, 75.6) years; waiting list = 71.9 (69.0, 75.0) N = 36 Definition: Symptomatic osteoarthritis of the knee diagnosed at least 6 months prior. Symptoms classified under the American College of Rheumatology criteria Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: High morbidity score (Yoga mean [95% CI]: 2.8 [1.7, 3.9].	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
		Waiting list control: 1.4 [0.8, 2.0])		
Cheung 2017[84]	**Supervised mixed modality exercise** (n=28) Aerobic-strength exercise involving 15 minutes of mild aerobic exercise and 30 minutes of isometric and isotonic strengthening exercises. Classes were once a week for 8 weeks, and home practice was 4 times per week for the aerobic exercise, and twice per week for the strengthening exercise Group or individual: Group session Type of exercise: Strength, aerobic **Other supervised exercise** (n=32) Hatha yoga for 45 minutes, once per week for 8 weeks, plus 30 minutes four times per week home practice. Sessions included poses, breathing and relaxation/ mindfulness training Group or individual: Group session Type of exercise: Mind-body **No treatment** (n=23)	**Knee osteoarthritis** Mean age (SD): 71.6 (8.1) years N = 83 Definition: A self-reported medical diagnosis of osteoarthritis of the knee for at least 6 months Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Low morbidity score (1.5 [1.5])	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment **Concomitant therapy:** No additional information			
Cochrane 2005[91]	**Other supervised exercise** (n=153) Aquatic exercise therapy, including flexibility, strength, and isotonic and endurance (aerobic) exertion. Sessions were 1 hour, twice per week for 1 year Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=159) No exercise control **Concomitant therapy:** No additional information	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 69.75 (6.54) years N = 312 Definition: Diagnosis confirmed by a general practitioner and confirmed by a member of the research team Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: High morbidity score	Quality of life at >3 months Pain at >3 months Physical function at >3 months	
Duman 2012[111]	**Other supervised exercise** (n=30) Proprioceptive exercise including strengthening, bicycling, walking, and heel-to-toe and toe-to-heel walking, five days per week for three weeks Group or individual: Not stated/unclear Type of exercise: Proprioception	**Knee osteoarthritis** Mean age (SD): 64 (3.7) years N = 54 Definition: Knee osteoarthritis according to the American College of Rheumatology criteria with grade 3 or higher Kellgren	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=24) No exercise treatment **Concomitant therapy:** All people received non-steroidal anti-inflammatory drugs (meloxicam 15mg/day) and physical therapy (infrared and short-wave therapy)	Lawrence scale radiographic changes Severity: Kellgren Lawrence grade 3-4, median grade 3 Duration of symptoms (mean [SD]): 7.9 (1.7) years Presence of multimorbidities: Not stated/unclear		
Foley 2003[134]	**Other supervised exercise** (n=35) Hydrotherapy, including strengthening exercises. One set of 10 repetitions increased to three sets of 10-15, plus weighted gaiters, for 12 weeks Group or individual: Individual session Type of exercise: Hydrotherapy **No treatment** (n=35) Fortnightly telephone calls to record any changes in their condition, drug use, or injuries and free exercise treatment at the hospital at the end of the study period **Concomitant therapy:** No additional information	**Hip and/or knee osteoarthritis** Mean age (SD): 70.9 (8.8) years N = 105 Definition: Radiological diagnosis of osteoarthritis of the hip or knee, or both Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Fransen 2007[145]	**Other supervised exercise** (n=111) Tai Chi or hydrotherapy, performed for 1 hour, twice per week for 12 weeks. Tai Chi included modification of 24 forms of the Sun style. Hydrotherapy was a standardised protocol Group or individual: Group session Type of exercise: Hydrotherapy for one group, Tai Chi for another group **No treatment** (n=41) Waiting list control **Concomitant therapy:** No additional information	**Hip or knee osteoarthritis** Mean age (SD): 70.2 (6.3) years N = 152 Definition: Diagnosis of osteoarthritis involving the hip or knee as per American College of Rheumatology criteria and current and chronic hip or knee pain (at least 1 year) Severity: Not stated Duration of symptoms (median): 6-10 years Presence of multimorbidities: High morbidity score (mean (SD): 4.7 (2.7)	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Psychological distress at ≤3 months	
Hinman 2007[180]	**Other supervised exercise** (n=36) Aquatic physical therapy programme including functional weight bearing and progressive exercises, twice a week for 6 weeks. Participants were then encouraged to continue self-directed aquatic physical therapy for 6 weeks Group or individual: Group session Type of exercise:	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 62.4 (8.8) years N = 61 Definition: Diagnosis was based on American College of Rheumatology classification criteria	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Hydrotherapy **No treatment** (n=35) No treatment for 6 weeks, then completed the aquatic physical therapy program over the next 6 weeks **Concomitant therapy:** People continued using their usual medication	Severity: Not stated Duration of symptoms (mean [SD]): 8 (10.0) years Presence of multimorbidities: Not stated/unclear		
Lee 2009[256]	**Other supervised exercise** (n=29) Tai Chi Qigong performed for 1 hour, repeated twice a week for 8 weeks Group or individual: Group session Type of exercise: mind-body **No treatment** (n=15) Concomitant therapy: No additional information	**Knee osteoarthritis** Mean age (SD): 69.1 (5.5) years N = 44 Definition: Symptomatic osteoarthritis with radiologic alterations in the knee joint of grade 2 or higher (Kellgren-Lawrence Scale) at least 6 months prior to study entry Severity: Median Kellgren-Lawrence grade 2-3 Duration of symptoms: at least 6 months Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	
Lin 2004[268]	**Other supervised exercise** (n=66)	**Knee osteoarthritis**	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Water exercise programme consisted of 1 hour sessions twice a week over a period of 12 months. Accounting for holidays, the programme was run for a total of 46 weeks. Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=40) Health education leaflet **Concomitant therapy:** No additional information	Mean age (SD): 69.2 (6.00) years N = 106 Definition: People treated for osteoarthritis of the knee/hip from their general practitioner, rheumatologist or orthopaedic surgeon Severity: Not stated Duration of symptoms (mean [SD]): 12.2 (11.1) years Presence of multimorbidities: Not stated/unclear	Physical function at ≤3 months	
McIlroy 2017[289]	**Other supervised exercise** (n=7) Aquatic therapy in 6x30 minute weekly group sessions Group or individual: Individual session Type of exercise: Hydrotherapy **No treatment** (n=7) **Concomitant therapy:** All people attended one 30 minute individual, self-management education session with a physiotherapist. This	**Knee osteoarthritis** Mean age (SD): 63.3 (7.8) years N = 14 Definition: Adults with persistent knee pain of at least >3 months duration Severity: Not stated Duration of symptoms: >3 months Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	comprised: information on the causes of persistent knee pain, physical activity/aerobic exercise and knee exercise (e.g. quadriceps strengthening exercises); footwear advice and the use of shock absorbing insoles; activity pacing; pain (e.g. thermotherapy) and weight management.			
Munukka 2016[310]	**Other supervised exercise** (n=43) 1 hour of supervised lower limb aquatic resistance training three times a week for 16 weeks. Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=44) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 64 (2) years N = 87 Definition: Mild knee osteoarthritis demonstrated through radiography grade 1-2 changes according to the Kellgren-Lawrence classification experiencing knee pain on most days Severity: Kellgren Lawrence grade 1-2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at >3 months Pain at >3 months Physical function at >3 months Serious adverse events at >3 months	
Nahayatbin 2018[314]	**Other supervised exercise** (n=16) 10 minutes of Yang style Tai Chi with 5 minutes of warm up and cool down.	**Knee osteoarthritis** Mean age (SD): 55.89 (5.97) years N = 48	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=16) Closed chain kinetic exercises with10 minutes of exercise and 5 minutes of warm up and cool down. Group or individual: Group session **No treatment** (n=16) **Concomitant therapy:** All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound	Definition: Knee osteoarthritis with grade 2-3 changes based on the Kellgren Lawrence classification Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Patrick 2001[348]	**Other supervised exercise** (n=125) Arthritis Foundation certified aquatic class twice weekly for 20 weeks Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=124)	**Knee osteoarthritis** Mean age: 65.7 years N = 249 Definition: Clinically confirmed diagnosis of osteoarthritis from a physician Severity: Not stated	Quality of life at >3 months Pain at >3 months Physical function at >3 months Psychological distress at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: No additional information	Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Rewald 2020[374]	**Other supervised exercise** (n=55) Aquatic exercise (cycling in an upright position). Participants exercised twice per week for 45 minutes Also, out-of-the-saddle positions, leg exercises and upper body exercises were incorporated. Duration 12 weeks. Group or individual: Individual session. Type of exercise: Hydrotherapy **No treatment** (n=47) People were not prohibited to follow treatment that they would have also received outside of the trail. Subsequent physical therapy was not obliged and was not considered as part of the study. Duration 12 weeks **Concomitant therapy:** No additional information.	**Knee osteoarthritis** Mean age (SD): 59.9 (8.7) years N = 111 Definition: Knee pain between 4 and 7 on a 10-point numeric rating scale and a Kellgren-Lawrence score between 1 and 3. Severity: Kellgren Lawrence score between 1-3. Duration of symptoms: Not stated/unclear. Presence of multimorbidities: High morbidity score (Mean count comorbidity: intervention = 2 (1.7), usual care = 1 (1.3).).	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Robbins 2022[377]	**Other supervised exercise** (n=86) Either stretching exercise and laser therapy or stretching	**Knee osteoarthritis** Mean age (SD): 66.6 (9.6) years N = 172	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	exercise alone, three times a week for 8 weeks. Group or individual: Group session Type of exercise: Flexibility **No treatment** (n=86) Either laser therapy alone or educational booklet only (no treatment control). **Concomitant therapy:** No additional information	Definition: Knee osteoarthritis diagnosed by an independent rehabilitation specialist (clinical and radiological) Severity: Osteoarthritis degree 2-4, median grade 3 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear		
Segal 2015[401]	**Other supervised exercise** (n=36) Gait training intervention completed in 24 biweekly 45 minute sessions Group or individual: Individual session Type of exercise: Neuromodulatory **No treatment** (n=22) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 69.3 (7.0) years N = 58 Definition: Symptomatic knee osteoarthritis, defined by a definite osteophyte or joint space narrowing in either tibiofemoral compartment or posteroanterior knee radiographs and an affirmative response to "Have you had pain or stiffness in one or both knees on most of the past 30 days" on both the telephone screen and screening visit and mobility disability [LLFDI advanced	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
		lower limb function score below 32 points]) Severity: Kellgren Lawrence grade 2-4, median grade 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Sekir 2005[403]	**Other supervised exercise** (n=12) Multistation exercise program including balance and proprioception exercises twice a week for 6 weeks Group or individual: Individual session Type of exercise: Proprioception **No treatment** (n=10) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 60.4 (8.7) years N = 22 Definition: People with bilateral complaints of knee osteoarthritis, who had grade 2 or 3 osteoarthritis, as judged by the criteria of the American College of Rheumatology based on weight-bearing radiographs Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months	
Song 2003[415] Subsidiary paper:	**Other supervised exercise** (n=22)	**Knee osteoarthritis** Mean age (SD): 63.7 (5.9) years	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Song 2007[414]	Sun style Tai Chi exercise taught over 12 weeks with an instructional audiotape to practice at home Group or individual: Group session Type of exercise: Mind-body **No treatment** (n=21) **Concomitant therapy:** No additional information	N = 43 Definition: Clinical and radiographic evidence of knee osteoarthritis according to the American College of Rheumatology criteria with a Kellgren Lawrence grade of at least 2 Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms (mean [SD]): 9.8 (7.2) years Presence of multimorbidities: Not stated/unclear		
Wang 2007[478]	**Other supervised exercise** (n=21) Aquatic exercises conducted over 12 weeks Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=21) **Concomitant therapy:** No additional information	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 66.2 (12.6) years N = 42 Definition: People diagnosed with osteoarthritis of the hip or knee Severity: Number of tender joints (mean [SD]): 6.8 (4.8). Duration of symptoms (mean [SD]): 13.5 (11.8) years Presence of multimorbidities:	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
		Not stated/unclear		
Wang 2011[479]	**Supervised mixed modality exercise** (n=28) Land based exercise included flexibility and aerobic training, three times a week for 12 weeks Group or individual: Group session Type of exercise: Upper body, lower body training, flexibility and aerobic **Other supervised exercise** (n=28) Aquatic exercise with a flexibility and aerobic training class, three times a week for 12 weeks Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=26) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 67.7 (5.9) years N = 82 Definition: Knee osteoarthritis diagnosed by physician assessment based on symptoms and X-ray Severity: Not stated Duration of symptoms (mean [SD]): 6.8 (6.4) years Presence of multimorbidities: Low morbidity score	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	
Wortley 2013[489]	**Other supervised exercise** (n=15) 1 hour group training session twice a week based on the 12 basic movements adapted from	**Knee osteoarthritis** Mean age (SD): 69.2 (6.0) years N = 39	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	the Yang Style Tai Ji for 10 weeks Group or individual: Group session Type of exercise: Mind-body **Supervised strength exercise** (n=15) Resistance training program consisting of two 1 hour sessions per week using ankle cuff weights for resistance. Completed over 10 weeks Group or individual: Group session **No treatment** (n=9) **Concomitant therapy:** Participants were asked not to alter their regular physical activity or pain medications	Definition: The Classification Criteria for Knee OA of the American College of Rheumatology and bilateral knee x-rays Severity: Kellgren Lawrence median grade 2-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Ye 2019[496]	**Other supervised exercise** (n=25) Banduajin Qigong training. 3 sessions per week with each session lasting 40 minutes. Duration 12 weeks. Group or individual : Group session Type of exercise: Mind-body **No treatment** (n=25)	**Knee osteoarthritis** Mean age (SD): 63.8 (6.2) years N = 50 Definition: People diagnosed with knee osteoarthritis according to the criteria of the American College of Rheumatology with radiographic grading of the severity between 2 and 3	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Patients in the control group were informed to maintain their unaltered lifestyle while refraining from other supervised exercise training program. Duration 12 weeks **Concomitant therapy:** No additional information	Severity: Radiographic grade 2-3. Duration of symptoms: Not stated/unclear. Presence of multimorbidities: Not stated/unclear		
Ye 2020[497]	**Other supervised exercise** (n=28) 12 weeks Banduanjin program. People were asked to perform three banduanjin sessions per week, with each session lasting for 40 minutes. Duration 12 weeks. Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Banduajin Qigong). **No treatment** (n=28) Usual care. Participants received the exercise intervention after 12 weeks. **Concomitant therapy:** All people received conventional therapies (acupuncture, massage and moxibustion), one hour each day, five days a week for the first four weeks.	**Knee osteoarthritis** Mean age (SD): 64.4 (5.1) years N = 56 Definition: Knee osteoarthritis diagnosed according to the criteria of the American College of Rheumatology, with a radiographic grading of the severity between 2 and 3 Severity: Radiographic grade between 2 and 3 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

1.1.5.11 Other unsupervised exercise compared to unsupervised strength exercise

Table 13: Summary of studies included in the evidence review for the other unsupervised exercise compared to unsupervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Chaipinyo 2009[77]	**Other unsupervised exercise** (n=24) Balance exercise including stepping forward, backward and sideways, 5 days per week. Bilateral mini squats were also done to strengthen the quadriceps Group or individual: Individual session Type of exercise: Neuromuscular **Unsupervised strength exercise** (n=24) Strengthening exercise of isometric knee extension, 30 repetitions performed 5 days per week Group or individual: Individual **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 66 (7.2) years N = 48 Definition: Knee osteoarthritis as per the American College of Rheumatology clinical criteria Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

1.1.5.12 Supervised mixed modality exercise compared to supervised strength exercise

Table 14: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to supervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Cantero-Tellez 2021[72]	**Supervised mixed modality exercise** (n=6) A supervised proprioceptive training program divided into three phases. Each phase was performed for 2 consecutive weeks. Group or individual: Individual session Type of exercise: Proprioception **Supervised strength exercise** (n=6) No additional treatment. Duration 12 weeks Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** Conservative treatments including a short opponens orthosis for night-time wear, self passive traction of the thumb CMC joint, self-massage to the thumb muscles, active resistance of the FDI muscle, and instruction for functional incorporation of the thumb for	**Hand osteoarthritis** Age (Range of means): 65.33-67.17 years N = 12 Definition: Thumb carpometacarpal joint osteoarthritis diagnosed as grade I or II by the Eaton Classification Stage Severity: Eaton Classification Stage grades I-II Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	activities of daily living. The exercise routine was performed on a home program basis 2 times per day (3 sets of 8-10 repetitions) and seen twice a week in the clinic to monitor and provide feedback for proper performance of the exercise routine			
Diracoglu 2005[106]	**Supervised mixed modality exercise** (n=33) Kinesthesia and balance exercises with strengthening exercises, 3 days a week for 8 weeks Group or individual: Group session Type of exercise: Strength, proprioception **Supervised strength exercise** (n=33) Isometric and isotonic strength exercises completed 3 days a week in groups of 5 people for 8 weeks Group or individual: Group sessions **Concomitant therapy:** Paracetamol was given as an escape medicine for pain control	**Knee osteoarthritis** Age range: 35-65 years N = 66 Definition: Primary osteoarthritis according to the criteria of American College of Rheumatology with radiological stage 1-2 bilateral knee osteoarthritis according to the Kellgren and Lawrence scale Severity: Kellgren and Lawrence grade 1-2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Diracoglud 2008[107]	**Supervised mixed modality exercise** (n=33)	**Knee osteoarthritis**	Quality of life at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Active range of motion exercises, active stretching and isometric strengthening, three days a week for eight weeks Group or individual: Group session Type of exercise: Strength, range of motion **Supervised strength exercise** (n=33) Strength exercise component only - no balance exercises Group or individual: Group session **Concomitant therapy:** No additional information	Mean age (SD): 50.5 (7.2) years N = 66 Definition: Primary knee osteoarthritis fulfilling the clinical and radiological criteria of the American College of Rheumatology Severity: Kellgren and Lawrence grade 1-2, median grade 1 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	
Hernandez 2019[179]	**Supervised mixed modality exercise** (n=53) Conventional exercises plus exercises aimed at the activation of the muscles considered important for core stability according to electromyography tests. Treatments were delivered in triweekly sessions for three months. Group or individual : Individual session	**Knee osteoarthritis** Mean age (SD): 62.3 (10.6) years N = 113 Definition: Medical diagnosis of knee osteoarthritis referred by the Orthopedics Department to the Physical Therapy Department of Hospital Durand. Confirmed by an orthopedist based on radiographic and clinical findings.	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months Serious adverse events at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Type of exercise: Other (Specific muscular exercises to increase core stability and strengthening exercises). **Supervised strength exercise** (n=60) No additional exercises. Duration 3 months. Group or individual : Individual session Type of exercise: Not applicable **Concomitant therapy:** All groups were offered conventional exercises including warm up and mobility as well as strengthening and stretching exercises.	Severity: Not stated/unclear Duration of symptoms (median [range]): Experimental group = 11.5 (1-120), control group = 8.5 (1-72) (units unclear) Presence of multimorbidities: Not stated/Unclear		
Joshi 2019[215]	**Supervised mixed modality exercise** (n=21) Retrowalking protocol on a treadmill, three session in a week for a total duration of six weeks. Group or individual: Individual session Type of exercise: Other (Aerobic and strengthening). **Supervised strength exercise** (n=21)	**Knee osteoarthritis** Mean age (SD): 52.5 (9.5) years N = 42 Definition: Knee osteoarthritis diagnosed by an orthopedician. Severity: Not stated/unclear Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Conventional exercise only. Three sessions per week for six weeks. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** Both groups received conventional exercise program which consisted of hot packs for 10 minutes followed by exercises. These consisted of range of motion exercises, muscle strengthening exercise in the form of isometric and isotonic exercises, muscle stretching exercises and flexibility exercises.			
Knoop 2013[237] Subsidiary papers: Knoop 2014[239] Knoop 2015[238]	**Supervised mixed modality exercise** (n=80) An exercise programme and home exercise programme with increasing intensity, knee load and difficulty, targeting knee joint stabilisation, muscle strength and performance of daily activities Group or individual: Individual session Type of exercise: Strength and proprioceptive **Supervised strength exercise** (n=79)	**Knee osteoarthritis** Mean age (SD): 62.0 (7.1) years N = 159 Definition: Diagnosis of knee osteoarthritis according to clinical American College of Rheumatology criteria Severity: Kellgren Lawrence grade 0-4, median grade 2	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Exercise targeting muscle strength and performance of daily activities Group or individual: Individual session **Concomitant therapy:** Not stated/unclear	Duration of symptoms (mean [SD]): 10.8 (9.3) years Presence of multimorbidities: Not stated/unclear		
Kumar 2013[248]	**Supervised mixed modality exercise** (n=22) Resistive exercise, including knee flexor and extensor strengthening, hip extensors, hip abductors and hip external rotators, and proprioceptive training. Group or individual: Individual session Type of exercise: Strength and proprioception **Supervised strength exercise** (n=22) Resistive exercise only, without proprioceptive training Group or individual: Individual session **Concomitant therapy:** Not stated/unclear	**Knee osteoarthritis** Mean age (SD): 53.3 (6.2) years N = 44 Definition: People referred with knee osteoarthritis (diagnosed radiologically or clinically) Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	
Pazit 2018[349]	**Supervised mixed modality exercise** (n=10)	**Knee osteoarthritis** Mean age (SD): 67.68 (6.68) years	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	High speed resistance training with balance training for 8 weeks. Group or individual: Group session Type of exercise: Strength and balance **Supervised strength exercise** (n=10) High speed resistance training for 8 weeks. Group or individual: Group session **No treatment** (n=10) **Concomitant therapy:** No additional information	N = 28 Definition: Knee osteoarthritis based on the presence of clinical symptoms of knee osteoarthritis as defined by the American College of Rheumatology criteria Severity: Not stated Duration of symptoms: At least 6 months Presence of multimorbidities: High morbidity score	Physical function at ≤3 months Serious adverse events at ≤3 months	
Rogers 2012[380]	**Supervised mixed modality exercise** (n=11) Kinaesthesia, balance and agility exercise training with resistance exercise training. Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory and strength	**Knee osteoarthritis** Mean age (SD): 70.4 (9.8) years N = 44 Definition: met the American College of Rheumatology diagnostic criteria for unilateral or bilateral symptomatic knee osteoarthritis as confirmed by the person's physician	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Other supervised exercise (n=11) Kinaesthesia, balance and agility exercise training Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory **Supervised strength exercise** (n=11) Resistance exercise training only for 8 week three times a week for 30-40 minutes. Group or individual: Not stated/unclear A fourth group (n=11) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review **Concomitant therapy:** No additional information	Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Sedghatnezhad 2020[400]	**Supervised mixed modality exercise** (n=15) Walking on an uphill treadmill, gradually increasing the slope for a total of 30 minutes (15 minutes before physical therapy,	**Knee osteoarthritis** Mean age (SD): 56.7 (8.0) years N = 30	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	15 minutes after physical therapy). Physical therapy (available to all participants) included a strengthening exercise program (see concomitant treatment). Duration 2 weeks. Group or individual: Individual session Type of exercise: Other (Strengthening and aerobic). **Supervised strength exercise** (n=15) Strengthening exercise component (and other physical therapy available to all participants) only. Duration 2 weeks. Group or individual: Individual session Type of exercise: Not applicable **Concomitant therapy:** Everyone received the following: a 201B ultrasound used for continuous ultrasound therapy (using a 1MHz head set to 1W/cm² applied for 6 minutes - 3 minutes on the anteromedial and 3 on the posterior of the knee); a transcutaneous nerve stimulation unit giving therapy for 20 minutes at 100Hz for a pulse duration of 50	Definition: Knee osteoarthritis according to the American College of Rheumatology (including radiographic findings) Severity: Kellgren Lawrence grade II-III Duration of symptoms: Not stated/unclear. Presence of multimorbidities: Not stated/unclear		

Study	Intervention and comparison	Population	Outcomes	Comments
	microseconds; two hot packs on the anterior and posterior aspects of the knees. This was followed by a muscle strengthening program performed individually in two sets, repeated from 10 up to 30 times between the first and fifth sessions, and then 30 for the remaining five sessions. Exercises included supine quadriceps setting, side lying hip abduction and standing heel raising on two legs.			
Vaghela 2020[458]	**Supervised mixed modality exercise** (n=43) Yoga therapy. This included six asanas, each consisting of ten repetitions with short intervals of rest in between for a total of 30 minutes per session, three times per week for 4 weeks. Group or individual: Individual session Type of exercise: Other (Yoga and strengthening). **Supervised strength exercise** (n=40) Conventional physiotherapy only. Duration 4 weeks. Group or individual: Individual session Type of exercise: Not applicable	**Knee osteoarthritis** Mean age (SD): 55.5 (9.4) years. N = 83 Definition: Bilateral osteoarthritis of the knee based on the clinical American College of Rheumatology criteria Severity: Not stated/unclear Duration of symptoms: Not stated/unclear. Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: Conventional physiotherapy program included the following: Transelectrical nerve stimulation (10 minutes), isometric quadriceps exercise, straight leg-raising exercise in supine, terminal knee extension or vastus medialis oblique strengthening exercise in supine and high sitting; straight leg abduction exercise in side lying. Each exercise was performed for a total of three sets, with each set made up of ten repetitions for 20 minutes, three times a week for 4 weeks.			

1.1.5.13 Supervised mixed modality exercise compared to unsupervised strength exercise

Table 15: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to unsupervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Lim 2010[265]	**Supervised mixed modality exercise** (n=26) Land based exercise program with generalised conditioning and knee-specific exercises conducted over 8 weeks in 40 minute sessions Group or individual: Individual session Type of exercise:	**Knee osteoarthritis** Mean age (SD): 65.0 (8.7) years N = 79 Definition: Definitive medial tibiofemoral osteophytes on x-ray with joint space narrowing greater in the medial tibiofemoral	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Strength, aerobic, stretching/range of motion **Other supervised exercise** (n=25) Aquatic exercise with 3x40 minute sessions per week for 8 weeks Group or individual: Individual session Type of exercise: Hydrotherapy **Unsupervised strength exercise** (n=24) Home-based exercise instructions including Q-sets exercise for strengthening of quadriceps muscles and a partial squatting along with behavioural correction of daily activities Group or individual: Individual session **Concomitant therapy:** No additional information	compartment compared to the lateral compartment Severity: Median radiographic severity – Moderate Duration of symptoms: At least 6 months Presence of multimorbidities: Not stated/unclear		

1.1.5.14 Supervised mixed modality exercise compared to supervised aerobic exercise

Table 16: Summary of studies included in the evidence review for the supervised strength exercise compared to supervised aerobic exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Hunt 2018[196]	**Supervised mixed modality exercise** (n=39) Toe-out gait modification programme, which involved training to perform walking with 15 degrees more toe-out than the self-selected amount using mirror guided biofeedback, and treadmill walking Group or individual: Individual session Type of exercise: Proprioceptive (gait adjustment) and aerobic **Supervised aerobic exercise** (n=40) Treadmill walking without instruction relating to toe-out walking Group or individual: Individual session **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 65.0 (8.7) years N = 79 Definition: Definitive medial tibiofemoral osteophytes on x-ray with joint space narrowing greater in the medial tibiofemoral compartment compared to the lateral compartment Severity: Median radiographic severity – Moderate Duration of symptoms: At least 6 months Presence of multimorbidities: Not stated/unclear	Pain at >3 months Physical function at >3 months	

1.1.5.15 Supervised mixed modality exercise compared to other supervised exercise

Table 17: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to other supervised exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Cheung 2017[84]	**Supervised mixed modality exercise** (n=28) Aerobic-strength exercise involving 15 minutes of mild aerobic exercise and 30 minutes of isometric and isotonic strengthening exercises. Classes were once a week for 8 weeks, and home practice was 4 times per week for the aerobic exercise, and twice per week for the strengthening exercise Group or individual: Group session Type of exercise: Strength, aerobic **Other supervised exercise** (n=32) Hatha yoga for 45 minutes, once per week for 8 weeks, plus 30 minutes four times per week home practice. Sessions included poses, breathing and relaxation/ mindfulness training Group or individual: Group session Type of exercise: Mind-body	**Knee osteoarthritis** Mean age (SD): 71.6 (8.1) years N = 83 Definition: A self-reported medical diagnosis of osteoarthritis of the knee for at least 6 months Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Low morbidity score (1.5 [1.5])	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=23) No treatment **Concomitant therapy:** No additional information			
Holm 2020[453] Subsidiary paper: (Holm 2021[182])	**Supervised mixed modality exercise** (n=45) The people in this group performed one set of low-intensity, high-repetition (30-60RM) knee extensions followed by 4 sets of high-intensity (8-12RM) leg-press in gym machines. Duration 12 weeks. Group or individual: Individual session (Initially group, but for the majority individual). Type of exercise: Other (Strength and neuromodulatory). **Other supervised exercise** (n=45) No additional therapy. Duration 12 weeks. Group or individual: Individual session (Initially group, but for the majority individual). Type of exercise: Neuromodulatory **Concomitant therapy:** Education was provided in the first week. Both groups received neuromuscular exercises twice	**Knee osteoarthritis** Mean age (SD): 64.7 (10.2) years N = 90 Definition: Symptomatic and radiographic (Kellgren and Lawrence at least 2) knee osteoarthritis deemed ineligible for knee replacement surgery by orthopedic surgeons in the orthopedic outpatient clinic at Naestved Hospital. Severity: Kellgren and Lawrence at least 2 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	weekly (60 minute sessions) for 12 weeks. All exercises were performed in 2-3 sets of 10-15 repetitions with three levels of difficulty.			
Lim 2010[265]	**Supervised mixed modality exercise** (n=26) Land based exercise program with generalised conditioning and knee-specific exercises conducted over 8 weeks in 40 minute sessions Group or individual: Individual session Type of exercise: Strength, aerobic, stretching/range of motion **Other supervised exercise** (n=25) Aquatic exercise with 3x40 minute sessions per week for 8 weeks Group or individual: Individual session Type of exercise: Hydrotherapy **Unsupervised strength exercise** (n=24) Home-based exercise instructions including Q-sets exercise for strengthening of quadriceps muscles and a	**Knee osteoarthritis** Mean age (SD): 65.0 (8.7) years N = 79 Definition: Definitive medial tibiofemoral osteophytes on x-ray with joint space narrowing greater in the medial tibiofemoral compartment compared to the lateral compartment Severity: Median radiographic severity – Moderate Duration of symptoms: At least 6 months Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	partial squatting along with behavioural correction of daily activities Group or individual: Individual session **Concomitant therapy:** No additional information			
Rogers 2012[380]	**Supervised mixed modality exercise** (n=11) Kinaesthesia, balance and agility exercise training with resistance exercise training. Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory and strength **Other supervised exercise** (n=11) Kinaesthesia, balance and agility exercise training Training for 8 weeks three times a week for 30-40 minutes. Group or individual: Not stated/unclear Type of exercise: Neuromodulatory **Supervised strength exercise**	**Knee osteoarthritis** Mean age (SD): 70.4 (9.8) years N = 44 Definition: met the American College of Rheumatology diagnostic criteria for unilateral or bilateral symptomatic knee osteoarthritis as confirmed by the person's physician Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	(n=11) Resistance exercise training only for 8 week three times a week for 30-40 minutes. Group or individual: Not stated/unclear A fourth group (n=11) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review **Concomitant therapy:** No additional information			
Silva 2008[408]	**Supervised mixed modality exercise** (n=32) Land based therapy based on strengthening and gait training exercises including stretching, isometric strengthening, isotonic strengthening and gait training exercises performed in groups of 5 to 8 people in 50 minute sessions 3 times a week for 18 weeks Group or individual: Group session Type of exercise: Strength and proprioception **Other supervised exercise**	**Knee osteoarthritis** Mean age (SD): 59 (6.8) years N = 64 Definition: Clinical and radiographic diagnosis of osteoarthritis of the knee according to the American College of Rheumatology criteria Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	(n=32) Hydrotherapy including stretching, isometric strengthening, isotonic strengthening and gait training exercises performed in groups of 5 to 8 people in 50 minute sessions 3 times a week for 18 weeks Group or individual: Group session Type of exercise: Hydrotherapy **Concomitant therapy:** People were instructed to take 50mg sodium diclofenac tablets as required, not surpassing a maximum dose of 150mg per day			
Wang 2011[479]	**Supervised mixed modality exercise** (n=28) Land based exercise included flexibility and aerobic training, three times a week for 12 weeks Group or individual: Group session Type of exercise: Upper body, lower body training, flexibility and aerobic **Other supervised exercise** (n=28)	**Knee osteoarthritis** Mean age (SD): 67.7 (5.9) years N = 82 Definition: Knee osteoarthritis diagnosed by physician assessment based on symptoms and X-ray Severity: Not stated Duration of symptoms (mean [SD]): 6.8 (6.4) years Presence of multimorbidities: Low morbidity score	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Aquatic exercise with a flexibility and aerobic training class, three times a week for 12 weeks Group or individual: Group session Type of exercise: Hydrotherapy **No treatment** (n=26) **Concomitant therapy:** No additional information			
Xiao 2020[491] Subsidiary paper: (Xiao 2021[492])	**Other supervised exercise** (n=49) Wu Qin Xi Qigong exercise program.. Each participant performed three repetitions, with a 2-minute rest period between the sets. Training took place in groups four times a week (each session 60 minutes) for 24 weeks. Group or individual: Individual session Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Qigong). **Supervised mixed modality exercise** (n=49) Conventional physical therapy consisting of muscle-strength training of the lower extremity	**Knee osteoarthritis** Mean age (SD): 70.4 (9.72) years N = 98 Definition: People with knee osteoarthritis diagnosed by senior physicians based on standard clinical, endoscopic, radiologic and histological criteria Severity: X-ray classification grade I-II, median grade II Duration of symptoms (SD): 12.44 (4.17) months Presence of multimorbidities: Not applicable	Pain at >3 months Physical function at >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	and aerobic training. The exercise program was conducted 4 days a week for 24 weeks. Group or individual: Individual session Type of exercise: Other (Strength and aerobic). **Concomitant therapy:** No additional information.			

1.1.5.16 Supervised mixed modality exercise compared to unsupervised mixed modality exercise

Table 18: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to unsupervised mixed modality exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Tunay 2010[451]	**Supervised mixed modality exercise** (n=30) Proprioceptive exercise training at hospital and a home program including application of a cold compress and strengthening exercises Group or individual: Individual session Type of exercise: Strength and proprioception **Unsupervised mixed modality exercise** (n=30) A home program including proprioceptive exercises,	**Knee osteoarthritis** Mean age (SD): 52.3 (8.8) years N = 60 Definition: People diagnosed with knee osteoarthritis Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	application of a cold compress and strengthening exercises Group or individual: Individual session Type of exercise: Strength and proprioception **Concomitant therapy:** No additional information			

1.1.5.17 Supervised mixed modality exercise compared to pharmacological treatments

Table 19: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to pharmacological treatments comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Holsgaard-Larsen 2017[185] Subsidiary papers: Clausen 2014[90] Holsgaard-Larsen 2018[184]	**Supervised mixed modality exercise** (n=47) Exercise including a warmup, functional, proprioceptive, and endurance strengthening exercise, and cool down Group or individual: Group session Type of exercise: Proprioceptive, functional, strengthening **Pharmacological treatment** (n=46) People received best information on how to use paracetamol and oral NSAIDs, in doses consistent with the	**Knee osteoarthritis** Mean age (SD): 58.1 (8.0) years N = 95 Definition: Clinical diagnosis of knee osteoarthritis in accordance with the American College of Rheumatology criteria, with or without radiographic changes Severity: Kellgren Lawrence grade 0-3, median grade 2 Duration of symptoms: Not stated	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Danish guidelines. Additional NSAIDs could be prescribed from their GP is over-the-counter paracetamol was not sufficient Class of medicine: Oral treatment **Concomitant therapy:** No additional information	Presence of multimorbidities: Not stated/unclear		
Saccomanno 2016[386]	**Supervised mixed modality exercise** (n=55) Rehabilitation exercises for a total of 20 sessions in 1 month including isometric and isotonic exercises, stretching and proprioceptive exercises Group or individual: Individual session Type of exercise: Strength, proprioception, stretching **Pharmacological treatment** (n=55) Intra-articular hyaluronic acid. Three injections (one injection every 2 weeks) of high molecular weight hyaluronic acid (Orthovisc 2mL, 15mg/mL) Class of medicine: Intra-articular treatment	**Knee osteoarthritis** Mean age (SD): 61.8 (11.2) years N = 165 Definition: Knee osteoarthritis according to the American College of Rheumatology diagnostic criteria with knee malalignment confirmed by radiographic examinations Severity: Kellgren Lawrence grade 1-3, median grade 1 Duration of symptoms (median [IQR]): between 24-36 (10-80). Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	A third group (n=55) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review. **Concomitant therapy:** No additional information			

1.1.5.18 Supervised mixed modality exercise compared to no treatment

Table 20: Summary of studies included in the evidence review for the supervised mixed modality exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Abbott 2013[2] Subsidiary paper: Abbott 2019[3]	**Supervised mixed modality exercise** (n=51) Programme of aerobic, muscle strengthening, stretching and neuromuscular control exercises. Additional home exercise 3 times per week Group or individual: Group session Type of exercise: Aerobic, strengthening, neuromodulatory **No treatment** (n=51) No trial physiotherapy A third group (n=104) was reported but was not included in the analysis as they did not fulfil	**Hip or knee osteoarthritis** Mean age (SD): 66.6 (6.9) years N = 206 Definition: Considered for hip or knee joint replacement based on the clinical criteria of the American College of Rheumatology Severity: Not stated Duration of symptoms (mean [SD]): 2.7 (1.4) years Presence of multimorbidities: Not stated/unclear	Pain at >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	the inclusion criteria for this review. **Concomitant therapy:** Usual care provided by GP and other healthcare providers			
Aglamis 2008[6] Subsidiary papers: Aglamis 2009[5]	**Supervised mixed modality exercise** (n=17) Exercise programme including walking at a comfortable pace, functional strengthening exercise using body weight resistance, and a static stretching programme, three times per week for 12 weeks. Group or individual: Not stated/unclear Type of exercise: Aerobic, strengthening, flexibility **No treatment** (n=17) No treatment control **Concomitant therapy:** No additional information	**Hip or knee osteoarthritis** Mean age (SD): 55.7 (5.0) years N = 34 Definition: Radiographic grade 2-4 Kellgren Lawrence knee osteoarthritis Severity: Kellgren Lawrence grade 2-4 Duration of symptoms (mean [SD]): Not stated Presence of multimorbidities: Low morbidity score (mean score <1.5)	Quality of life at ≤3 months Pain at ≤3 months	
Cheung 2017[84]	**Supervised mixed modality exercise** (n=28) Aerobic-strength exercise involving 15 minutes of mild aerobic exercise and 30 minutes of isometric and isotonic strengthening exercises. Classes were once a week for 8	**Knee osteoarthritis** Mean age (SD): 71.6 (8.1) years N = 83 Definition:	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	weeks, and home practice was 4 times per week for the aerobic exercise, and twice per week for the strengthening exercise Group or individual: Group session Type of exercise: Strength, aerobic **Other supervised exercise** (n=32) Hatha yoga for 45 minutes, once per week for 8 weeks, plus 30 minutes four times per week home practice. Sessions included poses, breathing and relaxation/ mindfulness training Group or individual: Group session Type of exercise: Mind-body **No treatment** (n=23) No treatment **Concomitant therapy:** No additional information	A self-reported medical diagnosis of osteoarthritis of the knee for at least 6 months Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Low morbidity score (1.5 [1.5])		
De Matos Brunelli Braghin 2019[100]	**Supervised mixed modality exercise** (n=15) Exercise sessions included a warmup, strengthening exercises and aerobic exercise on a stationary bike. Three	**Knee osteoarthritis** Mean age (SD): 71.6 (8.1) years N = 120 Definition:	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	stages with 4-5 sessions per stage Group or individual: Group session Type of exercise: Strength, aerobic **No treatment** (n=15) No treatment A third and fourth group (n=30) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review **Concomitant therapy:** No additional information	Radiographic diagnosis of knee osteoarthritis, grade 1-3 according to the Kellgren and Lawrence classification Severity: Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
De Rooij 2017[101]	**Supervised mixed modality exercise** (n=63) Individualised exercise programme, including lower extremity muscle strength training, aerobic training and training of daily activities, plus home exercise five times per week Group or individual: Group session Type of exercise: Strength, aerobic and activity based	**Knee osteoarthritis** Mean age (SD): 63.6 (10.6) years N = 126 Definition: Knee osteoarthritis according to the clinical criteria of the American College of Rheumatology Severity: Kellgren Lawrence grade 0-4, median grade 2	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months Psychological distress at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=63) Waiting list control **Concomitant therapy:** People continued their current medical care for knee osteoarthritis and comorbid disease	Duration of symptoms (mean [SD]): 9.0 (9.0) years Presence of multimorbidities: High morbidity score (inclusion criteria required a score of at least 2 on a comorbidity scale)		
French 2013[146] Subsidiary papers: French 2009[147]	**Supervised mixed modality exercise** (n=45) Exercise therapy including flexibility and strengthening exercises in 6-8 sessions of 30-minutes, over 8 weeks. A daily home exercise programme supplemented the clinic-based treatments. Aerobic exercise for 30 minutes, 5 days per week was also encouraged Group or individual: Individual session Type of exercise: Strength and flexibility **No treatment** (n=63) Waiting list for 8 weeks, then randomisation into the exercise or exercise and manual therapy groups at week 9 A third group (n=43) was reported but was not included in the analysis as they did not fulfil	**Hip osteoarthritis** Mean age (SD): 62.5 (9.9) years N = 131 Definition: Osteoarthritis of the hip according to the American College of Rheumatology clinical and radiographic criteria Severity: Not stated Duration of symptoms (mean [SD]): 34.5 (45.5) months Presence of multimorbidities: High morbidity score (2.2 [1.4])	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	the inclusion criteria for this review **Concomitant therapy:** All groups received standardised written information on hip osteoarthritis. All nonconsenting and excluded participants were treated as usual by the physiotherapy department in each trial centre. Participants were asked to avoid all other interventions for the duration of the RCT, apart from routine doctor care and analgesics			
Keefe 2004[228]	**Supervised mixed modality exercise** (n=16) Three supervised exercise sessions per week for 12 weeks, including cardiopulmonary endurance training; strength training and flexibility/range of motion training Group or individual: Group session Type of exercise: Strength and aerobic **No treatment** (n=18) No exercise care A third group (n=38) was reported but was not included in	**Knee osteoarthritis** Mean age (SD): 59.5 (11.36) years N = 72 Definition: Persistent knee pain due to osteoarthritis and were diagnosed as having osteoarthritis of the knees Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	the analysis as they did not fulfil the inclusion criteria for this review **Concomitant therapy:** People were allowed to continue to receive their routine care			
Kraus 2014[242] Subsidiary papers: Krauss 2011[244]	**Supervised mixed modality exercise** (n=71) Tübinger exercise therapy approach entailing a once-weekly group intervention and twice-weekly home exercise. The therapeutic program entailed education and social interaction as well as exercises to strengthen the muscles and to improve proprioception, balance and flexibility Group or individual: Group session Type of exercise: Strength, proprioception and flexibility **No treatment** (n=69) No exercise control A third group (n=78) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review	**Hip osteoarthritis** Mean age (SD): 59 (10) years N = 218 Definition: Osteoarthritis of one or both hip joints according to the clinical criteria of the American College of Rheumatology Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: No additional information			
Pazit 2018[349]	**Supervised mixed modality exercise** (n=10) High speed resistance training with balance training for 8 weeks. Group or individual: Group session Type of exercise: Strength and balance **Supervised strength exercise** (n=10) High speed resistance training for 8 weeks. Group or individual: Group session **No treatment** (n=10) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 67.68 (6.68) years N = 28 Definition: Knee osteoarthritis based on the presence of clinical symptoms of knee osteoarthritis as defined by the American College of Rheumatology criteria Severity: Not stated Duration of symptoms: At least 6 months Presence of multimorbidities: High morbidity score	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Peloquin 1999[350]	**Supervised mixed modality exercise** (n=59) Aerobic, muscle strengthening and stretching exercises delivered over 12 weeks. Group or individual: Group session	**Knee osteoarthritis** Mean age (SD): 66.05 (7.89) years N = 124 Definition: Knee osteoarthritis confirmed by radiographs	Quality of life at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Type of exercise: Strength, aerobic **No treatment** (n=65) **Concomitant therapy:** No additional information	Severity: Grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 7.11 (7.03) years Presence of multimorbidities: Not stated/unclear		
Rogind 1998[382]	**Supervised mixed modality exercise** (n=12) Training focussed on general fitness, balance, coordination, stretching and lower extremity muscle strength, including a daily home exercise program. Training by physiotherapists 2 days per week for 3 months. Group or individual: Group session Type of exercise: Coordination, flexibility, strength **No treatment** (n=13) **Concomitant therapy:** As far as possible the medication was kept constant, apart from small changes in mild analgesics (paracetamol). No intra-articular or periarticular injections were given during the entire study period	**Knee osteoarthritis** Mean age (SD): 71.2 (7.4) years N = 25 Definition: Fulfilling the American College of Rheumatology criteria of osteoarthritis of the knee and the radiograph of the knee had to be rated at least 3 on the Kellgren scale Severity: At least grade 3 on the Kellgren Lawrence scale Duration of symptoms: Not stated Presence of multimorbidities:	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Takacs 2017[430]	**Supervised mixed modality exercise** (n=20) Dynamic balance training consisting of progressive exercise training over 3 phases, with exercises emphasizing dynamic balance control, eccentric lower limb muscle strength and core stability. Asked to perform all exercises 4 times a week for 10 weeks. Group or individual: Not stated/unclear Type of exercise: Strength, balance **No treatment** (n=20) **Concomitant therapy:** Co-interventions included prescription pain medication (n=1), physiotherapy (n=2); hydrotherapy (n=1), and exercise circuit training (n=1)	**Knee osteoarthritis** Mean age (SD): 66.6 (7.3) years N = 40 Definition: Radiographically confirmed tibiofemoral knee OA (Kellgren and Lawrence [KL] grade ≥2) and knee pain Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms: Not stated/unclear Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	
van Baar 2001[459]	**Supervised mixed modality exercise** (n=99) Exercises for muscle functions (strength and length), mobility, and coordination and exercises for elementary movement abilities and locomotion abilities	**Mixed osteoarthritis (knee or hip)** Mean age (SD): 68.0 (8.8) years N = 201 Definition: Osteoarthritis of the hip or knee according to the clinical criteria of the	Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	completed in 30 minute sessions for 12 weeks. Group or individual: Individual session Type of exercise: Strength, coordination **No treatment** (n=102) **Concomitant therapy:** GP provided patient education (including a brochure) and drug treatment, if necessary	American College of Rheumatology with radiographic confirmation Severity: Not stated Duration of symptoms: Median 1 year, no more than 6 years Presence of multimorbidities: Not stated/unclear		
Wang 2011[479]	**Supervised mixed modality exercise** (n=28) Land based exercise included flexibility and aerobic training, three times a week for 12 weeks Group or individual: Group session Type of exercise: Upper body, lower body training, flexibility and aerobic **Other supervised exercise** (n=28) Aquatic exercise with a flexibility and aerobic training class, three times a week for 12 weeks Group or individual: Group session Type of exercise: Hydrotherapy	**Knee osteoarthritis** Mean age (SD): 67.7 (5.9) years N = 82 Definition: Knee osteoarthritis diagnosed by physician assessment based on symptoms and X-ray Severity: Not stated Duration of symptoms (mean [SD]): 6.8 (6.4) years Presence of multimorbidities: Low morbidity score	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment (n=26) **Concomitant therapy:** No additional information			

1.1.5.19 Unsupervised mixed modality exercise compared to unsupervised strength exercise

Table 21: Summary of studies included in the evidence review for the unsupervised mixed modality exercise compared to unsupervised strength exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Chen 2021[83]	**Unsupervised mixed modality exercise** (n=16) Backwards walking training in addition to conventional training. Backwards walking for 10 minutes with 5 minutes of warm-up and cool-down sessions 3 days a week for 4 weeks at their comfortable walking speed. Group or individual: Individual session Type of exercise: Other (Strength and aerobic). **Unsupervised strength exercise** (n=16) No additional treatment. Duration 4 weeks Group or individual: Individual session	**Knee osteoarthritis** Mean age (SD): 60.6 (7.4) years N = 32 Definition: Knee osteoarthritis diagnosed by the American College of Rheumatology clinical criteria enrolled from outpatients of the hospital (including radiological evidence) Severity: Kellgren Lawrence grade at least 1 in one or both knees Duration of symptoms (SD): 37.3 (36.5) months Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Type of exercise: Not applicable **Concomitant therapy:** Conventional treatment comprising acupotomy, medications and routine exercise, once a week for 4 weeks. Based on the previous method, the subjects in both groups were treated with needle-knife therapy at the dominant inserted points. All were prescribed with an oral medication, Celebrex capsules (0.2g/d, once a day) for the first 6 days, while no extra painkillers were used in the next 3 weeks. Additionally, straight leg raising, as a routine exercise, was prescribed to practice at home for both legs, 1 set of 10 repetitions, twice a day, and gradually increase exercise time to 3 sets over the 4-week period, according to their pain intensity.			
Fitzgerald 2011[129] Subsidiary papers: Teixeira 2011[440]	**Unsupervised mixed modality exercise** (n=91) Lower extremity muscle stretching and strengthening, long-sitting knee flexion and extension range of motion, treadmill walking, and agility training and perturbation techniques Group or individual:	**Knee osteoarthritis** Mean age (SD): 64.0 (8.7) years N = 183 Definition: Knee osteoarthritis meeting the 1986 American College of Rheumatology clinical criteria	Pain at ≤3 months and >3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Individual session Type of exercise: Agility and perturbation **Unsupervised strength exercise** (n=92) Lower extremity muscle stretching and strengthening, long-sitting knee flexion and extension range of motion, and treadmill walking Group or individual: Individual session **Concomitant therapy:** All participants also were instructed to continue a walking program of at least 30 minutes per day for at least 3 days a week for the home program	with grade 2 or greater radiographic changes in the tibiofemoral joint Severity: Not stated explicitly. Kellgren and Lawrence grade 2 or more Duration of symptoms: Median 5-10 years Presence of multimorbidities: High morbidity score		
Gondhalekar 2013[158]	**Unsupervised mixed modality exercise** (n=15) Three sessions of Retro-walking per day for 3 weeks, plus free exercises such as hip flexion, hip abduction, knee bending and quadricep exercises, 10 repetitions, 1 set twice per day progressing to 3 sets in the third week Group or individual: Individual session Type of exercise: Strength and aerobic	**Knee osteoarthritis** Mean age (SD): 64.43 (6.2) years N = 30 Definition: People fulfilling three of the six clinical criteria listed by the American College of Rheumatology diagnosed as knee osteoarthritis confirmed using radiological investigations	Pain at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Unsupervised strength exercise (n=15) Ten repetitions, 1 set twice a day for the first week, progressing to 3 sets twice a day in the third week Group or individual: Individual session **Concomitant therapy:** Deep heating modality (short wave diathermy) 250W for 20 minutes	Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		

1.1.5.20 Unsupervised mixed modality exercise compared to other unsupervised exercise

Table 22: Summary of studies included in the evidence review for the unsupervised mixed modality exercise compared to other unsupervised exercise comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Petrella 2000[355]	**Unsupervised mixed modality exercise** (n=91) A series of progressive, simple, range of motion and resistance exercises utilizing common items at home over 8 weeks Group or individual: Individual session Type of exercise: Strength, range of motion/flexibility	**Knee osteoarthritis** Mean age (SD): 73.7 (4.9) years N = 179 Definition: Radiographic evidence of knee osteoarthritis in the tibial-femoral compartment (grade 1-3) Severity:	Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Other unsupervised exercise (n=88) Non-weight bearing joint unloading and stretches which did not include resistance of progression Group or individual: Individual session Type of exercise: Stretching **Concomitant therapy:** All people were given oxaprozin 1200mg/day during the study period. All people were given paracetamol 325mg to be taken every 4-6 hours as needed for rescue pain therapy	Kellgren Lawrence grade 1-3, median grade 1 Duration of symptoms: Not stated Presence of multimorbidities: High morbidity score		

1.1.5.21 Unsupervised mixed modality exercise compared to pharmacological treatments

Table 23: Summary of studies included in the evidence review for the unsupervised mixed modality exercise compared to pharmacological treatments comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Karatosun 2006[223]	**Unsupervised mixed modality exercise** (n=53) Progressive exercise programme, including strengthening, stretching, range of motions, resistive, and proprioceptive exercises, and advice for daily living activities, with new exercises added at different stages	**Knee osteoarthritis** Mean age (SD): 56.5 (12.9) years N = 105 Definition: Primary osteoarthritis of the knee as defined by the American College of	Pain at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Group or individual: Individual session Type of exercise: Proprioceptive, strengthening, flexibility, range of motion **Pharmacological treatment** (n=52) Three injections of hyaluronic acid separated by one-week intervals. In bilateral cases, both knees were injected Class of medicine: Intraarticular treatment **Concomitant therapy:** No treatment with non-steroidal anti-inflammatory drugs	Rheumatology criteria. All people had Kellgren Lawrence grade 3 osteoarthritis with narrowing of joint space and sclerosis of the subchondral bone Severity: Kellgren Lawrence grade 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Karatosun 2008[224]	**Unsupervised mixed modality exercise** (n=15) An exercise program taught over 6 weeks, including progressive, simple, isometric, isotonic range of motion, resistance, closed kinetic chain and proprioceptive exercises Group or individual: Individual session Type of exercise: Other **Pharmacological treatment** (n=15)	**Ankle osteoarthritis** Mean age (SD): 55.1 (12.1) years N = 30 Definition: Secondary ankle osteoarthritis defined by the clinical and radiographic findings (However, they ultimately ended up including people with primary osteoarthritis, 17 primary: 13 secondary)	Pain at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Three injections of hyaluronic acid separated by one-week intervals Class of medicine: Intraarticular treatment **Concomitant therapy:** No additional information	Severity: Kellgren Lawrence grade 3 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear		
Kawasaki 2009[227]	**Unsupervised mixed modality exercise** (n=60) Isometric muscle exercises of the lower limbs, range of motion exercises, and recommendation to walk as much as they could without pain Group or individual: Individual session Type of exercise: Strength and range of motion **Pharmacological treatment** (n=60) Intraarticular injections of hyaluronate sodium in the affected knee once a week for the first 5 weeks, followed by a once-a-month injection to maintain effects Class of medicine: Intraarticular treatment **Concomitant therapy:**	**Knee osteoarthritis** Mean age (SD): 70.4 (7.8) years N = 120 Definition: Primary osteoarthritis of the medial femorotibial compartment of the knee according to the clinical and radiographic criteria of the American College of Rheumatology Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/unclear	Pain at >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	All people were supplied with 100mg sodium loxoprofen tablets for pain rescue analgesia (300mg/day maximum allowed use) in the treated knee only			

1.1.5.22 Unsupervised mixed modality exercise compared to no treatment

Table 24: Summary of studies included in the evidence review for the unsupervised mixed modality exercise compared to no treatment comparison

Study	Intervention and comparison	Population	Outcomes	Comments
Allen 2018[12] Subsidiary papers: Andersen 2019 [20] Pignato 2018 [356]	**Unsupervised mixed modality exercise** (n=142) Internet based therapy, including tailored exercises, exercise progression recommendations, video display of exercises, automated reminders and progress tracking Group or individual: Individual session Type of exercise: Strengthening, stretching, aerobic **No treatment** (n=68) Waiting list control A third group (n=140) was reported but was not included in the analysis as they did not fulfil the inclusion criteria for this review'.	**Knee osteoarthritis** Mean age (SD): 65.3 (11.1) years N = 350 Definition: Radiographic evidence, physician diagnosis, or self-report of physician diagnosis based on the American College of Rheumatology clinical criteria Severity: Not stated Duration of symptoms (mean [SD]): 13.1 (11.7) years Presence of multimorbidities: Not stated/unclear	Pain at >3 months Physical function at >3 months Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: No additional information			
Teirlinck 2016[439]	**Unsupervised mixed modality exercise** (n=101) Group or individual: Individual session Type of exercise: Strengthening, flexibility, aerobic **No treatment** (n=102) GP care and a brochure with information about hip osteoarthritis **Concomitant therapy:** No additional information	**Hip osteoarthritis** Mean age (SD): 65.5 (9.2) years N = 203 Definition: Fulfilling the clinical criteria for hip osteoarthritis of the American College of Rheumatology Severity: Kellgren Lawrence grade 0-4, median grade 2 Duration of symptoms (median [IQR]): 365 (810-189) days Presence of multimorbidities: High morbidity score	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	

See Appendix D for full evidence tables.

1.1.5.23 Summary matrices

Table 25: Summary matrix for all interventions at ≤3 months

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
Supervised strength exercise	Unsupervised strength exercise	No evidence identified	1 GRADE Outcome (2 studies) **N = 115** Moderate	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	1 GRADE Outcomes (2 studies) **N = 66** Very Low	1 GRADE Outcome (1 study) **N = 29** Very Low	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	11 GRADE Outcomes (6 studies) **N = 534** Moderate-Very Low	2 GRADE Outcomes (13 studies) **N = 797** Very Low	2 GRADE Outcomes (12 studies) **N = 499** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 108** Very Low
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	9 GRADE Outcomes (4 studies) **N = 229** Low-Very Low	1 GRADE Outcome (10 studies) **N = 525** Very Low	1 GRADE Outcome (9 studies) **N = 543** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (3 studies) **N = 113**

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
							Very Low
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	1 GRADE Outcome (1 study) **N = 166** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	14 GRADE Outcomes (14 studies) **N = 1045** Low-Very Low	2 GRADE Outcomes (30 studies) **N = 2214** Very Low	2 GRADE Outcomes (23 studies) **N = 1632** Very Low	2 GRADE Outcomes (1 study) **N = 121** Low	No evidence identified	1 GRADE Outcome (3 studies) **N = 180** Very Low
Unsupervised strength exercise	Supervised strength exercise	No evidence identified	1 GRADE Outcome (2 studies) **N = 115** Moderate	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	6 GRADE Outcomes (1 study) **N = 55** Low-Very Low	1 GRADE Outcome (1 study) **N = 55** Very Low	1 GRADE Outcome (1 study) **N = 55** Very Low	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
	Other supervised exercise	2 GRADE Outcomes (1 study) **N = 44** Low-Very Low	1 GRADE Outcome (4 studies) **N = 280** Low-Very Low	1 GRADE Outcome (1 study) **N = 120** Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 80** Very Low
	Other unsupervised exercise	1 GRADE Outcome (1 study) **N = 42** Very Low	1 GRADE Outcome (1 study) **N = 42** Very Low	1 GRADE Outcome (1 study) **N = 42** Very Low	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	2 GRADE Outcomes (1 study) **N = 42** Low	1 GRADE Outcome (1 study) **N = 42** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	2 GRADE Outcomes (3 studies) **N = 221** Very Low	1 GRADE Outcome (2 studies) **N = 191** Very Low	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	7 GRADE Outcomes (3 studies) **N = 324** Low-Very Low	2 GRADE Outcomes (9 studies) **N = 821** Low-Very Low	2 GRADE Outcomes (8 studies) **N = 680** Moderate-Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 89** Moderate

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
Supervised aerobic exercise	Supervised strength exercise	No evidence identified	1 GRADE Outcome (2 studies) **N = 66** Very Low	1 GRADE Outcome (1 study) **N = 29** Very Low	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 40** Very low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	1 GRADE Outcome (1 study) **N = 28**	2 GRADE Outcomes (2 studies) **N = 55**	2 GRADE Outcomes (2 studies) **N = 55**	No evidence identified	No evidence identified	1 GRADE

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
		Very Low	Low-Very Low	Very Low			Outcome (1 study) **N = 37** Very Low
Unsupervised aerobic exercise	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	6 GRADE Outcomes (1 study) **N = 55** Low-Very Low	1 GRADE Outcome (1 study) **N = 55** Very Low	1 GRADE Outcome (1 study) **N = 55** Very Low	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	7 GRADE Outcomes (2 studies) **N = 219** Moderate-Very Low	1 GRADE Outcome (3 studies) **N = 286** Very Low	1 GRADE Outcome (3 studies) **N = 284** Very Low	2 GRADE Outcomes (1 study) **N = 164** Moderate	No evidence identified	No evidence identified
Other supervised exercise	Supervised strength exercise	11 GRADE Outcomes (6 studies) **N = 534** Moderate-Very Low	2 GRADE Outcomes (13 studies) **N = 797** Very Low	2 GRADE Outcomes (12 studies) **N = 499** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 108** Very Low
	Unsupervised strength exercise	2 GRADE Outcomes (1 study) **N = 44** Low-Very Low	2 GRADE Outcomes (4 studies) **N = 280** Low-Very Low	1 GRADE Outcome (1 study) **N = 120** Moderate	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 80** Very Low
	Supervised aerobic exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 40** Very low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	4 GRADE Outcomes (4 studies) **N = 248** Moderate-Very Low	1 GRADE Outcome (6 studies) **N = 334** Low	1 GRADE Outcome (4 studies) **N = 224** Very Low	2 GRADE Outcomes (1 study) **N = 60** Very Low	No evidence identified	1 GRADE Outcome (1 study) **N = 90** Very Low
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	6 GRADE Outcomes (11 studies) **N = 648** Very Low	2 GRADE Outcomes (21 studies) **N = 1217** Moderate-Very Low	2 GRADE Outcomes (17 studies) **N = 979** Very Low	3 GRADE Outcomes (3 studies) **N = 359** Very Low	No evidence identified	1 GRADE Outcome (4 studies) **N = 180** Very Low
Other unsupervised exercise	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	1 GRADE Outcome (1 study) **N = 42** Very Low	1 GRADE Outcome (1 study) **N = 42** Very Low	1 GRADE Outcome (1 study) **N = 42** Very Low	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 179** Low	1 GRADE Outcome (1 study) **N = 179** Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 179** Very Low
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
Supervised mixed modality exercise	Supervised strength exercise	9 GRADE Outcomes (4 studies) **N = 229** Very Low	1 GRADE Outcome (10 studies) **N = 525** Very Low	1 GRADE Outcome (9 studies) **N = 543** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (3 studies) **N = 113**

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
							Very Low
	Unsupervised strength exercise	2 GRADE Outcomes (1 study) **N = 42** Low	1 GRADE Outcome (1 study) **N = 42** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	4 GRADE Outcomes (4 studies) **N = 248** Moderate-Very Low	1 GRADE Outcome (6 studies) **N = 334** Low	1 GRADE Outcome (4 studies) **N = 224** Very Low	2 GRADE Outcomes (1 study) **N = 60** Very Low	No evidence identified	1 GRADE Outcome (1 study) **N = 90** Very Low
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	8 GRADE Outcomes (1 study) **N = 80** Low-Very Low	1 GRADE Outcome (2 studies) **N = 140** Very Low	1 GRADE Outcome (1 study) **N = 80** Very Low	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	1 GRADE Outcome (1 study) **N = 93**	2 GRADE Outcomes (2 studies) **N = 197**	2 GRADE Outcomes (2 studies) **N = 197**	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
		Low	Low-Very Low	Moderate-Very Low			
	No treatment	23 GRADE Outcomes (8 studies) **N = 500** Low-Very Low	2 GRADE Outcomes (12 studies) **N = 807** Very Low	2 GRADE Outcomes (8 studies) **N = 531** Very Low	3 GRADE Outcomes (3 studies) **N = 173** Low-Very Low	No evidence identified	1 GRADE Outcome (3 studies) **N = 284** Very Low
Unsupervised mixed modality exercise	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	No evidence identified	2 GRADE Outcomes (3 studies) **N = 221** Very Low	1 GRADE Outcome (2 studies) **N = 191** Very Low	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 179** Low	1 GRADE Outcome (1 study) **N = 179** Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 179**

Intervention	Control	Quality of life at ≤3 months	Pain at ≤3 months	Physical function at ≤3 months	Psychological distress at ≤3 months	Osteoarthritis flares at ≤3 months	Serious adverse events at ≤3 months
							Very Low
	Supervised mixed modality exercise	9 GRADE Outcomes (1 study) **N = 80** Low-Very Low	1 GRADE Outcome (2 studies) **N = 140** Very Low	1 GRADE Outcome (1 study) **N = 80** Very Low	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	1 GRADE Outcome (1 Study) **N = 203** Very Low	1 GRADE Outcome (1 Study) **N = 203** Very Low	1 GRADE Outcome (1 Study) **N = 203** Very Low	No evidence identified	No evidence identified	No evidence identified

Table 26: Summary matrix for all interventions at >3 months

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
Supervised strength exercise	Unsupervised strength exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 36** Moderate	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	2 GRADE Outcomes (2 studies) **N = 161** Very Low	1 GRADE Outcome (1 study) **N = 100** Very Low	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	10 GRADE Outcomes (3 studies) **N = 164** Low-Very Low	1 GRADE Outcome (3 studies) **N = 166** Very low	1 GRADE Outcome (1 study) **N = 66** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 64** Very Low
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	1 GRADE Outcome (1 study) **N = 66** Very Low	1 GRADE Outcome (3 studies) **N = 268** Very Low	1 GRADE Outcome (3 studies) **N = 268** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 113** Very low
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	3 GRADE Outcomes (3 studies) **N = 468** Low-Very Low	1 GRADE Outcome (7 studies) **N = 842** Very Low	2 GRADE Outcomes (4 studies) **N = 580**	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
				Very Low			
Unsupervised strength exercise	Supervised strength exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 36** Moderate	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 36** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 142** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	9 GRADE Outcomes (2 studies) **N = 321**	2 GRADE Outcomes (4 studies) **N = 1934**	2 GRADE Outcomes (4 studies)	2 GRADE Outcomes (1 study) **N = 191**	No evidence identified	1 GRADE Outcome (1 study)

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
		Low-Very Low	Moderate-Very Low	**N = 1934** Moderate-Very Low	Low-Very Low		**N = 130** High
Supervised aerobic exercise	Supervised strength exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 100** Very Low	1 GRADE Outcome (1 study) **N = 100** Very Low	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 79** Low	1 GRADE Outcome (1 study) **N = 79** Low	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	2 GRADE Outcomes (2 studies) **N = 208** Low	1 GRADE Outcome (2 studies) **N = 206** Moderate	1 GRADE Outcome (2 studies) **N = 206** Moderate	No evidence identified	No evidence identified	No evidence identified
Unsupervised aerobic exercise	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	1 GRADE Outcome (1 study) **N = 146** Low	1 GRADE Outcome (1 study) **N = 147** Low	1 GRADE Outcome (1 study) **N = 147** Low	2 GRADE Outcomes (1 study) **N = 147** Low-Very Low	No evidence identified	No evidence identified
Other supervised exercise	Supervised strength exercise	10 GRADE Outcomes (3 studies) **N = 164** Low-Very Low	1 GRADE Outcome (3 studies) **N = 166** Very low	1 GRADE Outcome (1 study) **N = 66** Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 64** Low
	Unsupervised strength exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 36** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	2 GRADE Outcomes (3 studies) **N = 613** Low-Very Low	2 GRADE Outcomes (5 studies) **N = 745** Low-Very Low	2 GRADE Outcomes (4 studies) **N = 706** Low-Very Low	1 GRADE Outcome (1 study) **N = 214** Low	No evidence identified	1 GRADE Outcome (1 study) **N = 87** Very Low
Other unsupervised exercise	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Supervised mixed modality exercise	No evidence identified	1 GRADE Outcome (2 study) **N = 149** Very Low	1 GRADE Outcome (1 study) **N = 85** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 88** Low
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	No treatment	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
Supervised mixed modality exercise	Supervised strength exercise	1 GRADE Outcome (1 study) **N = 66** Very Low	1 GRADE Outcome (3 studies) **N = 268** Very Low	1 GRADE Outcome (3 studies) **N = 268** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 113** Very low
	Unsupervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 79** Low	1 GRADE Outcome (1 study) **N = 79** Low	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	1 GRADE Outcome (2 study) **N = 149** Very Low	1 GRADE Outcome (1 study) **N = 85** Very Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 88** Low
	Unsupervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	1 GRADE Outcome (1 study) **N = 93** Moderate	2 GRADE Outcomes (2 studies) **N = 197** Very Low	2 GRADE Outcomes (2 studies) **N = 197** Low	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 110** Low
	No treatment	No evidence identified	2 GRADE Outcomes (4 studies) **N = 416** Low-Very Low	1 GRADE Outcome (1 study) **N = 107** Very Low	1 GRADE Outcome (1 study) **N = 106** Very Low	No evidence identified	1 GRADE Outcome (1 study) **N = 102** Moderate
Unsupervised mixed	Supervised strength exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
modality exercise	Unsupervised strength exercise	No evidence identified	1 GRADE Outcome (1 study) **N = 142** Low	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Unsupervised aerobic exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other supervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Other unsupervised exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Supervised mixed modality exercise	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified	No evidence identified
	Pharmacological treatment	No evidence identified	2 GRADE Outcomes (3 studies) **N = 255** Low-Very Low	No evidence identified	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 120** Moderate
	No treatment	1 GRADE Outcome (1 study) **N = 203** Very Low	2 GRADE Outcomes (2 studies) **N = 413** Moderate-Low	2 GRADE Outcomes (2 studies) **N = 413**	No evidence identified	No evidence identified	1 GRADE Outcome (1 study) **N = 210**

Intervention	Control	Quality of life at >3 months	Pain at >3 months	Physical function at >3 months	Psychological distress at >3 months	Osteoarthritis flares at >3 months	Serious adverse events at >3 months
				Low			Very Low

1.1.6 Summary of the effectiveness evidence

1.1.6.1 Supervised strength exercise compared to unsupervised strength exercise, supervised aerobic exercise and no treatment

Table 27: Clinical evidence summary: supervised strength exercise compared to unsupervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with supervised strength exercise	
Pain (VAS, 0-100, high is poor, final value) at ≤3 months	115 (2 RCTs) follow up: mean 7 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 44	MD **19.77 lower** (22.32 lower to 17.23 lower)	MID = 16.9 (0.5 x median baseline SD)
Pain (VAS, 0-10, high is poor, final value) at >3 months	36 (1 RCT) follow up: 6 months	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 3.1	MD **2.3 lower** (2.47 lower to 2.13 lower)	MID = 0.5 SD (SMD)

[a] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Table 28: Clinical evidence summary: supervised strength exercise compared to supervised aerobic exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised aerobic exercise	Risk difference with supervised strength exercise	
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months	66 (2 RCTs) follow up: mean 7 weeks	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.45 SD higher** (0.04 lower to 0.94 higher)	MID = 0.5 SD (SMD)
Pain (Arthritis Self-Efficacy pain subscale, 0-100, high is poor) at >3 months	100 (1 RCT)	⊕◯◯◯ VERY LOW [a,b]	-	-	MD **11.1 higher** (0.1 higher to 22.1 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised aerobic exercise	Risk difference with supervised strength exercise	
	follow up: 12 months					
Pain (Intermittent and constant osteoarthritis pain total pain, 0-20, high is poor, final value) at >3 months	30 (1 RCT) follow up: 18 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was 10.1	MD **2.17 lower** (8.41 lower to 4.07 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	29 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 14.57	MD **1.51 higher** (6.88 lower to 9.9 higher)	MID = 0.5 SD (SMD)
Physical function (Arthritis Self-Efficacy function subscale, 0-100, high is poor) at >3 months	100 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,b]	-		MD **7.6 higher** (0.7 higher to 14.5 higher)	MID = 0.5 SD (SMD)

[a.] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b.] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 29: Clinical evidence summary: supervised strength exercise compared to pharmacological treatment

Outcomes	№ of participants (studies) Follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with pharmacological treatment	Risk difference with supervised strength exercise	
Quality of life (SF-36 total, scale range unclear, high is good, final values) at ≤3 months	166 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean health related quality of life was 83.4	MD **22 higher** (17.5 higher to 26.5 higher)	MID = 0.5 SD (SMD)

[a.] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Table 30: Clinical evidence summary: supervised strength exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised strength exercise	
Quality of life (KOOS, 0-100, high is good, change scores) at ≤3 months	98 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was -2.25	MD **15.94 higher** (4.44 lower to 36.32 higher)	MID = 5.92 (0.5 x median baseline SD)
Quality of life (EQ-5D, KOOS, HOOS, Assessment of Quality of Life Scale, AIMS [different scale ranges], high is good, final values) at ≤3 months	569 (6 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.42 SD higher** (0.01 lower to 0.86 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component summary, SF-12 physical score, 0-100, high is good, final values) at ≤3 months	218 (4 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 45.01	MD **5.33 higher** (8.19 lower to 18.85 higher)	MID = 6.6 (0.5 x median baseline SD)
Quality of life (SF-36 mental component summary, SF-12 mental score, 0-100, high is good, final values) at ≤3 months	218 (4 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 57.3	MD **9.45 higher** (0. 79 higher to 18.11 higher)	MID = 9.5 (0.5 x median baseline SD)
Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 18.9	MD **16.35 higher** (9.1 higher to 23.61 higher)	MID = 3 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised strength exercise	
Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 23.92	MD **14.47 higher** (5.21 higher to 23.73 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 15.2	MD **26.19 higher** (11.79 higher to 40.58 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 27.8	MD **9.83 higher** (0.44 higher to 19.22 higher)	MID = 2 (established value)
Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 29.3	MD **7.57 higher** (3.53 lower to 18.67 higher)	MID = 2 (established value)
Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 30.4	MD **10.12 higher** (3.98 lower to 24.22 higher)	MID = 3 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 31.4	MD **16.9 higher** (0.14 higher to 33.67 higher)	MID = 4 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised strength exercise	
Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at ≤3 months	160 (2 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 29.0	MD **15.4 higher** (4.24 higher to 26.56 higher)	MID = 3 (established value)
Quality of life (EQ-5D, KOOS [different scale ranges], high is good, final values) at >3 months	407 (2 RCTs) follow up: mean 15 months	⊕⊕◯◯ LOW [a]	-	-	SMD **0.06 SD lower** (0.25 lower to 0.14 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component summary, 0-100, high is good, change score) at >3 months	61 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was -0.7	MD **3.7 higher** (12.19 lower to 19.59 higher)	MID = 2 (established value)
Quality of life (SF-36 mental component summary, 0-100, high is good, change score) at >3 months	61 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was -0.7	MD **1.4 higher** (19.26 lower to 22.06 higher)	MID = 3 (established value)
Pain (KOOS, WOMAC, NRS, VAS [different scale ranges], high is poor, change scores) at ≤3 months	531 (7 RCTs) follow up: mean 8 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.55 SD lower** (0.73 lower to 0.37 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, HOOS, AUSCAN, WOMAC, NRS, VAS [different scale ranges], high is poor, final values) at ≤3 months	1733 (23 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.81 SD lower** (1.06 lower to 0.57 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, VAS [different scale ranges], high is poor,	842 (7 RCTs)	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **1 SD lower** (1.76 lower to 0.23 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised strength exercise	
change score and final values) at >3 months	follow up: mean 11 months					
Physical function (KOOS, WOMAC [different scale ranges], high is poor, change scores) at ≤3 months	301 (5 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.58 SD lower** (0.83 lower to 0.33 lower)	MID = 0.5 SD (SMD)
Physical function (KOOS, HOOS, AUSCAN, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at ≤3 months	1381 (19 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **1 SD lower** (1.37 lower to 0.63 lower)	MID = 0.5 SD (SMD)
Physical function (KOOS ADL, 0-100, high is good, change score) at >3 months	61 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was -1.9	MD **2.8 higher** (4.36 lower to 9.96 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at >3 months	519 (3 RCTs) follow up: mean 10 months	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.31 SD lower** (1.09 lower to 0.48 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months	121 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 6.54	MD **0.54 higher** (1.1 lower to 2.18 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months	121 (1 RCT)	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 7.13	MD **0.38 lower** (1.7 lower to 0.94 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised strength exercise	
	follow up: 12 weeks					
Serious adverse events at ≤3 months	180 (3 RCTs) follow up: mean 9 weeks	⨁◯◯◯ VERY LOW [d,e]	RD 0.06 (0.00 to 0.12)	0 per 1,000	**60 more per 1,000** (120 more to 0 fewer) [f]	Precision calculated through Optimal Information Size (OIS) due to zero events in some studies (0.8-0.9 = serious, <0.8 = very serious).

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis
d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)
e. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size
f. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

1.1.6.2 Unsupervised strength exercise compared to unsupervised aerobic exercise and no treatment

Table 31: Clinical evidence summary: unsupervised strength exercise compared to unsupervised aerobic exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised aerobic exercise	Risk difference with unsupervised strength exercise	
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 8.6	MD **20.9 higher** (18.56 higher to 23.24 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised aerobic exercise	Risk difference with unsupervised strength exercise	
Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 9	MD **0.8 higher** (0.89 lower to 2.49 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 14.6	MD **18.8 higher** (17.87 higher to 19.73 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 19.6	MD **12.3 higher** (9.93 higher to 14.67 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 6.9	MD **12.2 higher** (10.81 higher to 13.59 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 17.3	MD **0.2 lower** (2.32 lower to 1.92 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean pain was 3.4	MD **0.4 lower** (1.2 lower to 0.4 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	55 (1 RCT) follow up: 12 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean physical function was 10.2	MD **0.6 higher** (0.52 lower to 1.72 higher)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised aerobic exercise	Risk difference with unsupervised strength exercise	
b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs						

Table 32: Clinical evidence summary: unsupervised strength exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised strength exercise	
Quality of life (EQ-5D, Arthritis Impact Measurement Scale 2 - Short form [different scale ranges], high is good, final values) at ≤3 months	271 (2 RCTs) follow up: mean 12 weeks	⊕◯◯◯ VERY LOW a,b,c	-	-	SMD **0.20 SD higher** (0.23 lower to 0.63 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW a	-	The mean quality of life was 36.6	MD **7.1 lower** (10.06 lower to 4.14 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW a	-	The mean quality of life was 20.4	MD **10.6 lower** (12.3 lower to 8.9 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW a	-	The mean quality of life was 49.3	MD **15.9 lower** (16.93 lower to 14.87 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW a,c	-	The mean quality of life was 35.3	MD **3.4 lower** (5.91 lower to 0.89 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised strength exercise	
Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 27.9	MD **8.8 lower** (10.72 lower to 6.88 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months	53 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 19.2	MD **2.1 lower** (4.48 lower to 0.28 higher)	MID = 0.5 SD (SMD)
Quality of life (EQ-5D, 0-1, high is good, final value) at >3 months	130 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a,c]	-	The mean quality of life was 0.634	MD **0.07 higher** (0.00 lower to 0.14 higher)	MID = 0.03 (established value)
Quality of life (SF-36 physical functioning, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was -1.63	MD **4.31 higher** (0.41 lower to 9.03 higher)	MID = 3 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 0.16	MD **4.81 higher** (2.3 lower to 11.92 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was -7.59	MD **10.78 higher** (0.54 lower to 22.1 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 0.56	MD **1.91 higher** (3.53 lower to 7.35 higher)	MID = 2 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised strength exercise	
Quality of life (SF-36 general health, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was -0.7	MD **2.63 higher** (1.55 lower to 6.81 higher)	MID = 2 (established value)
Quality of life (SF-36 mental health, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was -2.91	MD **2.7 higher** (1.8 lower to 7.2 higher)	MID = 3 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 0.48	MD **1.37 higher** (14.87 lower to 17.61 higher)	MID = 4 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, change score) at >3 months	191 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 1.9	MD **0.01 lower** (10.3 lower to 10.28 higher)	MID = 3 (established value)
Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at ≤3 months	379 (5 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	-	SMD **1.1 SD lower** (1.32 lower to 0.88 lower)	MID = 0.5 SD (SMD)
Pain (WOMAC, NRS [different scale ranges], high is poor, final values) at ≤3 months	442 (4 RCTs) follow up: mean 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.37 SD lower** (0.81 lower to 0.08 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at >3 months	1686 (2 RCTs)	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.08 SD lower** (0.18 lower to 0.01 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised strength exercise	
	follow up: mean 6 months					
Pain (WOMAC, NRS, 0-100, high is poor, final values) at >3 months	248 (2 RCTs) follow up: 15 months	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 35.5	MD **1.75 lower** (7.31 lower to 3.8 higher)	MID = 9.5 (0.5 x median baseline SD)
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, change scores) at ≤3 months	379 (5 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	-	SMD **0.93 SD lower** (1.14 lower to 0.72 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at ≤3 months	301 (3 RCTs) follow up: mean 13 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.85 SD lower** (2.15 lower to 0.44 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months	1686 (2 RCTs) follow up: mean 6 months	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.1 SD lower** (0.2 lower to 0.01 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at >3 months	248 (2 RCTs) follow up: mean 15 months	⊕⊕⊕◯ MODERATE [a]	-	-	SMD **0.06 SD lower** (0.31 lower to 0.19 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS anxiety, 0-21, high is poor, change score) at >3 months	191 (1 RCT)	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 0.06	MD **0.63 lower** (1.54 lower to 0.28 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised strength exercise	
	follow up: 6 months					
Psychological distress (HADS depression, 0-21, high is poor, change score) at >3 months	191 (1 RCT) follow up: 6 months	⨁◯◯◯ VERY LOW [a,c]	-	The mean psychological distress was 0.11	MD **0.68 lower** (1.3 lower to 0.06 lower)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	89 (1 RCT) follow up: 12 weeks	⨁⨁⨁◯ MODERATE [a]	Peto OR 7.94 (1.32 to 47.77)	0 per 1,000	**110 more per 1,000** (from 10 more to 210 more) [d]	MID (precision) = Peto OR 0.8-1.25.
Serious adverse events at >3 months	130 (1 RCT) follow up: 6 months	⨁⨁⨁⨁ HIGH	Peto OR 8.29 (1.99 to 34.46)	0 per 1,000	**120 more per 1,000** (from 40 more to 210 more) [d]	MID (precision) = Peto OR 0.8-1.25.

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

c. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

1.1.6.3 Supervised aerobic exercise compared to no treatment

Table 33: Clinical evidence summary: supervised aerobic exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised aerobic exercise	
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months	28 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 3.8	MD **6.8 higher** (6.32 lower to 19.92 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component, 0-100, high is good, change score and final value) at >3 months	208 (2 RCTs) follow up: mean 17 months	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 24.8	MD **1.15 lower** (3.41 lower to 1.11 higher)	MID = 2 (established value)
Quality of life (SF-36 mental component, 0-100, high is good, change score) at >3 months	208 (2 RCTs) follow up: mean 17 months	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 27.2	MD **1.18 lower** (3.46 lower to 1.11 higher)	MID = 3 (established value)
Pain (KOOS, 0-100, high is good, change score) at ≤3 months	28 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was -0.9	MD **13.3 higher** (2.97 higher to 23.63 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months	27 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean pain was 7.31	MD **4.02 lower** (6.01 lower to 2.03 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months	206 (2 RCTs) follow up: mean 17 months	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 16.1	MD **1.3 higher** (3 lower to 5.59 higher)	MID = 8.1 (0.5 x median baseline SD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised aerobic exercise	
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months	28 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 0.8	MD **11.1 higher** (2.9 lower to 25.1 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	27 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 29.92	MD **15.35 lower** (24.02 lower to 6.68 lower)	MID = 0.5 SD (SMD)
Physical function (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months	206 (2 RCTs) follow up: mean 17 months	⊕⊕⊕◯ MODERATE [a]	-	The mean physical function was 12.8	MD **1.87 lower** (5.98 lower to 2.24 higher)	MID = 8.0 (0.5 x median baseline SD)
Serious adverse events at ≤3 months	37 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	Peto OR 7.86 (0.77 to 80.77)	0 per 1,000	**160 more per 1,000** (20 fewer to 340 more) [d]	MID (precision) = Peto OR 0.8-1.25.

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

1.1.6.4 Unsupervised aerobic exercise compared to no treatment

Table 34: Clinical evidence summary: unsupervised aerobic exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised aerobic exercise	
Quality of life (KOOS, 0-100, high is good, final value) at ≤3 months	165 (1 RCT) follow up: 13 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean quality of life was 47.3	MD **2.1 higher** (8.86 lower to 13.06 higher)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months	54 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 36.6	MD **28 lower** (30.77 lower to 25.23 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months	54 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 20.4	MD **11.4 lower** (13.13 lower to 9.67 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months	54 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 49.3	MD **34.7 lower** (35.51 lower to 33.89 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months	54 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 35.3	MD **15.7 lower** (17.95 lower to 13.45 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months	54 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 27.9	MD **21 lower** (23.06 lower to 18.94 lower)	MID = 0.5 SD (SMD)
Quality of life (Nottingham Health Profile social isolation subscale,	54 (1 RCT)	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 19.2	MD **1.9 lower** (4.21 lower to 0.41 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised aerobic exercise	
0-100, high is poor, final value) at ≤3 months	follow up: 12 weeks					
Quality of life (KOOS, 0-100, high is good, final value) at >3 months	146 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 47.5	MD **1.2 higher** (10.14 lower to 12.54 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months	286 (3 RCTs) follow up: mean 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **1.49 SD lower** (3.11 lower to 0.14 higher)	MID = 0.5 SD (SMD)
Pain (VAS, 0-10, high is poor, final value) at >3 months	147 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean pain was 3.8	MD **0.3 lower** (1.82 lower to 1.22 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months	284 (3 RCTs) follow up: mean 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **2.1 SD lower** (4.38 lower to 0.18 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-100, high is good, final value) at >3 months	147 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean physical functioning was 62.9	MD **5 higher** (7.45 lower to 17.45 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months	164 (1 RCT) follow up: 13 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean psychological distress was 4.2	MD **0.7 lower** (2.16 lower to 0.76 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised aerobic exercise	
Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months	164 (1 RCT) follow up: 13 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean psychological distress was 3.2	MD **0.6 lower** (2.16 lower to 0.96 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at >3 months	147 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 4.1	MD **1 lower** (2.63 lower to 0.63 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS depression, 0-21, high is poor, final value) at >3 months	147 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 3	MD **0.6 lower** (2.23 lower to 1.03 higher)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

1.1.6.5 Other supervised exercise compared to supervised strength exercise, unsupervised strength exercise and no treatment

Table 35: Clinical evidence summary: other supervised exercise compared to supervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Quality of life (KOOS, Assessment of Quality of Life Instrument version two, WHO Quality of Life total [different scale ranges], high is good, final values) at ≤3 months	148 (3 RCTs) follow up: mean 9 weeks	⊕⊕◯◯ LOW [a]	-	-	SMD **0 SD** (0.32 lower to 0.32 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Quality of life (SF-12 physical score, 0-100, high is good, final value) at ≤3 months	70 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 31.4	MD **5.7 higher** (0.25 lower to 11.65 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 mental component, SF-12 mental score, 0-100, high is good, final values) at ≤3 months	136 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 56.8	MD **4.97 lower** (9.23 lower to 0.7 lower)	MID = 6.6 (0.5 x median control group baseline SD)
Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 50.94	MD **16.56 higher** (13.52 higher to 19.6 higher)	MID = 3 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 46.93	MD **26.84 higher** (23.87 higher to 29.81 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 58.33	MD **28.11 higher** (19.78 higher to 36.44 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 53.2	MD **16.85 lower** (18.46 lower to 15.24 lower)	MID = 2 (established value)
Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 60.12	MD **17.35 higher** (13.07 higher to 21.63 higher)	MID = 2 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 52.27	MD **17.94 lower** (19.35 lower to 16.53 lower)	MID = 3 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 57.75	MD **27.66 higher** (20.17 higher to 35.15 higher)	MID = 4 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months	250 (1 RCT) follow up: 12 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 57.15	MD **6.89 higher** (4.49 higher to 9.29 higher)	MID = 3 (established value)
Quality of life (SF-36 mental component, 0-100, high is good, final value) at >3 months	66 (1 RCT) follow up: 15 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 51.9	MD **0.7 lower** (6.16 lower to 4.76 higher)	MID = 3 (established value)
Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 50.8	MD **6.7 higher** (13.31 lower to 26.71 higher)	MID = 3 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 54.8	MD **4.5 higher** (8.97 lower to 17.97 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⨁◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 51.4	MD **3.4 higher** (8.88 lower to 15.68 higher)	MID = 3 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 60.3	MD **4.2 higher** (5.21 lower to 13.61 higher)	MID = 2 (established value)
Quality of life (SF-36 general health, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 62	MD **1.2 lower** (11.35 lower to 8.95 higher)	MID = 2 (established value)
Quality of life (SF-36 mental health, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 65.6	MD **8.5 higher** (0.54 lower to 17.54 higher)	MID = 3 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 64.6	MD **3.5 lower** (24.37 lower to 17.37 higher)	MID = 4 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 67.3	MD **6.7 higher** (5.8 lower to 19.2 higher)	MID = 3 (established value)
Quality of life (WHO Quality of Life Total, 0-100, high is good, final value) at >3 months	34 (1 RCT) follow up: 6 months	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 96.24	MD **1.94 higher** (2.22 lower to 6.1 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score) at ≤3 months	15 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was -4	MD **1.33 higher** (3.39 lower to 6.05 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Pain (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months	797 (13 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.18 SD lower** (0.43 lower to 0.79 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months	166 (3 RCTs) follow up: mean 19 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.37 SD higher** (0.03 higher to 0.71 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, change score) at ≤3 months	30 (2 RCT) follow up: mean 7 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was -12.1	MD **4.92 lower** (13.86 lower to 4.02 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months	469 (10 RCTs) follow up: mean 8 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.03 SD lower** (0.4 lower to 0.33 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months	66 (1 RCT) follow up: 15 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 32.2	MD **0.4 higher** (5.18 lower to 5.98 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	108 (2 RCT) follow up: mean 11 weeks	⊕◯◯◯ VERY LOW [a,b,d]	RD 0.05 (-0.11 to 0.20)	227 per 1,000	**50 more per 1,000** (110 fewer to 200 more) [e]	Precision calculated through Optimal Information Size (OIS) due to zero events in some studies (0.8-0.9 = serious, <0.8 = very serious).

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with other supervised exercise	
Serious adverse events at >3 months	64 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,b]	Peto OR 7.39 (0.15 to 372.38)	0 per 1,000	**30 more per 1,000** (50 fewer to 110 more) [e]	MID (precision) = Peto OR 0.8-1.25.

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis
d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)
e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 36: Clinical evidence summary: other supervised exercise compared to unsupervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with other supervised exercise	
Quality of life (SF-36 physical component, 0-100, high is good, final value) at ≤3 months	44 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 36.9	MD **1.9 higher** (3.31 lower to 7.11 higher)	MID = 2 (established value)
Quality of life (SF-36 mental component, 0-100, high is good, final value) at ≤3 months	44 (1 RCT) follow up: 8 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 48.4	MD **6.4 higher** (0.79 lower to 13.59 higher)	MID = 3 (established value)
Pain (NRS, 0-10, high is poor, change score) at ≤3 months	80 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was -1.8	MD **0.5 lower** (1.29 lower to 0.29 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with other supervised exercise	
Pain (WOMAC, NRS [different scale ranges], high is poor, final scores) at ≤3 months	200 (3 RCTs) follow up: mean 8 weeks	⨁⨁◯◯ LOW [a]	-	-	SMD **1.03 SD lower** (1.33 lower to 0.74 lower)	MID = 0.5 SD (SMD)
Pain (VAS, 0-10, high is poor, final value) at >3 months	36 (1 RCT) follow up: 6 months	⨁⨁◯◯ LOW [a,b]	-	The mean pain was 3.1	MD **0.2 lower** (0.33 lower to 0.07 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	120 (1 RCT) follow up: 8 weeks	⨁⨁◯◯ LOW [a]	-	The mean physical function was 46.9	MD **20.8 lower** (26.68 lower to 14.92 lower)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	80 (1 RCT) follow up: 4 weeks	⨁◯◯◯ VERY LOW [a,b]	RR 0.57 (0.27 to 1.21)	350 per 1,000	**151 fewer per 1,000** (256 fewer to 73 more)	MID (precision) = RR 0.8-1.25.

[a] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 37: Clinical evidence summary: other supervised exercise compared to supervised aerobic exercise

Outcomes	№ of participants (studies) Follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised aerobic exercise	Risk difference with other supervised exercise	
Pain (VAS, 0-100, high is poor, final value) at ≤3 months	40 (1 RCT) follow-up: 2 months	⨁◯◯◯ VERY LOW [a,b]	-	The mean pain was 60.3	MD **20.5 lower** (40.01 lower to 0.99 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised aerobic exercise	Risk difference with other supervised exercise	

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 38: Clinical evidence summary: other supervised exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with other supervised exercise	
Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at ≤3 months	257 (4 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW a,b,c	-	-	SMD **0.44 SD higher** (0.14 lower to 1.02 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change scores and final values) at ≤3 months	370 (6 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW a,b,c	-	The mean quality of life was 24.2	MD **4 higher** (0.56 higher to 7.44 higher)	MID = 4.95 (0.5 x median baseline SD)
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change scores and final values) at ≤3 months	370 (6 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW a,b,c	-	The mean quality of life was 35.6	MD **3.37 higher** (0.11 lower to 6.85 higher)	MID = 5.95 (0.5 x median baseline SD)
Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months	21 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW a,c	-	The mean quality of life was 49.1	MD **12.1 higher** (7.12 lower to 31.32 higher)	MID = 2 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with other supervised exercise	
Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months	21 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 67	MD **9.4 higher** (0.97 lower to 19.77 higher)	MID = 3 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months	21 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 77.5	MD **2.5 lower** (25.05 lower to 20.05 higher)	MID = 3 (established value)
Quality of life (KOOS, 0-100, high is good, change score) at >3 months	84 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 3	MD **4 higher** (2 lower to 10 higher)	MID = 0.5 SD (SMD)
Quality of life (EQ-5D VAS, Quality of Well-being scale [different scale ranges], high is good, final values) at >3 months	529 (2 RCTs) follow up: mean 11 months	⊕⊕◯◯ LOW [a]	-	-	SMD **0.1 SD higher** (0.07 lower to 0.27 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at ≤3 months	272 (4 RCTs) follow up: mean 9 weeks	⊕⊕⊕◯ MODERATE [a,c]	-	-	SMD **0.79 SD lower** (1.04 lower to 0.54 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months	949 (17 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.5 SD lower** (0.63 lower to 0.36 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, 0-100, high is good, change scores) at >3 months	126 (2 RCTs)	⊕◯◯◯ VERY LOW [a,c]	-	The mean pain was 1.9	MD **3.76 higher** (0.12 lower to 7.64 higher)	MID = 5.7 (0.5 x median baseline SD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with other supervised exercise	
	follow up: mean 8 months					
Pain (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months	619 (3 RCTs) follow up: mean 12 months	⊕⊕◯◯ LOW [a]	-	-	SMD **0.12 SD lower** (0.28 lower to 0.04 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68 , high is poor, change scores) at ≤3 months	100 (3 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean physical function was 2.01	MD **9.26 lower** (13.77 lower to 4.74 lower)	MID = 7.0 (0.5 x median baseline SD)
Physical function (KOOS, WOMAC, Multidimensional Health Assessment Questionnaire [different scale ranges], high is poor, final values) at ≤3 months	879 (15 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.47 SD lower** (0.61 lower to 0.33 lower)	MID = 0.5 SD (SMD)
Physical function (KOOS, 0-100, high is good, change scores) at >3 months	84 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean physical function was 0	MD **4 higher** (0.13 higher to 7.87 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months	622 (3 RCTs) follow up: mean 12 months	⊕⊕◯◯ LOW [a]	-	-	SMD **0.22 SD lower** (0.38 lower to 0.06 lower)	MID = 0.5 SD (SMD)
Psychological distress (HADS anxiety subscale, DAS scale anxiety subscale [different scale	207 (2 RCTs)	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.33 SD lower** (0.63 lower to 0.03 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with other supervised exercise	
ranges], high is poor, final values) at ≤3 months	follow up: mean 10 weeks					
Psychological distress (HADS depression subscale, DAS scale depression subscale [different scale ranges], high is poor, final values) at ≤3 months	207 (2 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,c]	-	-	SMD **0.23 SD lower** (0.53 lower to 0.06 higher)	MID = 0.5 SD (SMD)
Psychological distress (DAS scale stress subscale, 0-48, high is poor, final value) at ≤3 months	152 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean psychological distress was 12.6	MD **5 lower** (8.68 lower to 1.32 lower)	MID = 0.5 SD (SMD)
Psychological distress (Centre for Epidemiological Studies Depression Scale, 0-60, high is poor, final value) at >3 months	214 (1 RCT) follow up: 20 weeks	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 8.092	MD **1.14 lower** (2.58 lower to 0.3 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	180 (4 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,d]	RD 0.07 (-0.10 to 0.25)	0 per 1,000	**70 more per 1,000** (100 fewer to 250 more) [e]	Precision calculated through Optimal Information Size (OIS) due to zero events in some studies (0.8-0.9 = serious, <0.8 = very serious).
Serious adverse events at >3 months	87 (1 RCT) follow up: 16 weeks	⊕◯◯◯ VERY LOW [a,c]	RR 2.05 (0.19 to 21.75)	23 per 1,000	**24 more per 1,000** (18 fewer to 472 more)	MID (precision) = RR 0.8-1.25.

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with other supervised exercise	

b. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis
c. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)
e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

1.1.6.6 Other unsupervised exercise compared to unsupervised strength exercise

Table 39: Clinical evidence summary: other unsupervised exercise compared to unsupervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with other unsupervised exercise	
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months	42 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 23	MD **17 lower** (28.24 lower to 5.76 lower)	MID = 0.5 SD (SMD)
Pain (KOOS, 0-100, high is good, change score) at ≤3 months	42 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was 11	MD **3 lower** (11.48 lower to 5.48 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months	42 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 13	MD **6 lower** (13.88 lower to 1.88 higher)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

1.1.6.7 Supervised mixed modality exercise compared to supervised strength exercise, unsupervised strength exercise, supervised aerobic exercise, other supervised exercise, unsupervised mixed modality exercise, pharmacological treatment and no treatment

Table 40: Clinical evidence summary: supervised mixed modality exercise compared to supervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with supervised mixed modality exercise	
Quality of life (AQoL, 0-1, high is good, final value) at ≤3 months	20 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 0.72	MD **0.01 lower** (0.16 lower to 0.14 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical function, 0-100, high is good, final values) at ≤3 months	143 (2 RCTs) follow up: mean 6 weeks	⊕⊕◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 31.9	MD **5.81 higher** (6.88 higher to 18.49 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, final values) at ≤3 months	143 (2 RCTs) follow up: mean 6 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 42.0	MD **8.15 higher** (9.2 lower to 25.5 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, final values) at ≤3 months	209 (3 RCTs) follow up: mean 7 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 32.6	MD **5.4 higher** (0.7 lower to 11.51 higher)	MID = 2 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months	83 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 3.35	MD **1.32 higher** (0.89 higher to 1.75 higher)	MID = 3 (established value)
Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months	83 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 17.42	MD **1.25 higher** (0.8 higher to 1.7 higher)	MID = 2 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months	83 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 7.575	MD **0.98 higher** (0.47 higher to 1.48 higher)	MID = 4 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months	83 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 7.35	MD **0.44 higher** (0.16 higher to 0.72 higher)	MID = 3 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months	83 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 31.15	MD **0.61 higher** (2.5 higher to 3.72 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months	66 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 43.82	MD **7.25 higher** (0.57 higher to 13.93 higher)	MID = 2 (established value)
Pain (WOMAC, VAS, NRS [different scale ranges], high is poor, final values) at ≤3 months	525 (10 RCTs) follow up: mean 8 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.67 SD lower** (1.09 lower to 0.24 lower)	MID = 0.5 SD (SMD)
Pain (WOMAC, VAS, NRS [different scale ranges], high is poor, final values) at >3 months	268 (3 RCTs) follow up: mean 39 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.3 SD lower** (0.65 lower to 0.05 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months	543 (9 RCTs) follow up: mean 8 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.83 lower** (1.3 lower to 0.36 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised strength exercise	Risk difference with supervised mixed modality exercise	
Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months	268 (3 RCTs) follow up: mean 39 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.5 SD lower** (1.08 lower to 0.08 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	193 (3 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,d]	RD 0.00 (-0.07 to 0.07)	0 per 1,000	**0 fewer per 1,000** (70 fewer to 70 more) [e]	Sample size used to determine precision: 75-150 = serious imprecision, <75 = very serious imprecision.
Serious adverse events at >3 months	113 (1 RCT) follow up: 6 months	⊕◯◯◯ VERY LOW [a,d]	RD 0.00 (-0.03 to 0.03)	0 per 1,000	**0 fewer per 1,000** (30 fewer to 30 more) [e]	Sample size used to determine precision: 75-150 = serious imprecision, <75 = very serious imprecision.

[a] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[b] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs
[c] Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis
[d] Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size
[e] Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 41: Clinical evidence summary: supervised mixed modality exercise compared to unsupervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 physical component, 0-100, high is good, final value) at ≤3 months	42 (1 RCT) Follow up: 8 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 36.9	MD **3.5 higher** (1.85 lower to 8.85 higher)	MID = 2 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 mental component, 0-100, high is good, final value) at ≤3 months	42 (1 RCT) Follow up: 8 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 48.4	MD **4.5 higher** (2.66 lower to 11.66 higher)	MID = 3 (established value)
Pain (BPI mean pain, 0-10, high is poor, final value) at ≤3 months	42 (1 RCT) Follow up: 8 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean pain was 4.55	MD **1.09 lower** (2.08 lower to 0.1 lower)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 42: Clinical evidence summary: supervised mixed modality exercise compared to supervised aerobic exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with supervised aerobic exercise	Risk difference with supervised mixed modality exercise	
Pain (WOMAC, 0-20, high is poor, change score) at >3 months	79 (1 RCT) follow up: 5 months	⨁⨁◯◯ LOW [a,b]	-	The mean pain was -1.5	MD **1 lower** (2.37 lower to 0.37 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months	79 (1 RCT) follow up: 5 months	⨁⨁◯◯ LOW [a,b]	-	The mean physical function was -6.6	MD **2.8 lower** (7.21 lower to 1.61 higher)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 43: Clinical evidence summary: supervised mixed modality exercise compared to other supervised exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with other supervised exercise	Risk difference with supervised mixed modality exercise	
Quality of life (KOOS, 0-100, high is good, final value) at ≤3 months	52 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 73	MD **1 higher** (5.26 lower to 7.26 higher)	MID = 0.5 SD (SMD)
Quality of life (EQ-5D, -0.11-1, high is good, final value) at ≤3 months	90 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 0.75	MD **0.03 lower** (0.08 lower to 0.02 higher)	MID = 0.03 (established MID)
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months	106 (2 RCTs) follow up: mean 8 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean quality of life was 40.2	MD **0.58 lower** (3.75 lower to 2.59 higher)	MID = 5.2 (0.5 x median baseline SD)
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months	106 (2 RCTs) follow up: mean 8 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 55	MD **1.63 lower** (4.98 lower to 1.72 higher)	MID = 5.0 (0.5 x median baseline SD)
Pain (KOOS, WOMAC, BPI, VAS [different scale ranges], high is poor, final values) at ≤3 months	334 (6 RCTs) follow up: mean 10 weeks	⊕⊕◯◯ LOW [a]	-	-	SMD **0.14 SD higher** (0.08 lower to 0.35 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months	149 (2 RCTs) follow up: mean 21 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.13 SD higher** (0.38 lower to 0.65 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, WOMAC [different scale	224 (4 RCTs)	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.03 SD lower** (0.58 lower to 0.64 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with other supervised exercise	Risk difference with supervised mixed modality exercise	
ranges], high is poor, final values) at ≤3 months	follow up: mean 10 weeks					
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months	85 (1 RCT) follow up: 26 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function as 18.8	MD **1.9 higher** (1.52 lower to 5.32 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS-anxiety, 0-21, high is poor, final value) at ≤3 months	60 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 3.8	MD **1.4 higher** (0.05 higher to 2.75 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS-depression, 0-21, high is poor, final value) at ≤3 months	60 (1 RCT) follow up: 8 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 3.8	MD **0.4 higher** (0.61 lower to 1.41 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	90 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	RR 0.60 (0.15 to 2.36)	111 per 1,000	**44 fewer per 1,000** (94 fewer to 151 more)	MID (precision) = RR 0.8-1.25.
Serious adverse events at >3 months	88 (1 RCT) follow up: 24 weeks	⊕⊕◯◯ LOW [a,d]	RD 0.00 (-0.04 to 0.04)	0 per 1,000	**0 fewer per 1,000** (40 fewer to 40 more) [e]	Sample size used to determine precision: 75-150 = serious imprecision, <75 = very serious imprecision.

[a.] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b.] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

[c.] Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

Table 44: Clinical evidence summary: supervised mixed modality exercise compared to unsupervised mixed modality exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised mixed modality exercise	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 physical function, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 60	MD **4.05 higher** (2.18 lower to 10.28 higher)	MID = 3 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 64.11	MD **9.99 higher** (2.2 higher to 17.78 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 61.84	MD **15.54 higher** (2.10 higher to 28.98 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 50	MD **1.67 higher** (8.34 lower to 11.68 higher)	MID = 2 (established value)
Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 57.89	MD **9.73 higher** (2.84 higher to 16.62 higher)	MID = 2 (established value)
Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 75.62	MD **0 higher** (7.5 lower to 7.5 higher)	MID = 3 (established value)
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 61.21	MD **25.98 higher** (11.58 higher to 40.38 higher)	MID = 4 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised mixed modality exercise	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 65.53	MD **59.58 lower** (67.03 lower to 52.13 lower)	MID = 3 (established value)
Pain (WOMAC, VAS [different scale ranges, high is poor, final values) at ≤3 months	140 (2 RCTs) follow up: mean 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.35 SD lower** (0.69 lower to 0.02 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	80 (1 RCT) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 18.89	MD **5.18 lower** (8.97 lower to 1.39 lower)	MID = 0.5 SD (SMD)

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

Table 45: Clinical evidence summary: supervised mixed modality exercise compared to pharmacological treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with pharmacological treatments	Risk difference with supervised mixed modality exercise	
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months	93 (1 RCT) follow up: 8 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 3.14	MD **1.36 lower** (6.58 lower to 3.86 higher)	MID = 0.5 SD (SMD)
Quality of life (KOOS, 0-100, high is good, final value) at >3 months	93 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	-	The mean quality of life was 8.7	MD **1.3 higher** (4.9 lower to 7.5 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with pharmacological treatments	Risk difference with supervised mixed modality exercise	
	follow up: 52 weeks					
Pain (KOOS, 0-100, high is good, change score) at ≤3 months	93 (1 RCT) follow up: 8 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean pain was 5.15	MD **2.08 higher** (2.28 lower to 6.44 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-500, high is poor, final value) at ≤3 months	104 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was 177.7	MD **23.1 lower** (60.11 lower to 13.91 higher)	MID = 0.5 SD (SMD)
Pain (KOOS, 0-100, high is good, change score) at >3 months	93 (1 RCT) follow up: 52 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean pain was 9.4	MD **4.2 higher** (1.45 lower to 9.85 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-500, high is poor, final value) at >3 months	104 (1 RCT) follow up: 26 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was 181.5	MD **19.9 lower** (56.08 lower to 16.28 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months	93 (1 RCT) follow up: 8 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean physical function was 7.46	MD **0.5 lower** (5.02 lower to 4.02 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-1800, high is poor, final value) at ≤3 months	104 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 685.7	MD **89.2 lower** (216.18 lower to 37.78 higher)	MID = 0.5 SD (SMD)
Physical function (KOOS, 0-100, high is good, change score) at >3 months	93 (1 RCT) follow up: 52 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean physical function was 7.9	MD **3.5 higher** (2.01 lower to 9.01 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with pharmacological treatments	Risk difference with supervised mixed modality exercise	
Physical function (WOMAC, 0-1800, high is poor, final value) at >3 months	104 (1 RCT) follow up: 6 months	⨁◯◯◯ VERY LOW [a,b]	-	The mean physical function was 691.4	MD **72.9 lower** (202.71 lower to 56.91 higher)	MID = 0.5 SD (SMD)
Serious adverse events at >3 months	110 (1 RCT) follow up: 26 weeks	⨁⨁◯◯ LOW [a,c]	RD 0.00 (-0.03 to 0.03)	0 per 1,000	**0 fewer per 1,000** (30 fewer to 30 more) [d]	Sample size used to determine precision: 75-150 = serious imprecision, <75 = very serious imprecision.

[a] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

[c] Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

[d] Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 46: Clinical evidence summary: supervised mixed modality exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at ≤3 months	72 (2 RCTs) follow up: mean 10 weeks	⨁◯◯◯ VERY LOW [a,b]	-	-	SMD **0.56 SD higher** (0.09 higher to 1.04 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months	139 (2 RCTs) follow up: mean 9 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 36.4	MD **1.66 higher** (1.57 lower to 4.89 higher)	MID = 5.1 (0.5 x median control group baseline SD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months	139 (2 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 50.7	MD **0.73 higher** (2.95 lower to 4.41 higher)	MID = 5.3 (0.5 x median control group baseline SD)
Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 19.2	MD **25.35 higher** (24.44 lower to 75.13 higher)	MID = 3 (established value)
Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at ≤3 months	163 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 10.1	MD **25.86 higher** (15.48 lower to 67.2 higher)	MID = 3 (established value)
Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 4.3	MD **41.88 higher** (42.4 lower to 126.15 higher)	MID = 3 (established value)
Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 12.2	MD **24.77 higher** (27.05 lower to 76.6 higher)	MID = 2 (established value)
Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 20	MD **19.57 higher** (14.21 lower to 53.36 higher)	MID = 2 (established value)
Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 24.2	MD **16.61 higher** (14.65 lower to 47.86 higher)	MID = 3 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 8.4	MD **34.83 higher** (37.46 lower to 107.12 higher)	MID = 4 (established value)
Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at ≤3 months	165 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 20.3	MD **27.94 higher** (29.14 lower to 85.03 higher)	MID = 3 (established value)
Quality of life (AIMS2 arm function, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 0.39	MD **0.13 lower** (0.44 lower to 0.18 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 arthritis pain, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 3.94	MD **0.85 lower** (1.52 lower to 0.18 lower)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 hand and finger function, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 0.62	MD **0.1 lower** (0.52 lower to 0.32 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 household tasks, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 0.35	MD **0.24 lower** (0.56 lower to 0.08 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 level of tension, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 3.45	MD **0.42 lower** (1.12 lower to 0.28 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Quality of life (AIMS2 mobility level, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 1.58	MD **0.5 lower** (0.93 lower to 0.07 lower)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 mood, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 1.7	MD **0.16 lower** (0.69 lower to 0.37 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 self-care tasks, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 0.06	MD **0.01 lower** (0.14 lower to 0.12 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 social activity, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 5.42	MD **0.08 lower** (0.63 lower to 0.47 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 support from family and friends, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 1.93	MD **0.08 lower** (0.82 lower to 0.66 higher)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 walking and bending, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 2.89	MD **1.25 lower** (2.08 lower to 0.42 lower)	MID = 0.5 SD (SMD)
Quality of life (AIMS2 work, 0-10, high is good, final value) at ≤3 months	124 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 1.28	MD **0.39 lower** (0.88 lower to 0.1 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Pain (WOMAC, VAS, 0-100, high is poor, change scores) at ≤3 months	331 (2 RCTs) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	MD **11.83 lower** (21.42 lower to 2.24 lower)	MID = 10.7 (0.5 x median baseline SD)
Pain (KOOS, WOMAC, AIMS, VAS, NRS [different scale ranges], high is poor, final values) at ≤3 months	476 (10 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.67 SD lower** (1.04 lower to 0.29 lower)	MID = 0.5 SD (SMD)
Pain (VAS, 0-100, high is poor, change scores) at >3 months	284 (2 RCTs) follow up: mean 44 weeks	⊕⊕◯◯ LOW [a,b]	-	-	MD **7.61 lower** (13.78 lower to 1.44 lower)	MID = 12 (0.5 x control group SD)
Pain (KOOS, NRS [different scale ranges], high is poor, final values) at >3 months	132 (2 RCTs) follow up: 42 weeks	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.63 SD lower** (0.98 lower to 0.27 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-100, high is poor, change score) at ≤3 months	139 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was -2.1	MD **6.3 lower** (10.67 lower to 1.93 lower)	MID = 0.5 SD (SMD)
Physical function (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months	392 (7 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.42 SD lower** (0.62 lower to 0.22 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months	107 (1 RCT) follow up: 32 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 31.4	MD **7.9 lower** (12.78 lower to 3.02 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months	139 (2 RCTs) follow up: 9 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 5.3	MD **0.71 higher** (0.43 lower to 1.85 higher)	MID = 1.8 (0.5 x median baseline SD)
Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months	139 (2 RCTs) follow up: 9 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean psychological distress was 4.6	MD **0.09 higher** (0.8 lower to 0.98 higher)	MID = 1.3 (0.5 x median baseline SD)
Psychological distress (AIMS psychological disability, 0-10, high is poor, final value) at ≤3 months	34 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 1.8	MD **0.08 higher** (0.56 lower to 0.72 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS, 0-21, high is poor, final value) at >3 months	106 (1 RCT) follow up: 32 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean psychological distress was 8	MD **1.6 higher** (0.91 lower to 4.11 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	284 (3 RCTs) follow up: mean 11 weeks	⊕◯◯◯ VERY LOW [a,d,e]	RD 0.01 (-0.02 to 0.04)	0 per 1,000	**10 fewer per 1,000** (40 fewer to 20 more) [f]	Precision calculated through Optimal Information Size (OIS) due to zero events in some studies (0.8-0.9 = serious, <0.8 = very serious).
Serious adverse events at >3 months	102 (1 RCT) follow up: 52 weeks	⊕⊕⊕◯ MODERATE [b]	Peto OR 0.14 (0.00 to 6.82)	20 per 1,000	**20 fewer per 1,000** (70 fewer to 30 more) [f]	MID (precision) = Peto OR 0.8-1.25.

[a.] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b.] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with supervised mixed modality exercise	

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis
d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)
e. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size
f. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

1.1.6.8 Unsupervised mixed modality exercise compared to unsupervised strength exercise, other unsupervised exercise, pharmacological treatment and no treatment

Table 47: Clinical evidence summary: unsupervised mixed modality exercise compared to unsupervised strength exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with unsupervised mixed modality exercise	
Pain (WOMAC, 0-20, high is poor, change score) at ≤3 months	32 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean pain was -1.88	MD **1.12 lower** (2.08 lower to 0.16 lower)	MID = 0.5 SD (SMD)
Pain (VAS, NRS, 0-10, high is poor, final values) at ≤3 months	189 (2 RCTs) follow up: 6 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean pain was 3.8	MD **0.05 lower** (1.17 lower to 1.06 higher)	MID = 0.9 (0.5 x median baseline SD)
Pain (NRS, 0-10, high is poor, final value) at >3 months	142 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean pain was 3.5	MD **0.1 higher** (0.86 lower to 1.06 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	191 (2 RCTs) follow up: mean 6 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean physical function was 7.84	MD **0.76 lower** (6.59 lower to 5.07 higher)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with unsupervised strength exercise	Risk difference with unsupervised mixed modality exercise	
a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis						

Table 48: Clinical evidence summary: unsupervised mixed modality exercise compared to other unsupervised exercise

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with other unsupervised exercise	Risk difference with unsupervised mixed modality exercise	
Pain (WOMAC, 0-100, high is poor, change score) at ≤3 months	179 (1 RCT) follow up: 8 weeks	⨁⨁◯◯ LOW [a]	-	The mean pain was 11	MD **7 higher** (4.64 higher to 9.36 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-100, high is poor, change score) at ≤3 months	179 (1 RCT) follow up: 8 weeks	⨁⨁◯◯ LOW [a]	-	The mean physical function was 5	MD **9 higher** (7.62 higher to 10.38 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	179 (1 RCT) follow up: 8 weeks	⨁◯◯◯ VERY LOW [a,b]	RR 0.60 (0.21 to 1.78)	91 per 1,000	**36 fewer per 1,000** (72 fewer to 71 more)	MID (precision) = RR 0.8-1.25.
a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs						

Table 49: Clinical evidence summary: unsupervised mixed modality exercise compared to pharmacological treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with pharmacological treatments	Risk difference with unsupervised mixed modality exercise	
Pain (HSS pain during activity, VAS [different scale ranges], high is poor, final values) at >3 months	135 (2 RCTs) follow up: mean 15 months	⊕◯◯◯ VERY LOW [a,b]	-	-	SMD **0.27 higher** (0.07 lower to 0.61 higher)	MID = 0.5 SD (SMD)
Pain (VAS, 0-100, high is poor, change score) at >3 months	120 (1 RCT) follow up: 24 weeks	⊕⊕◯◯ LOW [a]	-	The mean pain was -20.46	MD **0.83 lower** (12.32 lower to 10.66 higher)	MID = 0.5 SD (SMD)
Serious adverse events at >3 months	120 (1 RCT) follow up: 24 weeks	⊕⊕⊕◯ MODERATE [c]	RD 0.00 (-0.03 to 0.03)	0 per 1,000	**0 fewer per 1,000** (30 fewer to 30 more) [d]	Sample size used to determine precision: 75-150 = serious imprecision, <75 = very serious imprecision.

[a.] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

[b.] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

[c.] Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

[d.] Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 50: Clinical evidence summary: unsupervised mixed modality exercise compared to no treatment

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised mixed modality exercise	
Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at ≤3 months	203 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 0.777	MD **0** (0.04 lower to 0.05 higher)	MID = 0.03 (established value)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised mixed modality exercise	
Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at >3 months	203 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,b]	-	The mean quality of life was 0.784	MD **0** (0.05 lower to 0.05 higher)	MID = 0.03 (established value)
Pain (HOOS, 0-100, high is poor, final value) at ≤3 months	203 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean pain was 36.2	MD **4.4 lower** (9.44 lower to 0.64 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score) at >3 months	210 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean pain was -0.64	MD **0.51 lower** (1.43 lower to 0.41 higher)	MID = 0.5 SD (SMD)
Pain (HOOS, 0-100, high is poor, final value) at >3 months	203 (1 RCT) follow up: 12 months	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 34.6	MD **3 lower** (8.34 lower to 2.34 higher)	MID = 0.5 SD (SMD)
Physical function (HOOS, 0-100, high is poor, final value) at ≤3 months	203 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean physical function was 35.7	MD **6.9 lower** (12.45 lower to 1.35 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months	210 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean physical function was -1.51	MD **1.89 lower** (4.72 lower to 0.94 higher)	MID = 0.5 SD (SMD)
Physical function (HOOS, 0-100, high is poor, final value) at >3 months	203 (1 RCT) follow up: 12 months	⊕⊕◯◯ LOW [a,b]	-	The mean physical function was 34.2	MD **7.4 lower** (13.26 lower to 1.54 lower)	MID = 0.5 SD (SMD)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with no treatment	Risk difference with unsupervised mixed modality exercise	
Serious adverse events at >3 months	210 (1 RCT) follow up: 12 months	⨁◯◯◯ VERY LOW [a,b]	Peto OR 4.48 (0.54 to 36.96)	0 per 1,000	**30 more per 1,000** (10 fewer to 50 more) [c]	MID (precision) = Peto OR 0.8-1.25.

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

See Appendix F for full GRADE tables.

1.1.7 Economic evidence

1.1.7.1 Included studies

Four health economic studies in five papers were included in this review.[3, 234, 333, 357, 434]:

- exercise versus manual therapy, versus usual medical care.[3, 357];
- leaflet and advice versus joint protection versus hand exercises verses joint protection and hand exercises [333];
- supervised exercise versus usual GP care [434]; and
- individually tailored exercise versus targeted exercise therapy versus usual care [200].

These are summarised in the health economic evidence profiles below (**Table 52** to **Table 54**) and the health economic evidence tables in Appendix H.

Table 54

1.1.7.2 Excluded studies

One economic study relating to this review question was identified but was selectively excluded due to the availability of more applicable evidence.[216]. This is listed in Appendix J, with reasons for exclusion given.

See also the health economic study selection flow chart in Appendix G.

1.1.8 Summary of included economic evidence

Table 51: Health economic evidence profile: Abbott 2019 - Supervised exercise versus usual care.

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental QALYs	Cost effectiveness	Uncertainty
Abbott 2019[3] (Pinto 2013[357]) [New Zealand]	Partially applicable[(a)]	Potentially serious limitations[(b)]	• Within-RCT analysis (Abbott 2013[2] • Cost-utility analysis (QALYs) • Population: People with hip or knee osteoarthritis meeting American College of Rheumatology clinical diagnostic criteria for hip or knee OA. • Comparators: 1. Usual medical care 2. Supervised exercise plus usual care 3. Manual therapy plus usual care 4. Combination of exercise and manual therapy plus usual care • Time horizon: 2 years	2-1: saves £27 3-2: £1,052 4-3: saves £858 [(c)]	2-1: 0.15 3-2: -0.07 4-3: -0.01	Intervention 2 dominates all other interventions.	During sensitivity analysis, intervention 2 remained dominant over all other interventions when only complete case data were used and when participants who underwent joint replacement were excluded.

Abbreviations: ICER= incremental cost-effectiveness ratio; QALY= quality-adjusted life years; RCT= randomised controlled trial
(a) 2009 New Zealand resource use and unit costs may not reflect current UK NHS practice.
(b) Within trial analysis may not reflect full body of evidence available.
(c) 2009 New Zealand dollars converted to UK pounds.[334] Cost components incorporated: Medical and other healthcare consumed by participants during the trial.

Table 52: Health economic evidence profile: Oppong 2014- Exercise vs Leaflet and advice only

<table>
<tr><th>Study</th><th>Applicability</th><th>Limitations</th><th>Other comments</th><th>Incremental cost</th><th>Incremental effects</th><th>Cost effectiveness</th><th>Uncertainty</th></tr>
<tr><td>Oppong 2014[333] (UK)</td><td>Partially applicable [(a)]</td><td>Potentially serious limitations[(b)]</td><td>• Within-RCT analysis (Dziedzic 2015[114])
• Cost-utility analysis (QALYs)
• Population: Adults aged 50 years or older with hand osteoarthritis
• Comparators:
1. Leaflet and advice only
2. Joint protection only
3. Hand exercises only
4. Joint protection and hand exercises
Follow-up: 1 year</td><td colspan="3">Full incremental analysis [(c)]:
<table><tr><th>Int</th><th>Cost [(d)]</th><th>QALY</th><th>Inc cost</th><th>Inc QALY</th><th>ICER</th></tr><tr><td>4</td><td>£112</td><td>0.658</td><td colspan="3">Dominated</td></tr><tr><td>2</td><td>£92</td><td>0.659</td><td colspan="3">Dominated</td></tr><tr><td>1</td><td>£58</td><td>0.662</td><td colspan="3">Baseline</td></tr><tr><td>3</td><td>£65</td><td>0.681</td><td>£6</td><td>0.019</td><td>£318</td></tr></table></td><td>Probability Intervention 3 cost effective versus Intervention 1 (£20K threshold): 80%

Study explored different analytic methods to generate the cost effectiveness results. Conclusions unchanged by use of different analytic methods.</td></tr>
</table>

Abbreviations: ICER= incremental cost-effectiveness ratio; QALY= quality-adjusted life years; RCT= randomised controlled trial

(a) Study does not include all exercise treatment options.

(b) Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison – This trial was not included in the clinical review because it did not contain relevant outcomes.

(c) Intervention number in order of least to most effective in terms of QALYs. Costs rounded up.

(d) 2010/2011 costs. Cost components incorporated: Intervention, primary care (general practice and nurse); secondary care (orthopaedic surgeon, rheumatologist, plastic surgeon, physiotherapist, occupational therapist), other health care staff and prescribed medication. As all participants received the leaflet and advice, this cost was not included in the analysis.

Table 53: Health economic evidence profile: Supervised exercise versus usual care

Study	Applicability	Limitations	Other comments	Incremental cost (2 vs 1)	Incremental effects (2 vs 1)	Cost effectiveness (2 vs 1)	Uncertainty
Tan 2016[434] (The Netherlands)	Partially applicable [(a)]	Potentially serious limitations[(b)]	• Within-RCT analysis (Teirlinck 2016 [439]) • Cost-utility analysis (QALYs) • Population: Adults with hip osteoarthritis in primary care >45yrs • Comparators: 1. GP care (usual care) 2. Supervised exercise plus GP care Follow up: 1 year	Saves £83[(c)]	-0.006 QALYs	Supervised exercise saves extra £13,793 per QALY gained	None undertaken from healthcare perspective.

Abbreviations: ICER= incremental cost-effectiveness ratio; QALY= quality-adjusted life years; RCT= randomised controlled trial

(a) Dutch healthcare perspective may not reflect current UK NHS context. Study does not include all exercise treatment options

(b) Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison. No analysis of uncertainty.

(c) 2011 Euros converted to UK pounds.[334] Cost components incorporated: Healthcare professional visits in primary and secondary care, medical investigations/interventions and prescribed medications. Interventions - number and grade of staff involved and equipment use to deliver intervention as well as number of sessions attended.

Table 54: Health economic evidence profile: Individually tailored exercise versus targeted exercise therapy versus no treatment

<table>
<tr><th>Study</th><th>Applicability</th><th>Limitations</th><th>Other comments</th><th>Incremental cost</th><th>Incremental effects</th><th>Cost effectiveness</th><th>Uncertainty</th></tr>
<tr>
<td>Kigozi 2018[200] (UK)</td>
<td>Partially applicable [(a)]</td>
<td>Potentially serious limitations[(b)]</td>
<td>• Within-RCT analysis (Kigozi 2018[200])
• Cost-utility analysis (QALYs)
• Population: Adults with knee osteoarthritis in primary care
• Comparators:
1.No treatment
2.Individually tailored exercise
3.Targeted exercise therapy
Follow-up: 18 months</td>
<td colspan="3">Full incremental analysis [(c)]:
<table>
<tr><th>Int</th><th>Cost</th><th>QALY</th><th>Inc cost</th><th>Inc QALY</th><th>ICER</th></tr>
<tr><td>3</td><td>£524</td><td>1.019</td><td colspan="3">Dominated</td></tr>
<tr><td>2</td><td>£656</td><td>1.032</td><td colspan="3">Dominated</td></tr>
<tr><td>1</td><td>£383</td><td>1.035</td><td colspan="3">Baseline</td></tr>
</table></td>
<td>Probability Intervention 2 or 3 being cost effective compared to 1 (£20K threshold): <40%

Complete case analysis to assess impact of missing cost and EQ5D data. This resulted in the same conclusion that usual care was dominant.</td>
</tr>
</table>

Abbreviations: ICER= incremental cost-effectiveness ratio; QALY= quality-adjusted life years; RCT= randomised controlled trial

(a) Study does not include all exercise treatment options.

(b) Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison.

(c) Intervention number in order of least to most effective in terms of QALYs.

(d) 2012/2013 costs. Cost components incorporated: Primary care consultations (GP, nurse practitioners, community physical therapists), consultations with other health-care professionals (hospital consultants, hospital physical therapists, acupuncturists), hospital-based investigations (X-ray and MRI), procedures (injections, surgery), prescribed meds. Intervention costs - sessions. Also included a 47-minute initial assessment and treatment session, followed by 28-minute face to face treatment session, 11 min telephone call contacts (where applicable).

1.1.9 Economic model

This area was not prioritised for new cost-effectiveness analysis.

1.1.10 Unit costs

Relevant unit costs are provided below to aid consideration of cost effectiveness.

Resource	Unit costs (cost per hour)	Source
Community physiotherapist (band 5/6/7)	£38/£50/£60	PSSRU 2020[96]

1.1.11 Economic evidence statements

- One cost utility analysis reported that supervised exercise dominated usual medical care, manual therapy plus usual care and a combination of exercise and manual therapy plus usual care. This analysis was assessed as partially applicable with potentially serious limitations.
- One cost utility analysis reported that hand exercises alone was cost effective versus leaflets and advice only (ICER: £318). Hand exercises alone also dominated joint protection alone and joint protection and hand exercises combined. This analysis was assessed as partially applicable with potentially serious limitations.
- One cost utility analysis reported that GP care alone cost an extra £13,793 per QALY gained versus exercise plus GP care. This analysis was graded as partially applicable with potentially serious limitations.
- One cost utility analysis reported that no treatment dominated individually tailored exercise as well as targeted exercise therapy. This analysis was graded as partially applicable with potentially serious limitations.

1.1.12 The committee's discussion and interpretation of the evidence

1.1.12.1. The outcomes that matter most

The critical outcomes were quality of life, pain and physical function. These were considered critical due to their importance to people with osteoarthritis. The Osteoarthritis Research Society International (OARSI) consider that pain and physical function were the most important outcomes for evaluating interventions. Quality of life gives a broader perspective on the person's wellbeing, allowing for examination of the biopsychosocial impact of interventions. Psychological distress, osteoarthritis flare and serious adverse events were included as important outcomes.

The committee considered osteoarthritis flares to be important in the lived experience and management of osteoarthritis. However, these were also considered difficult to measure with no clear consensus on their definition. The Flares in OA OMERACT working group have proposed an initial definition and domains of OA flares through a consensus exercise; "it is a transient state, different from the usual state of the condition, with a duration of a few days, characterized by onset, worsening of pain, swelling, stiffness, impact on sleep, activity, functioning, and psychological aspects that can resolve spontaneously or lead to a need to adjust therapy.". However, this has been considered to have limitations and has not been widely adopted. Therefore, the committee included the outcome accepting any reasonable definition provided by any studies discussing the event.

Mortality was included as treatment adverse events rather than as a discreet outcome and categorised as an important outcome. Osteoarthritis as a disease process is not considered to cause mortality by itself and mortality is an uncommon outcome from osteoarthritis interventions. There was evidence available for all outcomes apart from osteoarthritis flares. However, there was only limited evidence available for psychological distress and serious adverse events throughout the literature.

1.1.12.2 The quality of the evidence

One-hundred and four studies were included in this review. The comparisons where evidence was present included:

- Supervised strength exercise compared to unsupervised strength exercise
- Supervised strength exercise compared to supervised aerobic exercise
- Supervised strength exercise compared to no treatment
- Unsupervised strength exercise compared to unsupervised aerobic exercise
- Unsupervised strength exercise compared to no treatment
- Supervised aerobic exercise compared to no treatment
- Unsupervised aerobic exercise compared to no treatment
- Other supervised exercise (for example: aquatic exercise, mind-body (including yoga, tai chi, qigong and baduanjin), neuromuscular (including exercises focussing on balance, proprioception and other specific interventions) and strength exercise while having whole body vibration) compared to supervised strength exercise
- Other supervised exercise compared to unsupervised strength exercise
- Other supervised exercise compared to no treatment
- Other unsupervised exercise compared to unsupervised strength exercise
- Supervised mixed modality exercise compared to supervised strength exercise
- Supervised mixed modality exercise compared to unsupervised strength exercise
- Supervised mixed modality exercise compared to supervised aerobic exercise
- Supervised mixed modality exercise compared to other supervised exercise
- Supervised mixed modality exercise compared to unsupervised mixed modality exercise
- Supervised mixed modality exercise compared to pharmacological treatment
- Supervised mixed modality exercise compared to no treatment
- Unsupervised mixed modality exercise compared to unsupervised strength exercise
- Unsupervised mixed modality exercise compared to other unsupervised exercise
- Unsupervised mixed modality exercise compared to pharmacological treatment
- Unsupervised mixed modality exercise compared to no treatment

The evidence varied from moderate to very low quality, with the majority of evidence being of very low quality. Outcomes were commonly downgraded for risk of bias, in particular for risk of performance bias, with studies not being blinded and so subjective outcomes were commonly downgraded accordingly. In more than thirty studies, different groups reported differences in baseline values of the outcomes. This made interpretation of the results more challenging for the committee. These factors, in addition to studies being downgraded for inconsistency and imprecision, led to the very low quality rating. When present, inconsistent results were not explained by subgroup analysis. The majority of comparisons consisted of studies with a small number of participants (less than 50) with a few studies that included a larger number of participants.

The committee agreed that there was sufficient evidence to compare different types of exercise to each other and to no treatment. There was limited evidence comparing to pharmacological treatments, making it difficult to draw a conclusion on the comparison of exercise to pharmacological treatments.

Supervised strength exercise

Evidence was available comparing supervised strength exercise to unsupervised strength exercise, supervised aerobic exercises, other supervised exercise, supervised mixed modality exercise, pharmacological treatment and no treatment.

- When compared to unsupervised strength exercise, evidence was of moderate quality. When downgrading occurred, this was due to risk of bias (in particular selection bias).
- When compared to supervised aerobic exercise, evidence was of very low quality. When downgrading occurred, this was due to risk of bias and imprecision.
- When compared to other supervised exercise, evidence was mainly of very low quality, but ranged from moderate to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to supervised mixed modality exercise, evidence was mainly of very low quality but ranged from low to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to pharmacological treatment, evidence was of low quality due to risk of bias.
- When compared to no treatment, evidence was mainly of very low quality, but ranged from low to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

Unsupervised strength exercise

Evidence was available comparing unsupervised strength exercise to supervised strength exercise, unsupervised aerobic exercise, other supervised exercise, other unsupervised exercise, supervised mixed modality exercise, unsupervised mixed modality exercise and no treatment.

- When compared to supervised strength exercise, evidence was of moderate quality. When downgrading occurred, this was due to risk of bias (in particular selection bias).
- When compared to unsupervised aerobic exercise, evidence was mostly of very low quality but ranged from low to very low quality. When downgrading occurred, this was due to a mixture of risk of bias and imprecision.
- When compared to other supervised exercise, evidence was of low to very low quality. When downgrading occurred, this was due to a mixture of risk of bias and imprecision.
- When compared to other unsupervised exercise, evidence was of very low quality due to risk of bias and imprecision.
- When compared to supervised mixed modality exercise, evidence was of low quality due to risk of bias and imprecision.
- When compared to unsupervised mixed modality exercise, evidence was of low to very low quality due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to no treatment, evidence was mostly of low quality, but ranged from high to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

Supervised aerobic exercise

Evidence was available comparing supervised aerobic exercise to supervised strength exercise, other supervised exercise and no treatment.

- When compared to supervised strength exercise, evidence was of very low quality. When downgrading occurred, this was due to risk of bias and imprecision.
- When compared to other supervised exercise, evidence was of very low quality due to risk of bias and imprecision.

- When compared to no treatment, evidence ranged from low to very low quality. When downgrading occurred, this was due to a mixture of risk of bias and imprecision.

Unsupervised aerobic exercise

Evidence was available comparing unsupervised aerobic exercise to unsupervised strength exercise and no treatment.

- When compared to unsupervised strength exercise, evidence was mostly of very low quality but ranged from low to very low quality. When downgrading occurred, this was due to a mixture of risk of bias and imprecision.
- When compared to no treatment, evidence was mostly of very low quality but ranged from moderate to very low quality. When downgrading occurred, this was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

Other supervised exercise

Evidence was available comparing other supervised exercise to supervised strength exercise, unsupervised strength exercise, supervised aerobic exercise, supervised mixed modality exercise and no treatment.

- When compared to supervised strength exercise, evidence was mainly of very low quality, but ranged from moderate to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to unsupervised strength exercise, evidence was of low to very low quality. When downgrading occurred, this was due to a mixture of risk of bias and imprecision.
- When compared to supervised aerobic exercise, evidence was of very low quality due to risk of bias and imprecision.
- When compared to supervised mixed modality exercise, evidence was mostly of very low quality but ranged from moderate to very low quality. When downgrading occurred, this was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to no treatment, evidence was mostly of very low quality, but ranged from moderate to very low quality. When downgrading occurred, this was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

Other unsupervised exercise

Evidence was available comparing other unsupervised exercise to unsupervised strength exercise and unsupervised mixed modality exercise.

- When compared to unsupervised strength exercise, evidence was of very low quality due to risk of bias and imprecision.
- When compared to unsupervised mixed modality exercise, evidence was between low and very low quality. This was due to a mixture of risk of bias and imprecision.

Supervised mixed modality exercise

Evidence was available comparing supervised mixed modality exercise to supervised strength exercise, unsupervised strength exercise, other supervised exercise, unsupervised mixed modality exercise, pharmacological treatment and no treatment.

- When compared to supervised strength exercise, evidence was mainly of very low quality but ranged from low to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

- When compared to unsupervised strength exercise, evidence was of low quality due to risk of bias and imprecision.
- When compared to other supervised exercise, evidence was mostly of very low quality but ranged from moderate to very low quality. When downgrading occurred, this was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to unsupervised mixed modality exercise, evidence was mostly of very low quality but ranged from low to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to pharmacological treatment, evidence was mostly of low quality but ranged between moderate and very low quality. This was due to a mixture of risk of bias and imprecision.
- When compared to no treatment, evidence was mostly of very low quality but ranged from moderate to very low quality. This was due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.

Unsupervised mixed modality exercise

Evidence was available comparing unsupervised mixed modality exercise to unsupervised strength exercise, other unsupervised exercise, supervised mixed modality exercise and no treatment.

- When compared to unsupervised strength exercise, evidence was of low to very low quality due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to other unsupervised exercise, evidence was between low and very low quality. This was due to a mixture of risk of bias and imprecision.
- When compared to supervised mixed modality exercise, evidence was of low to very low quality due to a mixture of risk of bias, imprecision and inconsistency due to heterogeneity unresolved by subgroup analysis.
- When compared to no treatment, evidence was mostly due to low quality but ranged from moderate to very low quality. This was due to a mixture of risk of bias and imprecision.

1.1.12.3 Benefits and harms

Key uncertainties

The committee agreed to separate out the types of exercise by whether they were supervised or unsupervised and by the mechanism of exercise, leading to four categories of strength, aerobic, other and mixed modality exercise. The committee noted that there were other classifications that could be used to interpret the evidence, including intensity level. As including this would have made the evidence too sparse, the committee agreed to only look at the former factors. However, this could influence the efficacy of exercise techniques which the committee were not able to analyse from these results.

The committee discussed that generally the adverse events data for these trials was limited as this was generally found in small studies with a short follow up time and so it is unclear whether this is representative of the events expected to be seen in real life practice. Given this, the committee considered the evidence for serious adverse events to be unclear throughout the review reflecting this in their weighting of findings while making recommendations. The committee noted throughout the evidence that the number of adverse events was often low and where events were reported they were transient in nature (such as increased pain). Given this, while the committee acknowledged where clinically important differences were highlighted in the evidence, but also considered the nature and true number of these events.

The comparison to no treatment means that is a potential for performance bias. Unlike other interventions covered in this guideline, the committee agreed that there was no appropriate method for a sham/placebo comparison for exercise. This was due to any form of joint movement used in a sham exercise having the potential to replicate the same mechanism of treatment as the exercise itself. Given this, it was agreed that no treatment was the best comparison to use. Furthermore, the diversity in the interventions which were classified in as no treatment in this review led to additional challenges in interpretation (as no treatment could vary from absolutely no intervention to several different modalities of treatment which were available to all intervention arms). The committee agreed that this definition would be comparable (with there being no singular agreed standard of care for people with osteoarthritis) but were aware that this could introduce an element of uncertainty while making recommendations.

The comparison to pharmacological treatments was difficult for the committee to interpret. This was due to the limited number of pharmacological interventions compared to exercise in the studies included in the review. The studies included compared exercise to intra-articular hyaluronic acid, non-steroidal anti-inflammatory drugs and glucosamine. As this is not a true representation of all of the types of pharmacological treatments used for osteoarthritis, the committee could not comment on the differences in the two interventions.

Studies varied in the time that outcomes were reported in relation to the length of time when an intervention was given (for example: some studies reported immediately post-intervention, while others may include treatment for 12 weeks and follow up for an additional 40 weeks after this). This made it difficult to interpret treatment effects, in particular those reporting outcomes at more than 3 months. Given this, while the outcomes at more than 3 months show no clinically important difference for critical outcomes in most comparisons, the committee were not able to draw clear conclusions due to the lack of consistency in intervention duration. Due to this, the committee relied more on the effect at less than 3 months, where the follow up times were more likely to be similar to the intervention duration. The committee used their expert opinion to conclude that the effects of exercise were not likely to be reduced over time, and so recommended that people need to continue exercise long-term in order to maintain the effect (see recommendation 1.3.3).

There was very limited information for people with osteoarthritis of joints other than knee. Where heterogeneity was present and outcomes contained studies including people with sites of osteoarthritis other than the knee, the subgrouping of studies by the joint affected did not resolve the heterogeneity. The committee's expert opinion was that the effects of exercise were likely to be beneficial for people with osteoarthritis in other joints.

Supervised strength exercise

Studies comparing supervised strength exercise to unsupervised strength exercise, supervised aerobic exercise, other supervised exercise, supervised mixed modality exercise and no treatment were included in the analysis. This evidence came from a population with osteoarthritis of the knee or hip. The type of strength exercise varied from singular exercises (for example: quadriceps strengthening exercises) to a program of exercises.

The results showed that supervised strength exercise (when compared to no treatment) led to a clinically important benefit in pain and physical function at less than or equal to 3 months. There was an unclear effect seen for quality of life where 8 outcomes showed a clinically important benefit, while 4 showed no clinically important difference. There was no clinically important difference in psychological distress and a clinically important harm in the protocol outcome of serious adverse events. The adverse events recorded were increases in pain and inflammation. The committee agreed that the adverse events were likely to be mild and transient. They also noted, that while some people reported these adverse events, given that the pain score reduced then these adverse events could be outweighed against the potential benefits. Finally, the committee discussed that generally the adverse events data for these trials was limited as this was generally found in small studies with a short follow up

time and so it is unclear whether this is representative of the events expected to be seen in real life practice. Given this, the committee considered the evidence for serious adverse events to be unclear throughout the review reflecting this in their weighting of findings while making recommendations.

The clinically important benefit in pain was retained at more than 3 months. However, the effect on physical function was not, with no clinically important difference being seen at this time period. At more than 3 months there was no clinically important difference in quality of life.

In general, supervised strength exercise did not appear to have any clinically important difference to the other forms of exercise mentioned above. However, the supervised strength exercise when compared to other supervised exercise comparison showed a clinically important benefit in serious adverse events based on 1 small study (N=90). The events seen in this study were increased pain and inflammation, which the committee agreed was consistent with the findings when compared to no treatment and were likely to be transient and outweighed by the potential long-term benefits. An unclear potential clinically important benefit of other supervised exercise and supervised mixed modality exercises was seen for quality of life with supervised strength exercise having a smaller effect.

Unsupervised strength exercise

Studies compared unsupervised strength exercise to supervised strength exercise, unsupervised aerobic exercise, other supervised exercise, supervised mixed modality exercise, unsupervised mixed modality exercise and no treatment. This evidence came from a population with osteoarthritis of the knee, hand or hip. The type of strength exercise varied in the number of exercises and the use of additional equipment (for example: resistance weights).

The results showed that unsupervised strength exercise (when compared to no treatment) led to a clinically important benefit in physical function at less than 3 months. There was an unclear effect seen for quality of life, where 5 outcomes showed a clinically important benefit, while 2 showed no clinically important difference, and pain where 1 outcome showed a clinically important benefit and 1 outcome showed no clinically important difference. There was a clinically important harm in the protocol outcome of serious adverse events seen in 1 outcome including 1 study at less than 3 months.

At more than 3 months, the clinically important benefit in physical function was not retained, instead showing no clinically important difference. The effect on quality of life remained unclear, with 5 outcomes showing a clinically important benefit and 4 showing no clinically important difference. The effect on pain and psychological distress showed no clinically important difference. However, the clinically important harm in serious adverse events was retained at more than 3 months. The adverse events included pain, inflammation and events the committee agreed were likely unrelated (varicose veins). The committee discussed that generally the adverse events data for these trials was limited as this was generally found in small studies with a short follow up time and so it is unclear whether this is representative of the events expected to be seen in real life practice. Given this, the committee considered the evidence for serious adverse events to be unclear throughout the review reflecting this in their weighting of findings while making recommendations.

In general, supervised strength exercise showed greater clinically important benefits when compared to unsupervised exercise. Otherwise, comparison to other supervised exercise interventions showed mixed results. This included other supervised exercise, where there was unclear results for quality of life and pain, including outcomes that showed no clinically important difference and outcomes that showed a clinically important harm, and a clinically important harm in physical function. This also included supervised mixed modality exercise where there were clinically important harms seen with unsupervised strength exercise in

quality of life and pain. The committee noted that these were seen in low-very low quality outcomes in 1 small study for each.

However, when compared to other unsupervised forms of exercise there was generally no clinically important difference seen in pain and physical function. There were potential harms seen in quality of life, where when compared to unsupervised aerobic exercise 2 outcomes showed no clinically important difference and 4 outcomes showed no clinically important difference, and when compared to other unsupervised exercise there was a clinically important harm in 1 outcome (both with results from 1 study). Therefore, the committee did not conclude that any type of unsupervised exercise was superior to any other.

Supervised aerobic exercise

Studies compared supervised aerobic exercise to supervised strength exercise, supervised mixed modality exercise and no treatment. The studies included people with osteoarthritis of the knee and hip. The type of aerobic exercise varied from walking programmes (including Nordic walking) to treadmill training.

When compared to no treatment, supervised aerobic exercise showed a clinically important benefit in pain and physical function at less than 3 months. There was no clinically important difference in quality of life and a clinically important harm in the protocol outcome of serious adverse events. However, the evidence came from a limited number of studies (at most 2) and included a small number of participants (at most 55). The adverse events seen were knee and wrist pain, which the committee agreed would be transient and were outweighed by the otherwise clinically important benefits observed for pain overall. These effects were not retained long term, with quality of life, pain and physical function being found to have no clinically important difference. When compared to other forms of exercise there was no clinically important difference seen in pain and physical function (at less than and more than 3 months when compared to supervised strength exercise, and at more than 3 months only when compared to supervised mixed modality exercise).

The committee discussed that generally the adverse events data for these trials was limited as this was generally found in small studies with a short follow up time and so it is unclear whether this is representative of the events expected to be seen in real life practice. Given this, the committee considered the evidence for serious adverse events to be unclear throughout the review reflecting this in their weighting of findings while making recommendations.

Unsupervised aerobic exercise

Studies compared unsupervised aerobic exercise to unsupervised strength exercise and no treatment. The studies included people with osteoarthritis of the knee or hip. The type of aerobic exercise including walking programs, a website supported activity-based program, and treadmill training.

When compared to no treatment, unsupervised aerobic exercise showed a clinically important benefit in pain and physical function at less than 3 months. There was an unclear effect on quality of life, with 5 outcomes showing a clinically important benefit and 2 showing no clinically important difference. No clinically important difference was seen in psychological distress. Any benefit was not retained at more than 3 months with quality of life, pain, physical function and psychological distress showing no clinically important difference. When compared to unsupervised strength exercise, there was an unclear effect on quality of life with 4 outcomes showing a clinically important benefit and 2 outcomes showing no clinically important difference. In pain and physical function, there was no clinically important difference seen at less than 3 months.

Other supervised exercise

Studies compared other supervised exercise to supervised strength exercise, unsupervised strength exercise, supervised mixed modality exercise, unsupervised mixed modality exercise and no treatment. The studies included people with osteoarthritis of the knee and hip. The type of exercises included: aquatic exercises, mind-body (including yoga, tai chi, qigong and baduanjin), neuromuscular (including exercises focussing on balance, proprioception [perception and awareness of position and movement in the body] and other specific interventions) and strength exercise while having whole body vibration.

When compared to no treatment, other supervised exercise showed a clinically important benefit in pain at less than 3 months. There was an unclear effect on quality of life, physical function and psychological distress with 1-2 outcomes showing a clinically important benefit and 1-4 outcomes showing no clinically important difference. There was no clinically important difference in serious adverse events (based on 1 outcome including 3 studies with 78 participants in total). No positive effects were retained long term with no clinically important difference in quality of life, pain, physical function, psychological distress and serious adverse events at more than 3 months.

When compared to other intervention, mostly there was no clinically important difference. There was no difference seen when compared to supervised exercises apart from quality of life (where there was an unclear benefit) and serious adverse events (where there was a clinically important benefit of supervised strength exercise rather than other supervised exercise) when compared to supervised strength exercise, and an unclear possible greater benefit in psychological distress for people receiving supervised mixed modality exercise when compared to other supervised exercise. The unclear benefit in quality of life with supervised strength exercise was retained at more than 3 months while the clinically important harm in adverse events was not, showing no clinically important difference at more than 3 months. There were benefits and unclear benefits seen when compared to unsupervised strength exercise. Unsupervised mixed modality exercise showed a clinically important benefit in pain and physical function when compared to other supervised exercise. However, there was no clinically important difference in adverse events.

Other unsupervised exercise

Other unsupervised exercise was compared to unsupervised strength exercise and supervised mixed modality exercise. The studies included people with osteoarthritis of the knee only. The types of exercise included stretching and neuromuscular exercises.

When compared to unsupervised strength exercise there was a clinically important difference in quality of life at less than 3 months (based on 1 outcome including 1 study with 42 participants). There was no clinically important difference in pain and physical function. When compared to supervised mixed modality exercise there was no difference in pain at more than 3 months.

Supervised mixed modality exercise

Supervised mixed modality exercise was compared to supervised strength exercise, unsupervised strength exercise, supervised aerobic exercise other supervised exercise, other unsupervised exercise, unsupervised mixed modality exercise, pharmacological treatment and no treatment. The studies included people with osteoarthritis of the knee and hip. The types of exercises making up combinations included: strength and aerobic; strength, aerobic and neuromuscular; strength, flexibility and aerobic; neuromuscular and strength; neuromuscular and aerobic and strength and range of motion.

When compared to no treatment, there was a clinically important benefit in pain at less than 3 months. There was an unclear different in quality of life, with 10 outcomes showing a clinically important benefit and 13 outcomes showing no clinically important difference. There

was no clinically important difference in physical function, psychological distress and adverse events. At more than 3 months, there was a clinically important benefit in physical function. There was an unclear effect on pain, with a clinically important benefit in 1 outcome and no clinically important difference in 1 outcome. There was no clinically important difference in psychological distress and serious adverse events.

When compared to other forms of exercise there was no clinically important difference when compared to supervised exercises, with the exception of unclear potential benefits in quality of life when compared to supervised strength exercise and psychological distress when compared to other supervised exercise. When compared to unsupervised exercise, supervised mixed modality exercise showed a clinically important benefit in quality of life and pain at less than 3 months. There was a clinically important benefit in physical function when compared to unsupervised mixed modality exercise. There was an unclear effect on quality of life for this comparison, where 5 outcomes showed a clinically important benefit, 2 outcomes showed no clinically important difference and 1 outcome showed a clinically important harm. There was no clinically important difference in pain at less than 3 months.

When compared to other forms of exercise at more than 3 months, mostly there was no clinically important difference with the exception of quality of life when compared to supervised strength exercise where there was a clinically important benefit. Finally, when compared to pharmacological treatment there was no clinically important difference at less than and more than 3 months in quality of life, pain, physical function and serious adverse events.

Unsupervised mixed modality exercise

Unsupervised mixed modality exercise was compared to unsupervised strength exercise, other supervised exercise, supervised mixed modality exercise, pharmacological treatment and no treatment. The studies included people with osteoarthritis of the knee, hip and ankle. The types of exercise included: strength and range of motion; strength and neuromuscular; strength, aerobic and stretching; strength, neuromuscular, flexibility and range of motion.

When compared to no treatment, there was no clinically important difference seen in quality of life, pain and physical function at less than 3 months, and quality of life, pain, physical function and adverse events at more than 3 months. When compared to other interventions, mostly there was no clinically important difference seen. The exceptions where other supervised exercise where there was a clinically important benefit in pain and physical function (based on 1 study with 179 participants) and supervised mixed modality exercise where there was an unclear effect on quality of life for this comparison, where 5 outcomes showed a clinically important benefit, 2 outcomes showed no clinically important difference and 1 outcome showed a clinically important harm, and physical function where there was a clinically important harm at less than 3 months.

Weighing up the clinical benefits and harms

Given this information, the committee acknowledged the benefit from exercise compared to no treatment. They concluded that the benefits in terms of pain, physical function and quality of life outweighed any possible harms. They noted that there did not appear to be a difference between different types of exercise. Therefore, the committee to recommend therapeutic exercise that could include strength and aerobic exercises (see recommendation 1.3.1) but no specific type of program. The committee agreed that therapeutic exercise, where exercise specifically aims at preventing progression and managing symptoms, was important. Therefore, exercise provision should be tailored to the needs of the person, with joint site-specific exercises to achieve this.

There was some evidence showing that supervised exercise was superior to unsupervised exercise. However, this was limited to small studies and was of low quality. However, expert consensus among the committee recommended healthcare professionals advise that

supervised exercise is likely to be of greater benefit than unsupervised exercise to people with osteoarthritis. The reasons behind this included that supervised exercise may enable tailored exercise, social support, and may lead to greater therapeutic rapport and exercise habit formation. Therefore, the committee made recommendation 1.3.2 recommending that supervised exercise could be considered, recognising the potential additional benefits seen from supervised exercise in some comparisons. The committee provided additional information to advise people starting therapeutic exercise in recommendation 1.3.3.

1.1.12.4 Cost effectiveness and resource use

Overall, the clinical review indicated that exercise interventions can improve pain, function and quality of life. However, exercise interventions vary greatly in their intensity and resource use. This can range from advice to exercise up to supervised one-to-one exercise. Although more costly, supervised exercise could be cost effective if the additional quality of life improvement is big. Provision of supervised exercise programmes is variable across the NHS.

Four economic evaluations were included. They compared the following:

- exercise versus manual therapy, versus usual medical care for hip and knee osteoarthritis[3]
- leaflet and advice versus joint protection versus hand exercises verses joint protection and hand exercises for hand osteoarthritis[333]
- supervised exercise versus usual GP care for hip osteoarthritis[434]
- supervised individually tailored exercise versus a mixture of supervised and unsupervised targeted exercise therapy versus usual care in knee osteoarthritis[200]

The committee concluded that the benefits of exercise in general were very clear. However, the cost effectiveness of supervised exercise was uncertain and would depend on the specifics of the programme and the patient selection.

The committee recommended that patients with osteoarthritis be directed towards therapeutic exercise. Given the absence of economic evidence comparing supervised to unsupervised exercise and the low quality of the clinical evidence for this, the committee decided that they could not make a strong recommendation specifically in favour of supervised exercise. They made research recommendations instead.

1.1.12.5 Other factors the committee took into account

There were generally no clinically important differences seen when comparing supervised and unsupervised exercises. The committee did not make a recommendation on the level of supervision required for exercise. Different patients may respond better to different approaches and may require more or less supervision to get the same effect. The committee agreed it was not possible to define which groups would be more likely to benefit from supervised exercise. Clinicians should work collaboratively with people with osteoarthritis to achieve the best approach for them.

The committee discussed the use of telehealth interventions to deliver exercise. Evidence for in person and virtual delivery of exercise were included in this review but were not examined as to whether one approach was superior to another. The committee agreed that the preference of the person should be taken into account and that virtual exercise may better suit some people and that in person delivery may be better for others. This should be discussed with the person when tailoring exercise to their needs.

The committee noted that the research identified does not appear to represent the diverse population of people with osteoarthritis. They agreed that any further research should be representative of the population, including people from different family backgrounds, and

socioeconomic backgrounds, disabled people, and people of different ages and genders. Future work should be done to consider the different experiences of people from diverse communities to ensure that the approach taken can be made equitable for everyone. With this in mind the committee subgrouped their research recommendation by these protected characteristics where appropriate while suggesting that people from each group should be included in the research to ensure that it is applicable to the entire population.

1.1.13 Recommendations supported by this evidence review

This evidence review supports recommendations 1.3.1 to 1.3.3 and the research recommendation on exercise. Other evidence supporting these recommendations can be found in the evidence review C.

1.1.14 References

1. Aaboe J, Henriksen M, Bartholdy C, Leonardis J, Rider P, Jørgensen L et al. The effect of quadriceps-strengthening exercise on quadriceps and knee biomichanics during walking in adults with knee osteoarthritis: a randomized controlled trial. Osteoarthritis and Cartilage. 2014; 22:S80-S81

2. Abbott JH, Robertson MC, Chapple C, Pinto D, Wright AA, Leon de la Barra S et al. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee: a randomized controlled trial. 1: clinical effectiveness. Osteoarthritis and Cartilage. 2013; 21(4):525-534

3. Abbott JH, Wilson R, Pinto D, Chapple CM, Wright AA, team MOAT. Incremental clinical effectiveness and cost effectiveness of providing supervised physiotherapy in addition to usual medical care in patients with osteoarthritis of the hip or knee: 2-year results of the MOA randomised controlled trial. Osteoarthritis and Cartilage. 2019; 27(3):424-434

4. Aebischer B, Elsig S, Taeymans J. Effectiveness of physical and occupational therapy on pain, function and quality of life in patients with trapeziometacarpal osteoarthritis - A systematic review and meta-analysis. Hand Therapy. 2016; 21(1):5-15

5. Aglamis B, Toraman NF, Yaman H. Change of quality of life due to exercise training in knee osteoarthritis: sF-36 and Womac. Journal of Back and Musculoskeletal Rehabilitation. 2009; 22(1):43-48

6. Aglamis B, Toraman NF, Yaman H. The effect of a 12-week supervised multicomponent exercise program on knee OA in Turkish women. Journal of Back and Musculoskeletal Rehabilitation. 2008; 21(2):121-128

7. Ahern M, Skyllas J, Wajon A, Hush J. The effectiveness of physical therapies for patients with base of thumb osteoarthritis: Systematic review and meta-analysis. Musculoskeletal Science & Practice. 2018; 35:46-54

8. Alayli G, Kuru O, Bilgici A. The effects of aerobic exercise and home exercise on pain and disability in patients with knee osteoarthritis. Journal of rheumatology and medical rehabilitation. 2007; 18(2):46-50

9. Alghadir AH, Anwer S, Sarkar B, Paul AK, Anwar D. Effect of 6-week retro or forward walking program on pain, functional disability, quadriceps muscle strength, and performance in individuals with knee osteoarthritis: a randomized controlled trial (retro-walking trial). BMC Musculoskeletal Disorders. 2019; 20(1):159

10. Alkatan M, Baker JR, Machin DR, Park W, Akkari AS, Pasha EP et al. Improved function and reduced pain after swimming and cycling training in patients with osteoarthritis. Journal of Rheumatology. 2016; 43(3):666-672

11. Allegrante JP, Kovar PA, MacKenzie CR, Peterson MG, Gutin B. USA impact of a supervised walking program and education on functional status: results from a controlled trial in patients with osteoarthritis of the knee. Patient Education and Counseling. 1991; 18:283-284

12. Allen KD, Arbeeva L, Callahan LF, Golightly YM, Goode AP, Heiderscheit BC et al. Physical therapy vs internet-based exercise training for patients with knee osteoarthritis: results of a randomized controlled trial. Osteoarthritis and Cartilage. 2018; 26(3):383-396

13. Allen KD, Bongiorni D, Bosworth HB, Coffman CJ, Datta SK, Edelman D et al. Group versus individual physical therapy for veterans with knee osteoarthritis: Randomized clinical trial. Physical Therapy. 2016; 96(5):597-608

14. Allen KD, Woolson S, Hoenig HM, Bongiorni D, Byrd J, Caves K et al. Stepped exercise program for patients with knee osteoarthritis : A randomized controlled trial. Annals of Internal Medicine. 2021; 174(3):298-307

15. Allen KD, Yancy WS, Jr., Bosworth HB, Coffman CJ, Jeffreys AS, Datta SK et al. A combined patient and provider intervention for management of osteoarthritis in veterans: A randomized clinical trial. Annals of Internal Medicine. 2016; 164(2):73-83

16. Alonso-Rodriguez AM, Sanchez-Herrero H, Nunes-Hernandez S, Criado-Hernandez B, Gonzalez-Lopez S, Solis-Munoz M. Efficacy of hydrotherapy versus gym treatment in primary total knee prosthesis due to osteoarthritis: a randomized controlled trial. [Spanish]. Anales del sistema sanitario de Navarra. 2021; 44(2):225-241

17. Alrushud AS, Rushton AB, Kanavaki AM, Greig CA. Effect of physical activity and dietary restriction interventions on weight loss and the musculoskeletal function of overweight and obese older adults with knee osteoarthritis: a systematic review and mixed method data synthesis. BMJ Open. 2017; 7(6):e014537

18. An B, Dai K, Zhu Z, Wang Y, Hao Y, Tang T et al. Baduanjin alleviates the symptoms of knee osteoarthritis. Journal of Alternative and Complementary Medicine. 2008; 14(2):167-174

19. An BC, Jiang X, Lu HS, Fang ZY, Wang Y, Dai KR. Effects of baduanjin exercise on knee osteoarthritis: a one-year study. Chinese Journal of Integrative Medicine. 2013; 19(2):143-148

20. Anderson ML, Allen KD, Golightly YM, Arbeeva LS, Goode A, Huffman KM et al. Fall risk and utilization of balance training for adults with symptomatic knee osteoarthritis: Secondary analysis from a randomized clinical trial. Journal of Geriatric Physical Therapy. 2019; 42(2):E39-E44

21. Ansanay RL, Tinduh D, Kusharyaningsih RH. The effect of dexterity and perturbation exercise on knee osteoarthritis through functional balance and power improvement of quadriceps and hamstring. Indian Journal of Forensic Medicine and Toxicology. 2020; 14(2):2307-2312

22. Anwer S, Alghadir A. Effect of isometric quadriceps exercise on muscle strength, pain, and function in patients with knee osteoarthritis: a randomized controlled study. Journal of Physical Therapy Science. 2014; 26(5):745-748

23. Anwer S, Alghadir A, Brismee JM. Effect of home exercise program in patients with knee osteoarthritis: A systematic review and meta-analysis. Journal of Geriatric Physical Therapy. 2016; 39(1):38-48

24. Aoki O, Tsumura N, Kimura A, Okuyama S, Takikawa S, Hirata S. Home stretching exercise is effective for improving knee range of motion and gait in patients with knee osteoarthritis. Journal of Physical Therapy Science. 2009; 21(2):113-119

25. Apparao P, Sudhakar S, Pundarikaksha P, Geetha Mounika R. A comparative study on the effectiveness of neuromuscular training and proprioceptive exercises on pain, strength and function in subjects with knee osteoarthritis. Biomedicine (india). 2017; 37(4):537-544

26. Armagan O, Yilmazer S, Calisir C, Ozgen M, Tascioglu F, Oner S et al. Comparison of the symptomatic and chondroprotective effects of glucosamine sulphate and

exercise treatments in patients with knee osteoarthritis. Journal of Back and Musculoskeletal Rehabilitation. 2015; 28(2):287-293

27. Arnold CM, Faulkner RA. The effect of aquatic exercise and education on lowering fall risk in older adults with hip osteoarthritis. Journal of Aging & Physical Activity. 2010; 18(3):245-260

28. Ashworth N, Chad K, Harrison E, Reeder B, Marshall S. Home versus center based physical activity programs in older adults. Cochrane Database of Systematic Reviews 2005, Issue 1. Art. No.: CD004017. DOI: 10.1002/14651858.CD004017.pub2.

29. Avelar NC, Simao AP, Tossige-Gomes R, Neves CD, Rocha-Vieira E, Coimbra CC et al. The effect of adding whole-body vibration to squat training on the functional performance and self-report of disease status in elderly patients with knee osteoarthritis: a randomized, controlled clinical study. Journal of Alternative and Complementary Medicine. 2011; 17(12):1149-1155

30. Azizi S, Dadarkhah A, Rezasoltani Z, Raeissadat SA, Mofrad RK, Najafi S. Randomized controlled trial of aquatic exercise for treatment of knee osteoarthritis in elderly people. Interventional Medicine and Applied Science. 2020; 11(3):161-167

31. Baker KR, Nelson ME, Felson DT, Layne JE, Sarno R, Roubenoff R. The efficacy of home based progressive strength training in older adults with knee osteoarthritis: a randomized controlled trial. Journal of Rheumatology. 2001; 28(7):1655-1665

32. Bartels E, Juhl C, Christensen R, Hagen K, Danneskiold-Samsøe B, Dagfinrud H et al. Aquatic exercise for the treatment of knee and hip osteoarthritis. Cochrane Database of Systematic Reviews 2016, Issue 3. Art. No.: CD005523. DOI: 10.1002/14651858.CD005523.pub3.

33. Bartholdy C, Juhl C, Christensen R, Lund H, Zhang W, Henriksen M. The role of muscle strengthening in exercise therapy for knee osteoarthritis: A systematic review and meta-regression analysis of randomized trials. Seminars in Arthritis and Rheumatism. 2017; 47(1):9-21

34. Bartholdy C, Klokker L, Bandak E, Bliddal H, Henriksen M. A standardized "rescue" exercise program for symptomatic flare-up of knee osteoarthritis: Description and safety considerations. Journal of Orthopaedic and Sports Physical Therapy. 2016; 46(11):942-946

35. Bautch JC, Malone DG, Vailas AC. Effects of exercise on knee joints with osteoarthritis: a pilot study of biologic markers. Arthritis and Rheumatism. 1997; 10(1):48-55

36. Bearne LM, Walsh NE, Jessep S, Hurley MV. Feasibility of an exercise-based rehabilitation programme for chronic hip pain. Musculoskeletal Care. 2011; 9(3):160-168

37. Beasley J, Ward L, Knipper-Fisher K, Hughes K, Lunsford D, Leiras C. Conservative therapeutic interventions for osteoarthritic finger joints: A systematic review. Journal of Hand Therapy. 2019; 32(2):153-164 e152

38. Beckwee D, Vaes P, Raeymaeckers S, Shahabpour M, Scheerlinck T, Bautmans I. Exercise in knee osteoarthritis: do treatment outcomes relate to bone marrow lesions? A randomized trial. Disability and Rehabilitation. 2017; 39(18):1847-1855

39. Benli Kucuk E, Ozyemisci Taskiran O, Tokgoz N, Meray J. Effects of isokinetic, isometric, and aerobic exercises on clinical variables and knee cartilage volume using magnetic resonance imaging in patients with osteoarthritis. Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi. 2018; 64(1):8-16

40. Bennell KL, Ahamed Y, Jull G, Bryant C, Hunt MA, Forbes AB et al. Physical therapist-delivered pain coping skills training and exercise for knee osteoarthritis: Randomized controlled trial. Arthritis Care and Research. 2016; 68(5):590-602

41. Bennell KL, Campbell PK, Egerton T, Metcalf B, Kasza J, Forbes A et al. Telephone coaching to enhance a home-based physical activity program for knee osteoarthritis: A randomized clinical trial. Arthritis Care and Research. 2017; 69(1):84-94

42. Bennell KL, Egerton T, Bills C, Gale J, Kolt GS, Bunker SJ et al. Addition of telephone coaching to a physiotherapist-delivered physical activity program in people with knee osteoarthritis: a randomised controlled trial protocol. BMC Musculoskeletal Disorders. 2012; 13:246

43. Bennell KL, Hinman RS, Metcalf BR, Buchbinder R, McConnell J, McColl G et al. Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. Annals of the Rheumatic Diseases. 2005; 64(6):906-912

44. Bennell KL, Hunt MA, Wrigley TV, Hunter DJ, Hinman RS. The effects of hip muscle strengthening on knee load, pain, and function in people with knee osteoarthritis: a protocol for a randomised, single-blind controlled trial. BMC Musculoskeletal Disorders. 2007; 8:121

45. Bennell KL, Hunt MA, Wrigley TV, Hunter DJ, McManus FJ, Hodges PW et al. Hip strengthening reduces symptoms but not knee load in people with medial knee osteoarthritis and varus malalignment: a randomised controlled trial. Osteoarthritis and Cartilage. 2010; 18(5):621-628

46. Bennell KL, Kyriakides M, Metcalf B, Egerton T, Wrigley TV, Hodges PW et al. Neuromuscular versus quadriceps strengthening exercise in patients with medial knee osteoarthritis and varus malalignment: a randomized controlled trial. Arthritis & Rheumatology. 2014; 66(4):950-959

47. Bennell KL, Nelligan RK, Kimp AJ, Schwartz S, Kasza J, Wrigley TV et al. What type of exercise is most effective for people with knee osteoarthritis and co-morbid obesity?: The TARGET randomized controlled trial. Osteoarthritis and Cartilage. 2020; 28(6):755-765

48. Bennell KL, Nelligan RK, Rini C, Keefe FJ, Kasza J, French S et al. Effects of internet-based pain coping skills training before home exercise for individuals with hip osteoarthritis (HOPE trial): a randomised controlled trial. Pain. 2018; 159(9):1833-1842

49. Bennell KL, Rini C, Keefe F, French S, Nelligan R, Kasza J et al. Effects of adding an internet-based pain coping skills training protocol to a standardized education and exercise program for people with persistent hip pain (hope trial): Randomized controlled trial protocol. Physical Therapy. 2015; 95(10):1408-1422

50. Beydagi MG, Bazancir Z, Bozgeyik S, Ulger O. Is therapeutic exercise clinically effective in reducing pain intensity in patients with knee osteoarthritis? a systematic review. Topics in Geriatric Rehabilitation. 2021; 37(2):89-103

51. Bezalel T, Carmeli E, Katz-Leurer M. The effect of a group education programme on pain and function through knowledge acquisition and home-based exercise among patients with knee osteoarthritis: a parallel randomised single-blind clinical trial. Physiotherapy. 2010; 96(2):137-143

52. Bieler T, Siersma V, Magnusson SP, Kjaer M, Christensen HE, Beyer N. In hip osteoarthritis, Nordic Walking is superior to strength training and home-based

exercise for improving function. Scandinavian Journal of Medicine and Science in Sports. 2017; 27(8):873-886

53. Bilgici A, Akdeniz O, Kuru O, Unlu S, Ulusoy H. The effect of aerobic exercise program versus a home based exercise therapy on pain and functional disability in patients with knee osteoarthritis. Journal of rheumatology and medical rehabilitation. 2005; 16(1):10-17

54. Boeer J, Mueller O, Krauss I, Haupt G, Axmann D, Horstmann T. Effects of a sensory-motor exercise program for older adults with osteoarthritis or prosthesis of the hip using measurements made by the Posturomed oscillatory platform. Journal of Geriatric Physical Therapy. 2010; 33(1):10-15

55. Bokaeian HR, Bakhtiary AH, Mirmohammadkhani M, Moghimi J. The effect of adding whole body vibration training to strengthening training in the treatment of knee osteoarthritis: A randomized clinical trial. Journal of Bodywork and Movement Therapies. 2016; 20(2):334-340

56. Bokaeian HR, Esfandiarpour F, Zahednejad S, Mohammadi HK, Farahmand F. Effects of an exercise therapy targeting knee kinetics on pain, function, and gait kinetics in patients with knee osteoarthritis: A randomized clinical trial. Adapted Physical Activity Quarterly. 2021; 38(3):377-395

57. Borjesson M, Robertson E, Weidenhielm L, Mattsson E, Olsson E. Physiotherapy in knee osteoarthrosis: effect on pain and walking. Physiotherapy Research International. 1996; 1(2):89-97

58. Bossen D, Buskermolen M, Veenhof C, de Bakker D, Dekker J. Adherence to a web-based physical activity intervention for patients with knee and/or hip osteoarthritis: a mixed method study. Journal of Medical Internet Research. 2013; 15(10):e223

59. Bossen D, Veenhof C, Van Beek KE, Spreeuwenberg PM, Dekker J, De Bakker DH. Effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis: randomized controlled trial. Journal of Medical Internet Research. 2013; 15(11):e257

60. Bove AM, Smith KJ, Bise CG, Fritz JM, Childs JD, Brennan GP et al. Exercise, manual therapy, and booster sessions in knee osteoarthritis: Cost-effectiveness analysis from a multicenter randomized controlled trial. Physical Therapy. 2018; 98(1):16-27

61. Braghin RMB, Libardi EC, Junqueira C, Nogueira-Barbosa MH, de Abreu DCC. Exercise on balance and function for knee osteoarthritis: A randomized controlled trial. Journal of Bodywork and Movement Therapies. 2018; 22(1):76-82

62. Brandao GS, Sampaio AAC, Damas Andrade L, Fonseca AL, Campos FKR, Silva AS et al. Home physical exercise improves functional mobility and quality of life in the elderly: A CONSORT-prospective, randomised controlled clinical trial. International Journal of Clinical Practice. 2021; 75(8):e14347

63. Bressel E, Wing JE, Miller AI, Dolny DG. High-intensity interval training on an aquatic treadmill in adults with osteoarthritis: effect on pain, balance, function, and mobility. Journal of Strength and Conditioning Research. 2014; 28(8):2088-2096

64. Bricca A, Juhl CB, Steultjens M, Wirth W, Roos EM. Impact of exercise on articular cartilage in people at risk of, or with established, knee osteoarthritis: a systematic review of randomised controlled trials. British Journal of Sports Medicine. 2018; 22:22

65. Brismee JM, Paige RL, Chyu MC, Boatright JD, Hagar JM, McCaleb JA et al. Group and home-based tai chi in elderly subjects with knee osteoarthritis: a randomized controlled trial. Clinical Rehabilitation. 2007; 21(2):99-111

66. Brosseau L, Wells GA, Kenny GP, Reid R, Maetzel A, Tugwell P et al. The implementation of a community-based aerobic walking program for mild to moderate knee osteoarthritis: a knowledge translation randomized controlled trial: part II: clinical outcomes. BMC Public Health. 2012; 12:1073

67. Bruce-Brand RA, Walls RJ, Ong JC, Emerson BS, O'Byrne JM, Moyna NM. Effects of home-based resistance training and neuromuscular electrical stimulation in knee osteoarthritis: a randomized controlled trial. BMC Musculoskeletal Disorders. 2012; 13:118

68. Bryk FF, Dos Reis AC, Fingerhut D, Araujo T, Schutzer M, Cury Rde P et al. Exercises with partial vascular occlusion in patients with knee osteoarthritis: a randomized clinical trial. Knee Surgery, Sports Traumatology, Arthroscopy. 2016; 24(5):1580-1586

69. Burrows NJ, Booth J, Sturnieks DL, Barry BK. Acute resistance exercise and pressure pain sensitivity in knee osteoarthritis: a randomised crossover trial. Osteoarthritis and Cartilage. 2014; 22(3):407-414

70. Cadmus L, Patrick MB, Maciejewski ML, Topolski T, Belza B, Patrick DL. Community-based aquatic exercise and quality of life in persons with osteoarthritis. Medicine and Science in Sports and Exercise. 2010; 42(1):8-15

71. Callaghan MJ, Oldham JA, Hunt J. An evaluation of exercise regimes for patients with osteoarthritis of the knee: a single-blind randomized controlled trial. Clinical Rehabilitation. 1995; 9(3):213-218

72. Cantero-Tellez R, Naughton N, Algar LA, Medina-Porqueres I, Cruz-Gambero L, Valdes KA. Proprioceptive neuromuscular facilitation protocol for thumb osteoarthritis: A pilot study. Hand (New York, NY). 2021:1558944721990785

73. Carlson NL, Christopherson Z, Arnall E, Mohn S, Holton K, Marshall L et al. A pilot study on the effects of strength and aerobic conditioning in patients with hip osteoarthritis. Osteoarthritis and Cartilage. 2011; 1:S212

74. Carmona-Teres V, Lumillo-Gutierrez I, Jodar-Fernandez L, Rodriguez-Blanco T, Moix-Queralto J, Pujol-Ribera E et al. Effectiveness and cost-effectiveness of a health coaching intervention to improve the lifestyle of patients with knee osteoarthritis: cluster randomized clinical trial. BMC Musculoskeletal Disorders. 2015; 16:38

75. Casilda-Lopez J, Valenza MC, Cabrera-Martos I, Diaz-Pelegrina A, Moreno-Ramirez MP, Valenza-Demet G. Effects of a dance-based aquatic exercise program in obese postmenopausal women with knee osteoarthritis: a randomized controlled trial. Menopause. 2017; 24(7):768-773

76. Ceballos-Laita L, Estebanez-de-Miguel E, Martin-Nieto G, Bueno-Gracia E, Fortun-Agud M, Jimenez-Del-Barrio S. Effects of non-pharmacological conservative treatment on pain, range of motion and physical function in patients with mild to moderate hip osteoarthritis. A systematic review. Complementary Therapies in Medicine. 2019; 42:214-222

77. Chaipinyo K, Karoonsupcharoen O. No difference between home-based strength training and home-based balance training on pain in patients with knee osteoarthritis: a randomised trial. Australian Journal of Physiotherapy. 2009; 55(1):25-30

78. Chamberlain MA, Care G, Harfield B. Physiotherapy in osteoarthrosis of the knees. A controlled trial of hospital versus home exercises. International Rehabilitation Medicine. 1982; 4(2):101-106

79. Chang TF, Liou TH, Chen CH, Huang YC, Chang KH. Effects of elastic-band exercise on lower-extremity function among female patients with osteoarthritis of the knee. Disability and Rehabilitation. 2012; 34(20):1727-1735

80. Chao J, Jing Z, Xuehua B, Peilei Y, Qi G. Effect of systematic exercise rehabilitation on patients with knee osteoarthritis: A randomized controlled trial. Cartilage. 2020:1947603520903443

81. Chen H, Zheng X, Huang H, Liu C, Wan Q, Shang S. The effects of a home-based exercise intervention on elderly patients with knee osteoarthritis: a quasi-experimental study. BMC Musculoskeletal Disorders. 2019; 20(1):160

82. Chen PY, Song CY, Yen HY, Lin PC, Chen SR, Lu LH et al. Impacts of tai chi exercise on functional fitness in community-dwelling older adults with mild degenerative knee osteoarthritis: a randomized controlled clinical trial. BMC Geriatrics. 2021; 21(1):449

83. Chen Z, Ye X, Wang Y, Shen Z, Wu J, Chen W et al. The efficacy of backward walking on static stability, proprioception, pain, and physical function of patients with knee osteoarthritis: A randomized controlled trial. Evidence-Based Complementary & Alternative Medicine: eCAM. 2021; 2021:5574966

84. Cheung C, Wyman JF, Bronas U, McCarthy T, Rudser K, Mathiason MA. Managing knee osteoarthritis with yoga or aerobic/strengthening exercise programs in older adults: a pilot randomized controlled trial. Rheumatology International. 2017; 37(3):389-398

85. Cheung C, Wyman JF, Resnick B, Savik K. Yoga for managing knee osteoarthritis in older women: a pilot randomized controlled trial. BMC Complementary and Alternative Medicine. 2014; 14:160

86. Cheung RTH, Ho KKW, Au IPH, An WW, Zhang JHW, Chan ZYS et al. Immediate and short-term effects of gait retraining on the knee joint moments and symptoms in patients with early tibiofemoral joint osteoarthritis: a randomized controlled trial. Osteoarthritis and Cartilage. 2018; 26(11):1479-1486

87. Cho Y, Kim M, Lee W. Effect of proprioceptive training on foot posture, lower limb alignment, and knee adduction moment in patients with degenerative knee osteoarthritis: a randomized controlled trial. Journal of Physical Therapy Science. 2015; 27(2):371-374

88. Chopp-Hurley JN, Brenneman EC, Wiebenga EG, Bulbrook B, Keir PJ, Maly MR. Randomized controlled trial investigating the role of exercise in the workplace to improve work ability, performance, and patient-reported symptoms among older workers with osteoarthritis. Journal of Occupational and Environmental Medicine. 2017; 59(6):550-556

89. Christensen R, Henriksen M, Leeds AR, Gudbergsen H, Christensen P, Sorensen TJ et al. Effect of weight maintenance on symptoms of knee osteoarthritis in obese patients: a twelve-month randomized controlled trial. Arthritis Care and Research. 2015; 67(5):640-650

90. Clausen B, Holsgaard-Larsen A, Sondergaard J, Christensen R, Andriacchi TP, Roos EM. The effect on knee-joint load of instruction in analgesic use compared with neuromuscular exercise in patients with knee osteoarthritis: study protocol for a

randomized, single-blind, controlled trial (the EXERPHARMA trial). Trials [Electronic Resource]. 2014; 15:444

91. Cochrane T, Davey RC, Matthes Edwards SM. Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis. Health Technology Assessment. 2005; 9(31):iii-iv, ix-xi, 1-114

92. Coleman S, Briffa NK, Carroll G, Inderjeeth C, Cook N, McQuade J. A randomised controlled trial of a self-management education program for osteoarthritis of the knee delivered by health care professionals. Arthritis Research & Therapy. 2012; 14(1):R21

93. Cotofana S, Ring-Dimitriou S, Hudelmaier M, Himmer M, Wirth W, Sanger AM et al. Effects of exercise intervention on knee morphology in middle-aged women: a longitudinal analysis using magnetic resonance imaging. Cells Tissues Organs. 2010; 192(1):64-72

94. Coudeyre E, Jegu AG, Giustanini M, Marrel JP, Edouard P, Pereira B. Isokinetic muscle strengthening for knee osteoarthritis: A systematic review of randomized controlled trials with meta-analysis. Annals of Physical and Rehabilitation Medicine. 2016; 59(3):207-215

95. Cuperus N, Hoogeboom TJ, Kersten CC, den Broeder AA, Vlieland TP, van den Ende CH. Randomized trial of the effectiveness of a non-pharmacological multidisciplinary face-to-face treatment program on daily function compared to a telephone-based treatment program in patients with generalized osteoarthritis. Osteoarthritis and Cartilage. 2015; 23(8):1267-1275

96. Curtis L, Burns A. Unit costs of health and social care 2021. Canterbury. University of Kent, 2021. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

97. da Silva FS, de Melo FE, do Amaral MM, Caldas VV, Pinheiro IL, Abreu BJ et al. Efficacy of simple integrated group rehabilitation program for patients with knee osteoarthritis: Single-blind randomized controlled trial. Journal of Rehabilitation Research and Development. 2015; 52(3):309-322

98. Danazumi MS, Ibrahim SU, Yakasai AM, Dermody G, Bello B, Kaka B. A comparison between the effect of combined chain exercises plus kinesio taping with combined chain exercises alone in knee osteoarthritis: A randomized clinical trial. American Journal of Physical Medicine and Rehabilitation. 2021; 100(11):1070-1077

99. Davenport BJ, Jansen V, Yeandle N. Pilot randomized controlled trial comparing specific dynamic stability exercises with general exercises for thumb carpometacarpal joint osteoarthritis. Hand Therapy. 2012; 17(3):60-67

100. de Matos Brunelli Braghin R, Libardi EC, Junqueira C, Rodrigues NC, Nogueira-Barbosa MH, Renno ACM et al. The effect of low-level laser therapy and physical exercise on pain, stiffness, function, and spatiotemporal gait variables in subjects with bilateral knee osteoarthritis: a blind randomized clinical trial. Disability and Rehabilitation. 2019; 41(26):3165-3172

101. de Rooij M, van der Leeden M, Cheung J, van der Esch M, Hakkinen A, Haverkamp D et al. Efficacy of tailored exercise therapy on physical functioning in patients with knee osteoarthritis and comorbidity: A randomized controlled trial. Arthritis Care and Research. 2017; 69(6):807-816

102. de Vos BC, Landsmeer MLA, van Middelkoop M, Oei EHG, Krul M, Bierma-Zeinstra SMA et al. Long-term effects of a lifestyle intervention and oral glucosamine sulphate

in primary care on incident knee OA in overweight women. Rheumatology. 2017; 56(8):1326-1334

103. Deepeshwar S, Tanwar M, Kavuri V, Budhi RB. Effect of yoga based lifestyle intervention on patients with knee osteoarthritis: A randomized controlled trial. Frontiers in psychiatry Frontiers Research Foundation. 2018; 9:180

104. Dias JM, Cisneros L, Dias R, Fritsch C, Gomes W, Pereira L et al. Hydrotherapy improves pain and function in older women with knee osteoarthritis: a randomized controlled trial. Brazilian Journal of Physical Therapy. 2017; 21(6):449-456

105. Dias RC, Dias JM, Ramos LR. Impact of an exercise and walking protocol on quality of life for elderly people with OA of the knee. Physiotherapy Research International. 2003; 8(3):121-130

106. Diracoglu D, Aydin R, Baskent A, Celik A. Effects of kinesthesia and balance exercises in knee osteoarthritis. JCR: Journal of Clinical Rheumatology. 2005; 11(6):303-310

107. DiracogluD, BaskentA, CelikA, IsseverH, AydinR. Long-term effects of kinesthesia/balance and strengthening exercises on patients with knee osteoarthritis: a one-year follow-up study. Journal of Back and Musculoskeletal Rehabilitation. 2008; 21(4):253-262

108. Doi T, Akai M, Fujino K, Iwaya T, Kurosawa H, Hayashi K et al. Effect of home exercise of quadriceps on knee osteoarthritis compared with nonsteroidal antiinflammatory drugs: a randomized controlled trial. American Journal of Physical Medicine and Rehabilitation. 2008; 87(4):258-269

109. Dong R, Wu Y, Xu S, Zhang L, Ying J, Jin H et al. Is aquatic exercise more effective than land-based exercise for knee osteoarthritis? Medicine. 2018; 97(52):e13823

110. Dong Y, Wu W, Zheng J, Chen S, Qiao J, Wang X. Whole body vibration exercise for chronic musculoskeletal pain: A Systematic review and meta-analysis of randomized controlled trials. Archives of Physical Medicine and Rehabilitation. 2019; 100(11):2167-2178

111. Duman I, Taskaynatan MA, Mohur H, Tan AK. Assessment of the impact of proprioceptive exercises on balance and proprioception in patients with advanced knee osteoarthritis. Rheumatology International. 2012; 32(12):3793-3798

112. Durmus D, Alayli G, Aliyazicioglu Y, Buyukakincak O, Canturk F. Effects of glucosamine sulfate and exercise therapy on serum leptin levels in patients with knee osteoarthritis: preliminary results of randomized controlled clinical trial. Rheumatology International. 2013; 33(3):593-599

113. Durmus D, Alayli G, Bayrak IK, Canturk F. Assessment of the effect of glucosamine sulfate and exercise on knee cartilage using magnetic resonance imaging in patients with knee osteoarthritis: a randomized controlled clinical trial. Journal of Back and Musculoskeletal Rehabilitation. 2012; 25(4):275-284

114. Dziedzic K, Nicholls E, Hill S, Hammond A, Handy J, Thomas E et al. Self-management approaches for osteoarthritis in the hand: a 2x2 factorial randomised trial. Annals of the Rheumatic Diseases. 2015; 74(1):108-118

115. Dziedzic KS, Hill S, Nicholls E, Hammond A, Myers H, Whitehurst T et al. Self management, joint protection and exercises in hand osteoarthritis: a randomised controlled trial with cost effectiveness analyses. BMC Musculoskeletal Disorders. 2011; 12:156

116. Ebnezar J, Nagarathna R, Bali Y, Nagendra HR. Effect of an integrated approach of yoga therapy on quality of life in osteoarthritis of the knee joint: A randomized control study. International Journal of Yoga. 2011; 4(2):55-63

117. Ebnezar J, Nagarathna R, Yogitha B, Nagendra HR. Effect of integrated yoga therapy on pain, morning stiffness and anxiety in osteoarthritis of the knee joint: A randomized control study. International Journal of Yoga. 2012; 5(1):28-36

118. Ebnezar J, Nagarathna R, Yogitha B, Nagendra HR. Effects of an integrated approach of hatha yoga therapy on functional disability, pain, and flexibility in osteoarthritis of the knee joint: a randomized controlled study. Journal of Alternative and Complementary Medicine. 2012; 18(5):463-472

119. The effects of closed kinetic chain exercise on articular cartilage morphology: myth or reality? A randomized controlled clinical trial. Turkiye Fiziksel Tip ve Rehabilitasyon Dergisi. 2016; 62(1):28-36

120. Efficacy of exercise rehabilitation in patients with knee osteoarthritis. Zhonghua yi xue za zhi. 2019; 99(41):3255-3259

121. Ettinger WH, Jr., Burns R, Messier SP, Applegate W, Rejeski WJ, Morgan T et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). JAMA. 1997; 277(1):25-31

122. Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. Rheumatology International. 2002; 22(3):103-106

123. Farr JN, Going SB, McKnight PE, Kasle S, Cussler EC, Cornett M. Progressive resistance training improves overall physical activity levels in patients with early osteoarthritis of the knee: a randomized controlled trial. Physical Therapy. 2010; 90(3):356-366

124. Fernandes L, Storheim K, Sandvik L, Nordsletten L, Risberg MA. Efficacy of patient education and supervised exercise vs patient education alone in patients with hip osteoarthritis: a single blind randomized clinical trial. Osteoarthritis and Cartilage. 2010; 18(10):1237-1243

125. Fernandopulle S, Perry M, Manlapaz D, Jayakaran P. Effect of land-based generic physical activity interventions on pain, physical function, and physical performance in hip and knee osteoarthritis: A systematic review and meta-analysis. American Journal of Physical Medicine and Rehabilitation. 2017; 96(11):773-792

126. Ferreira GE, Robinson CC, Wiebusch M, Viero CC, da Rosa LH, Silva MF. The effect of exercise therapy on knee adduction moment in individuals with knee osteoarthritis: A systematic review. Clinical Biomechanics. 2015; 30(6):521-527

127. Fisken AL, Waters DL, Hing WA, Steele M, Keogh JW. Comparative effects of 2 aqua exercise programs on physical function, balance, and perceived quality of life in older adults with osteoarthritis. Journal of geriatric physical therapy (2001). 2015; 38(1):17-27

128. Fitzgerald GK, Fritz JM, Childs JD, Brennan GP, Talisa V, Gil AB et al. Exercise, manual therapy, and use of booster sessions in physical therapy for knee osteoarthritis: a multi-center, factorial randomized clinical trial. Osteoarthritis and Cartilage. 2016; 24(8):1340-1349

129. Fitzgerald GK, Piva SR, Gil AB, Wisniewski SR, Oddis CV, Irrgang JJ. Agility and perturbation training techniques in exercise therapy for reducing pain and improving

function in people with knee osteoarthritis: a randomized clinical trial. Physical Therapy. 2011; 91(4):452-469

130. Focht BC, Garver MJ, Devor ST, Dials J, Lucas AR, Emery CF et al. Group-mediated physical activity promotion and mobility in sedentary patients with knee osteoarthritis: results from the IMPACT-pilot trial. Journal of Rheumatology. 2014; 41(10):2068-2077

131. Focht BC, Garver MJ, Lucas AR, Devor ST, Emery CF, Hackshaw KV et al. A group-mediated physical activity intervention in older knee osteoarthritis patients: effects on social cognitive outcomes. Journal of Behavioral Medicine. 2017; 40(3):530-537

132. Focht BC, Gauvin L, Rejeski WJ. The contribution of daily experiences and acute exercise to fluctuations in daily feeling states among older, obese adults with knee osteoarthritis. Journal of Behavioral Medicine. 2004; 27(2):101-121

133. Focht BC, Rejeski WJ, Ambrosius WT, Katula JA, Messier SP. Exercise, self-efficacy, and mobility performance in overweight and obese older adults with knee osteoarthritis. Arthritis and Rheumatism. 2005; 53(5):659-665

134. Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis--a randomised controlled trial comparing a gym based and a hydrotherapy based strengthening programme. Annals of the Rheumatic Diseases. 2003; 62(12):1162-1167

135. Foroughi N, Smith RM, Lange AK, Baker MK, Fiatarone Singh MA, Vanwanseele B. Lower limb muscle strengthening does not change frontal plane moments in women with knee osteoarthritis: A randomized controlled trial. Clinical Biomechanics. 2011; 26(2):167-174

136. Foroughi N, Smith RM, Lange AK, Singh MA, Vanwanseele B. Progressive resistance training and dynamic alignment in osteoarthritis: A single-blind randomised controlled trial. Clinical Biomechanics. 2011; 26(1):71-77

137. Foster NE, Healey EL, Holden MA, Nicholls E, Whitehurst DG, Jowett S et al. A multicentre, pragmatic, parallel group, randomised controlled trial to compare the clinical and cost-effectiveness of three physiotherapy-led exercise interventions for knee osteoarthritis in older adults: the BEEP trial protocol (ISRCTN: 93634563). BMC Musculoskeletal Disorders. 2014; 15:254

138. Fransen M, Crosbie J, Edmonds J. Physical therapy is effective for patients with osteoarthritis of the knee: a randomized controlled clinical trial. Journal of Rheumatology. 2001; 28(1):156-164

139. Fransen M, McConnell S, Bell M. Exercise for osteoarthritis of the hip or knee. Cochrane Database of Systematic Reviews 2003, Issue 3. Art. No.: CD004286.

140. Fransen M, McConnell S, Harmer A, Van dEM, Simic M, Bennell K. Exercise for osteoarthritis of the knee. Cochrane Database of Systematic Reviews 2015, Issue 1. Art. No.: CD004376. DOI: 10.1002/14651858.CD004376.pub3.

141. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. British Journal of Sports Medicine. 2015; 49(24):1554-1557

142. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Does land-based exercise reduce pain and disability associated with hip osteoarthritis? A meta-analysis of randomized controlled trials. Osteoarthritis and Cartilage. 2010; 18(5):613-620

143. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. Cochrane Database of Systematic Reviews 2014, Issue 4. Art. No.: CD007912. DOI: 10.1002/14651858.CD007912.pub2.

144. Fransen M, Nairn L, Winstanley J, Lam P. Hydrotherapy or taichi for osteoarthritis. Internal Medicine Journal. 2006; 36:A69

145. Fransen M, Nairn L, Winstanley J, Lam P, Edmonds J. Physical activity for osteoarthritis management: a randomized controlled clinical trial evaluating hydrotherapy or Tai Chi classes. Arthritis and Rheumatism. 2007; 57(3):407-414

146. French HP, Cusack T, Brennan A, Caffrey A, Conroy R, Cuddy V et al. Exercise and manual physiotherapy arthritis research trial (EMPART) for osteoarthritis of the hip: a multicenter randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2013; 94(2):302-314

147. French HP, Cusack T, Brennan A, White B, Gilsenan C, Fitzpatrick M et al. Exercise and manual physiotherapy arthritis research trial (EMPART): a multicentre randomised controlled trial. BMC Musculoskeletal Disorders. 2009; 10:9

148. French HP, Galvin R, Abbott JH, Fransen M. Adjunctive therapies in addition to land-based exercise therapy for osteoarthritis of the hip or knee. Cochrane Database of Systematic Reviews 2015, Issue Art. No.: CD011915. DOI: http://dx.doi.org/10.1002/14651858.CD011915.

149. Fukumoto Y, Tateuchi H, Ikezoe T, Tsukagoshi R, Akiyama H, So K et al. Effects of high-velocity resistance training on muscle function, muscle properties, and physical performance in individuals with hip osteoarthritis: a randomized controlled trial. Clinical Rehabilitation. 2014; 28(1):48-58

150. Fukumoto Y, Tateuchi H, Tsukagoshi R, Okita Y, Akiyama H, So K et al. Effects of high- and low-velocity resistance training on gait kinematics and kinetics in individuals with hip osteoarthritis: A randomized controlled trial. American Journal of Physical Medicine and Rehabilitation. 2017; 96(6):417-423

151. Garfinkel MS, Schumacher HR, Jr., Husain A, Levy M, Reshetar RA. Evaluation of a yoga based regimen for treatment of osteoarthritis of the hands. Journal of Rheumatology. 1994; 21(12):2341-2343

152. Ghroubi S, Elleuch H, Kaffel N, Echikh T, Abid M, Elleuch MH. Contribution of exercise and diet in the management of knee osteoarthritis in the obese. Annales de readaptation et de medecine physique. 2008; 51(8):663-670

153. Gill SD, McBurney H, Schulz DL. Land-based versus pool-based exercise for people awaiting joint replacement surgery of the hip or knee: results of a randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2009; 90(3):388-394

154. Goh SL, Persson MSM, Stocks J, Hou Y, Lin J, Hall MC et al. Efficacy and potential determinants of exercise therapy in knee and hip osteoarthritis: A systematic review and meta-analysis. Annals of Physical and Rehabilitation Medicine. 2019; 62(5):356-365

155. Goh SL, Persson MSM, Stocks J, Hou Y, Welton NJ, Lin J et al. Relative efficacy of different exercises for pain, function, performance and quality of life in knee and hip osteoarthritis: Systematic review and network meta-analysis. Sports Medicine. 2019; 49(5):743-761

156. Goksen A, Can F, Yilmaz S, Korkusuz F. Comparison of different neuromuscular facilitation techniques and conventional physiotherapy in knee osteoarthritis. Turkish journal of medical sciences. 2021; 51(6):3089-3097

157. Gomiero AB, Kayo A, Abraao M, Peccin MS, Grande AJ, Trevisani VF. Sensory-motor training versus resistance training among patients with knee osteoarthritis: randomized single-blind controlled trial. Sao Paulo Medical Journal = Revista Paulista de Medicina. 2018; 136(1):44-50

158. Gondhalekar GA, Deo MV. Retrowalking as an adjunct to conventional treatment versus conventional treatment alone on pain and disability in patients with acute exacerbation of chronic knee osteoarthritis: a randomized clinical trial. North American Journal of Medical Sciences. 2013; 5(2):108-112

159. Goonasegaran AR, Suhaimi A, Mokhtar AH. Retro-walking improves symptoms, pain, and function in primary knee osteoarthritis: a randomised control trial. Journal of Sports Medicine and Physical Finess. 2022; 62(2):229-237

160. Green J, McKenna F, Chamberlain MA. Osteoarthritis of the hip - are home exercises as useful as hydrotherapy? Clinical Rehabilitation. 1988; 2:253-258

161. Green J, McKenna F, Redfern EJ, Chamberlain MA. Home exercises are as effective as outpatient hydrotherapy for osteoarthritis of the hip. British Journal of Rheumatology. 1993; 32(9):812-815

162. Gudbergsen H, Boesen M, Lohmander LS, Christensen R, Henriksen M, Bartels EM et al. Weight loss is effective for symptomatic relief in obese subjects with knee osteoarthritis independently of joint damage severity assessed by high-field MRI and radiography. Osteoarthritis and Cartilage. 2012; 20(6):495-502

163. Gur H, Cakin N, Akova B, Okay E, Kucukoglu S. Concentric versus combined concentric-eccentric isokinetic training: effects on functional capacity and symptoms in patients with osteoarthrosis of the knee. Archives of Physical Medicine and Rehabilitation. 2002; 83(3):308-316

164. Halbert J, Crotty M, Weller D, Ahern M, Silagy C. Primary care-based physical activity programs: effectiveness in sedentary older patients with osteoarthritis symptoms. Arthritis and Rheumatism. 2001; 45(3):228-234

165. Hale LA, Waters D, Herbison P. A randomized controlled trial to investigate the effects of water-based exercise to improve falls risk and physical function in older adults with lower-extremity osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2012; 93(1):27-34

166. Hall M, Castelein B, Wittoek R, Calders P, Van Ginckel A. Diet-induced weight loss alone or combined with exercise in overweight or obese people with knee osteoarthritis: A systematic review and meta-analysis. Seminars in Arthritis and Rheumatism. 2019; 48(5):765-777

167. Hanada K, Hara M, Hirakawa Y, Hoshi K, Ito K, Gamada K. Immediate effects of leg-press exercises with tibial internal rotation on individuals with medial knee osteoarthritis. Physiotherapy Research International. 2018; 23(4):e1725

168. Handa N, Yamamoto H, Tani T, Kawakami T, Takemasa R. The effect of trunk muscle exercises in patients over 40 years of age with chronic low back pain. Journal of Orthopaedic Science. 2000; 5(3):210-216

169. Harris-Hayes M, Steger-May K, A MB, Mueller MJ, Clohisy JC, Fitzgerald GK. One-year outcomes following physical therapist-led intervention for chronic hip-related groin pain: Ancillary analysis of a pilot multicenter randomized clinical trial. Journal of Orthopaedic Research. 2021; 39(11):2409-2418

170. Hartman CA, Manos TM, Winter C, Hartman DM, Li B, Smith JC. Effects of T'ai Chi training on function and quality of life indicators in older adults with osteoarthritis. Journal of the American Geriatrics Society. 2000; 48(12):1553-1559

171. Hasegawa M, Yamazaki S, Kimura M, Nakano K, Yasumura S. Community-based exercise program reduces chronic knee pain in elderly Japanese women at high risk of requiring long-term care: a non-randomized controlled trial. Geriatrics & gerontology international. 2013; 13(1):167-174

172. Hay EM, Foster NE, Thomas E, Peat G, Phelan M, Yates HE et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomised trial. BMJ. 2006; 333(7576):995

173. Hennig T, Haehre L, Hornburg VT, Mowinckel P, Norli ES, Kjeken I. Effect of home-based hand exercises in women with hand osteoarthritis: a randomised controlled trial. Annals of the Rheumatic Diseases. 2015; 74(8):1501-1508

174. Henriksen M, Hansen JB, Klokker L, Bliddal H, Christensen R. Comparable effects of exercise and analgesics for pain secondary to knee osteoarthritis: a meta-analysis of trials included in Cochrane systematic reviews. Journal of Comparative Effectiveness Research. 2016; 5(4):417-431

175. Henriksen M, Hansen JB, Klokker L, Bliddal H, Christensen R. Exercise versus analgesics for knee osteoarthritis pain: a meta-epidemiological study of cochrane systematic reviews. Osteoarthritis and Cartilage. 2015; 23:A172

176. Henriksen M, Klokker L, Bartholdy C, Schjoedt-Jorgensen T, Bandak E, Bliddal H. No effects of functional exercise therapy on walking biomechanics in patients with knee osteoarthritis: exploratory outcome analyses from a randomised trial. BMJ Open Sport & Exercise Medicine. 2016; 2(1):e000230

177. Henriksen M, Klokker L, Graven-Nielsen T, Bartholdy C, Schjodt Jorgensen T, Bandak E et al. Association of exercise therapy and reduction of pain sensitivity in patients with knee osteoarthritis: a randomized controlled trial. Arthritis Care and Research. 2014; 66(12):1836-1843

178. Hermann A, Holsgaard-Larsen A, Zerahn B, Mejdahl S, Overgaard S. Preoperative progressive explosive-type resistance training is feasible and effective in patients with hip osteoarthritis scheduled for total hip arthroplasty--a randomized controlled trial. Osteoarthritis and Cartilage. 2016; 24(1):91-98

179. Hernandez D, Dimaro M, Navarro E, Dorado J, Accoce M, Salzberg S et al. Efficacy of core exercises in patients with osteoarthritis of the knee: A randomized controlled clinical trial. Journal of Bodywork and Movement Therapies. 2019; 23(4):881-887

180. Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. Physical Therapy. 2007; 87(1):32-43

181. Hiyama Y, Yamada M, Kitagawa A, Tei N, Okada S. A four-week walking exercise programme in patients with knee osteoarthritis improves the ability of dual-task performance: a randomized controlled trial. Clinical Rehabilitation. 2012; 26(5):403-412

182. Holm PM, Petersen KK, Wernbom M, Schroder HM, Arendt-Nielsen L, Skou ST. Strength training in addition to neuromuscular exercise and education in individuals with knee osteoarthritis-the effects on pain and sensitization. European Journal of Pain. 2021; 25(9):1898-1911

183. Holm PM, Schroder HM, Wernbom M, Skou ST. Low-dose strength training in addition to neuromuscular exercise and education in patients with knee osteoarthritis in secondary care - a randomized controlled trial. Osteoarthritis and Cartilage. 2020; 28(6):744-754

184. Holsgaard-Larsen A, Christensen R, Clausen B, Sondergaard J, Andriacchi TP, Roos EM. One year effectiveness of neuromuscular exercise compared with instruction in analgesic use on knee function in patients with early knee osteoarthritis: the EXERPHARMA randomized trial. Osteoarthritis and Cartilage. 2018; 26(1):28-33

185. Holsgaard-Larsen A, Clausen B, Sondergaard J, Christensen R, Andriacchi TP, Roos EM. The effect of instruction in analgesic use compared with neuromuscular exercise on knee-joint load in patients with knee osteoarthritis: a randomized, single-blind, controlled trial. Osteoarthritis and Cartilage. 2017; 25(4):470-480

186. Horstmann T, Mayer F, Heitkamp HC, Merk J, Axmann D, Bork H et al. Individual isokinetic strength training in patients with gonarthrosis. Zeitschrift für Rheumatologie. 2000; 59(2):93-100

187. Howe TE. Exercise for osteoarthritis of the hip and knee. Annual Review of Gerontology and Geriatrics. 2016; 36(1):155-168

188. Hu L, Wang Y, Liu X, Ji X, Ma Y, Man S et al. Tai Chi exercise can ameliorate physical and mental health of patients with knee osteoarthritis: systematic review and meta-analysis. Clinical Rehabilitation. 2021; 35(1):64-79

189. Hu X, Lai Z, Wang L. Effects of Taichi exercise on knee and ankle proprioception among individuals with knee osteoarthritis. Research in Sports Medicine. 2020; 28(2):268-278

190. Huang L, Guo B, Xu F, Zhao J. Effects of quadriceps functional exercise with isometric contraction in the treatment of knee osteoarthritis. International Journal of Rheumatic Diseases. 2018; 21(5):952-959

191. Huang MH, Lin YS, Lee CL, Yang RC. Use of ultrasound to increase effectiveness of isokinetic exercise for knee osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2005; 86(8):1545-1551

192. Huang MH, Lin YS, Yang RC, Lee CL. A comparison of various therapeutic exercises on the functional status of patients with knee osteoarthritis. Seminars in Arthritis and Rheumatism. 2003; 32(6):398-406

193. Huang MH, Yang RC, Lee CL, Chen TW, Wang MC. Preliminary results of integrated therapy for patients with knee osteoarthritis. Arthritis and Rheumatism. 2005; 53(6):812-820

194. Hughes SL, Seymour RB, Campbell R, Pollak N, Huber G, Sharma L. Impact of the fit and strong intervention on older adults with osteoarthritis. Gerontologist. 2004; 44(2):217-228

195. Hughes SL, Seymour RB, Campbell RT, Huber G, Pollak N, Sharma L et al. Long-term impact of Fit and Strong! on older adults with osteoarthritis. Gerontologist. 2006; 46(6):801-814

196. Hunt MA, Charlton JM, Krowchuk NM, Tse CTF, Hatfield GL. Clinical and biomechanical changes following a 4-month toe-out gait modification program for people with medial knee osteoarthritis: a randomized controlled trial. Osteoarthritis and Cartilage. 2018; 26(7):903-911

197. Hunt MA, Pollock CL, Kraus VB, Saxne T, Peters S, Huebner JL et al. Relationships amongst osteoarthritis biomarkers, dynamic knee joint load, and exercise: results from a randomized controlled pilot study. BMC Musculoskeletal Disorders. 2013; 14:115

198. Hunter DJ, Beavers DP, Eckstein F, Guermazi A, Loeser RF, Nicklas BJ et al. The Intensive Diet and Exercise for Arthritis (IDEA) trial: 18-month radiographic and MRI outcomes. Osteoarthritis and Cartilage. 2015; 23(7):1090-1098

199. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E et al. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. Arthritis and Rheumatism. 2007; 57(7):1211-1219

200. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Williamson E, Jones RH et al. Economic evaluation of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain. Arthritis and Rheumatism (Arthritis Care and Research). 2007; 57(7):1220-1229

201. Husby VS, Helgerud J, Bjorgen S, Husby OS, Benum P, Hoff J. Early maximal strength training is an efficient treatment for patients operated with total hip arthroplasty. Archives of Physical Medicine and Rehabilitation. 2009; 90(10):1658-1667

202. Imoto AM, Pardo JP, Brosseau L, Taki J, Desjardins B, Thevenot O et al. Evidence synthesis of types and intensity of therapeutic land-based exercises to reduce pain in individuals with knee osteoarthritis. Rheumatology International. 2019; 39(7):1159-1179

203. Imoto AM, Peccin MS, Trevisani VF. Quadriceps strengthening exercises are effective in improving pain, function and quality of life in patients with osteoarthritis of the knee. Acta Ortopédica Brasileira. 2012; 20(3):174-179

204. Isaramalai SA, Hounsri K, Kongkamol C, Wattanapisitkul P, Tangadulrat N, Kaewmanee T et al. Integrating participatory ergonomic management in non-weight-bearing exercise and progressive resistance exercise on self-care and functional ability in aged farmers with knee osteoarthritis: a clustered randomized controlled trial. Clinical Interventions in Aging. 2018; 13:101-108

205. Jan MH, Lin CH, Lin YF, Lin JJ, Lin DH. Effects of weight-bearing versus nonweight-bearing exercise on function, walking speed, and position sense in participants with knee osteoarthritis: a randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2009; 90(6):897-904

206. Jan MH, Lin JJ, Liau JJ, Lin YF, Lin DH. Investigation of clinical effects of high- and low-resistance training for patients with knee osteoarthritis: a randomized controlled trial. Physical Therapy. 2008; 88(4):427-436

207. Jan MH, Tang PF, Lin JJ, Tseng SC, Lin YF, Lin DH. Efficacy of a target-matching foot-stepping exercise on proprioception and function in patients with knee osteoarthritis. Journal of Orthopaedic and Sports Physical Therapy. 2008; 38(1):19-25

208. Jansen MJ, Viechtbauer W, Lenssen AF, Hendriks EJ, de Bie RA. Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: a systematic review. Journal of Physiotherapy. 2011; 57(1):11-20

209. Jegu AG, Pereira B, Andant N, Coudeyre E. Effect of eccentric isokinetic strengthening in the rehabilitation of patients with knee osteoarthritis: Isogo, a randomized trial. Trials [Electronic Resource]. 2014; 15:106

210. Jenkinson CM, Doherty M, Avery AJ, Read A, Taylor MA, Sach TH et al. Effects of dietary intervention and quadriceps strengthening exercises on pain and function in overweight people with knee pain: randomised controlled trial. BMJ. 2009; 339:b3170

211. Jeong HS, Lee SC, Song JB, Chang HS, Lee SY. Proprioceptive training and outcomes of patients with knee osteoarthritis: A meta-analysis of randomized controlled trials. Journal of Athletic Training. 2019; 54(4):418-428

212. Jigami H, Sato D, Tsubaki A, Tokunaga Y, Ishikawa T, Dohmae Y et al. Effects of weekly and fortnightly therapeutic exercise on physical function and health-related quality of life in individuals with hip osteoarthritis. Journal of Orthopaedic Science. 2012; 17(6):737-744

213. Jordan J, Holden M, Mason E, Foster N. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD005956. DOI: 10.1002/14651858.CD005956.pub2.

214. Jorge RT, Souza MC, Chiari A, Jones A, Fernandes Ada R, Lombardi Junior I et al. Progressive resistance exercise in women with osteoarthritis of the knee: a randomized controlled trial. Clinical Rehabilitation. 2015; 29(3):234-243

215. Joshi S, Singh SK, Vij JS. Effect of retrowalking, a non-pharmacological treatment on pain, disability, balance and gait in knee osteoarthritis: A randomized controlled trial. Indian journal of public health research and development. 2019; 10(2):214-219

216. Juhakoski R, Tenhonen S, Malmivaara A, Kiviniemi V, Anttonen T, Arokoski JP. A pragmatic randomized controlled study of the effectiveness and cost consequences of exercise therapy in hip osteoarthritis. Clinical Rehabilitation. 2011; 25(4):370-383

217. Juhl C, Christensen R, Roos EM, Zhang W, Lund H. Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. Arthritis & Rheumatology. 2014; 66(3):622-636

218. Kabiri S, Halabchi F, Angoorani H, Yekaninejad S. Comparison of three modes of aerobic exercise combined with resistance training on the pain and function of patients with knee osteoarthritis: A randomized controlled trial. Physical Therapy in Sport 2018; 32:22-28

219. Kamalakannan M, Karthik. Efficacy of proprioception training for tibiofemoral arthritis in relation with pain and functional disability. Research journal of pharmacy and technology. 2019; 12(3):995-998

220. Kan L, Zhang J, Yang Y, Wang P. The effects of yoga on pain, mobility, and quality of life in patients with knee osteoarthritis: A systematic review. Evidence-Based Complementary & Alternative Medicine: eCAM. 2016; 2016:6016532

221. Kang TW, Lee JH, Park DH, Cynn HS. Effects of a finger exercise program on hand function in automobile workers with hand osteoarthritis: A randomized controlled trial. Hand Surgery and Rehabilitation. 2019; 38(1):59-66

222. Karadag S, Tasci S, Dogan N, Demir H, Kilic Z. Application of heat and a home exercise program for pain and function levels in patients with knee osteoarthritis: A randomized controlled trial. International Journal of Nursing Practice. 2019; 25(5):e12772

223. Karatosun V, Unver B, Gocen Z, Sen A, Gunal I. Intra-articular hyaluranic acid compared with progressive knee exercises in osteoarthritis of the knee: a prospective randomized trial with long-term follow-up. Rheumatology International. 2006; 26(4):277-284

224. Karatosun V, Unver B, Ozden A, Ozay Z, Gunal I. Intra-articular hyaluronic acid compared to exercise therapy in osteoarthritis of the ankle. A prospective randomized trial with long-term follow-up. Clinical and Experimental Rheumatology. 2008; 26(2):288-294

225. Kars Fertelli T, Mollaoglu M, Sahin O. Aquatic exercise program for individuals with osteoarthritis: Pain, stiffness, physical function, self-efficacy. Rehabilitation Nursing Journal. 2019; 44(5):290-299

226. Kars FT, Mollaoglu M, Sahin O. Aquatic exercise program for individuals with osteoarthritis: Pain, stiffness, physical function, self-efficacy. Rehabilitation Nursing Journal. 2019; 44(5):E15

227. Kawasaki T, Kurosawa H, Ikeda H, Takazawa Y, Ishijima M, Kubota M et al. Therapeutic home exercise versus intraarticular hyaluronate injection for osteoarthritis of the knee: 6-month prospective randomized open-labeled trial. Journal of Orthopaedic Science. 2009; 14(2):182-191

228. Keefe FJ, Blumenthal J, Baucom D, Affleck G, Waugh R, Caldwell DS et al. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. Pain. 2004; 110(3):539-549

229. Kelley GA, Kelley KS. Exercise reduces depressive symptoms in adults with arthritis: Evidential value. World Journal of Rheumatology. 2016; 6(2):23-29

230. Kelley GA, Kelley KS, Callahan LF. Brief report: Exercise and anxiety in adults with arthritis and other rheumatic diseases: Support for evidential value. BioMed Research International. 2018; 2018(2984671)

231. Keogh JW, Grigg J, Vertullo CJ. Is high-intensity interval cycling feasible and more beneficial than continuous cycling for knee osteoarthritic patients? Results of a randomised control feasibility trial. PeerJ. 2018; 6:e4738

232. Keshtkaran Z, Ghodsbin F, Solouki S, Razeghi M, Zare N. The impact of self care education on quality of life of those clients suffering from osteoarthritis in rehabilitation centers of Shiraz University of Medical Science (Iran). Journal of babol university of medical sciences. 2010; 12(1):8-15

233. Khruakhorn S, Chiwarakranon S. Effects of hydrotherapy and land-based exercise on mobility and quality of life in patients with knee osteoarthritis: a randomized control trial. Journal of Physical Therapy Science. 2021; 33(4):375-383

234. Kigozi J, Nicholls E, Tooth S, Foster NE, Holden MA, Healey EL et al. Cost-utility analysis of interventions to improve effectiveness of exercise therapy for adults with knee osteoarthritis: The BEEP trial. Rheumatology Advances in Practice. 2018; 2(2):1-11

235. Kim IS, Chung SH, Park YJ, Kang HY. The effectiveness of an aquarobic exercise program for patients with osteoarthritis. Applied Nursing Research. 2012; 25(3):181-189

236. Kloek CJJ, Bossen D, Spreeuwenberg PM, Dekker J, de Bakker DH, Veenhof C. Effectiveness of a blended physical therapist intervention in people with hip osteoarthritis, knee osteoarthritis, or both: A cluster-randomized controlled trial. Physical Therapy. 2018; 98(7):560-570

237. Knoop J, Dekker J, van der Leeden M, van der Esch M, Thorstensson CA, Gerritsen M et al. Knee joint stabilization therapy in patients with osteoarthritis of the knee: a randomized, controlled trial. Osteoarthritis and Cartilage. 2013; 21(8):1025-1034

238. Knoop J, Steultjens MP, Roorda LD, Lems WF, van der Esch M, Thorstensson CA et al. Improvement in upper leg muscle strength underlies beneficial effects of exercise therapy in knee osteoarthritis: secondary analysis from a randomised controlled trial. Physiotherapy. 2015; 101(2):171-177

239. Knoop J, van der Leeden M, Roorda LD, Thorstensson CA, van der Esch M, Peter WF et al. Knee joint stabilization therapy in patients with osteoarthritis of the knee and knee instability: subgroup analyses in a randomized, controlled trial. Journal of Rehabilitation Medicine. 2014; 46(7):703-707

240. Konishi I, Tanabe N, Seki N, Suzuki H, Okamura T, Shinoda K et al. Physiotherapy program through home visits for community-dwelling elderly Japanese women with mild knee pain. Tohoku Journal of Experimental Medicine. 2009; 219(2):91-99

241. Kovar PA, Allegrante JP, MacKenzie CR, Peterson MG, Gutin B, Charlson ME. Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. Annals of Internal Medicine. 1992; 116(7):529-534

242. Kraus I, Steinhilber B, Haupt G, Miller R, Martus P, Jansen P. Exercise therapy in hip osteoarthritis--a randomized controlled trial. Deutsches Arzteblatt International. 2014; 111(35-36):592-599

243. Krauss I, Mueller G, Haupt G, Steinhilber B, Janssen P, Jentner N et al. Effectiveness and efficiency of an 11-week exercise intervention for patients with hip or knee osteoarthritis: a protocol for a controlled study in the context of health services research. BMC Public Health. 2016; 16:367

244. Krauss I, Steinhilber B, Haupt G, Miller R, Grau S, Janssen P. Efficacy of conservative treatment regimes for hip osteoarthritis--evaluation of the therapeutic exercise regime "Hip School": a protocol for a randomised, controlled trial. BMC Musculoskeletal Disorders. 2011; 12:270

245. Kreindler H, Lewis CB, Rush S, Schaefer K. Effects of three exercise protocols on strength of persons with osteoarthritis of the knee. Topics geriatric rehabilitation. 1989; 4(3):32-39

246. Kroon FPB, Carmona L, Schoones JW, Kloppenburg M. Efficacy and safety of non-pharmacological, pharmacological and surgical treatment for hand osteoarthritis: a systematic literature review informing the 2018 update of the EULAR recommendations for the management of hand osteoarthritis. RMD Open. 2018; 4(2):e000734

247. Kudo M, Watanabe K, Otsubo H, Kamiya T, Kaneko F, Katayose M et al. Analysis of effectiveness of therapeutic exercise for knee osteoarthritis and possible factors affecting outcome. Journal of Orthopaedic Science. 2013; 18(6):932-939

248. Kumar S, Kumar A, Kumar R. Proprioceptive training as an adjunct in osteoarthritis of knee. Journal of musculoskeletal research. 2013; 16(1):1350002

249. Kuptniratsaikul V, Kittichaikarn C, Suntornpiyapan P, Kovintaset K, Inthibal S. Is four-week underwater treadmill exercise regimen compared to home exercise efficacious for pain relief and functional improvement in obese patients with knee osteoarthritis? A randomized controlled trial. Clinical Rehabilitation. 2019; 33(1):85-93

250. Kuptniratsaikul V, Tosayanonda O, Nilganuwong S, Thamalikitkul V. The efficacy of a muscle exercise program to improve functional performance of the knee in patients

with osteoarthritis. Chotmaihet thangphaet [journal of the medical association of thailand]. 2002; 85(1):33-40

251. Kuru Colak T, Kavlak B, Aydogdu O, Sahin E, Acar G, Demirbuken I et al. The effects of therapeutic exercises on pain, muscle strength, functional capacity, balance and hemodynamic parameters in knee osteoarthritis patients: a randomized controlled study of supervised versus home exercises. Rheumatology International. 2017; 37(3):399-407

252. Lai Z, Zhang Y, Lee S, Wang L. Effects of strength exercise on the knee and ankle proprioception of individuals with knee osteoarthritis. Research in Sports Medicine. 2018; 26(2):138-146

253. Lange AK, Vanwanseele B, Foroughi N, Baker MK, Shnier R, Smith RM et al. Resistive Exercise for Arthritic Cartilage Health (REACH): a randomized double-blind, sham-exercise controlled trial. BMC Geriatrics. 2009; 9:1

254. Lee AC, Harvey WF, Price LL, Han X, Driban JB, Iversen MD et al. Dose-response effects of Tai Chi and physical therapy exercise interventions in symptomatic knee osteoarthritis. Pm & R. 2018; 10(7):712-723

255. Lee B-Y, Shin W-Y, An M-J, Yoon S-R, Choe Y. Effect of dynamic balance exercise in elderly patients with unilateral knee osteoarthritis. Clinical Pain. 2019; 18(1):16-23

256. Lee HJ, Park HJ, Chae Y, Kim SY, Kim SN, Kim ST et al. Tai Chi Qigong for the quality of life of patients with knee osteoarthritis: a pilot, randomized, waiting list controlled trial. Clinical Rehabilitation. 2009; 23(6):504-511

257. Lee HY. Comparison of effects among Tai-Chi exercise, aquatic exercise, and a self-help program for patients with knee osteoarthritis. Taehan kanho hakhoe chi. 2006; 36(3):571-580

258. Lee HY, Lee KJ. Effects of Tai Chi exercise in elderly with knee osteoarthritis. Taehan kanho hakhoe chi. 2008; 38(1):11-18

259. Li LC, Sayre EC, Xie H, Clayton C, Feehan LM. A community-based physical activity counselling program for people with knee osteoarthritis: Feasibility and preliminary efficacy of the track-oa study. JMIR MHealth and UHealth. 2017; 5(6):e86

260. Li X, Liu Y, Wang XQ. Whole-body vibration training for knee osteoarthritis: A systematic review and meta-analysis. Physiotherapy (United Kingdom). 2015; 1:eS868

261. Li Y, Su Y, Chen S, Zhang Y, Zhang Z, Liu C et al. The effects of resistance exercise in patients with knee osteoarthritis: a systematic review and meta-analysis. Clinical Rehabilitation. 2016; 30(10):947-959

262. Liao CD, Liou TH, Huang YY, Huang YC. Effects of balance training on functional outcome after total knee replacement in patients with knee osteoarthritis: a randomized controlled trial. Clinical Rehabilitation. 2013; 27(8):697-709

263. Liebs TR, Herzberg W, Ruther W, Haasters J, Russlies M, Hassenpflug J et al. Multicenter randomized controlled trial comparing early versus late aquatic therapy after total hip or knee arthroplasty. Archives of Physical Medicine and Rehabilitation. 2012; 93(2):192-199

264. Lim BW, Hinman RS, Wrigley TV, Sharma L, Bennell KL. Does knee malalignment mediate the effects of quadriceps strengthening on knee adduction moment, pain, and function in medial knee osteoarthritis? A randomized controlled trial. Arthritis and Rheumatism. 2008; 59(7):943-951

265. Lim JY, Tchai E, Jang SN. Effectiveness of aquatic exercise for obese patients with knee osteoarthritis: a randomized controlled trial. Pm & R. 2010; 2(8):723-731; quiz 793

266. Lin DH, Lin CH, Lin YF, Jan MH. Efficacy of 2 non-weight-bearing interventions, proprioception training versus strength training, for patients with knee osteoarthritis: a randomized clinical trial. Journal of Orthopaedic and Sports Physical Therapy. 2009; 39(6):450-457

267. Lin DH, Lin YF, Chai HM, Han YC, Jan MH. Comparison of proprioceptive functions between computerized proprioception facilitation exercise and closed kinetic chain exercise in patients with knee osteoarthritis. Clinical Rheumatology. 2007; 26(4):520-528

268. Lin SY, Davey RC, Cochrane T. Community rehabilitation for older adults with osteoarthritis of the lower limb: a controlled clinical trial. Clinical Rehabilitation. 2004; 18(1):92-101

269. Lin YT, Lee WC, Hsieh RL. Active video games for knee osteoarthritis improve mobility but not WOMAC score: A randomized controlled trial. Annals of Physical and Rehabilitation Medicine. 2020; 63(6):458-465

270. Liu C, Latham N. Progressive resistance strength training for improving physical function in older adults. Cochrane Database of Systematic Reviews 2009, Issue 3. Art. No.: CD002759. DOI: 10.1002/14651858.CD002759.pub2.

271. Loew L, Brosseau L, Kenny GP, Durand-Bush N, Poitras S, De Angelis G et al. An evidence-based walking program among older people with knee osteoarthritis: the PEP (participant exercise preference) pilot randomized controlled trial. Clinical Rheumatology. 2017; 36(7):1607-1616

272. Lorenc A, Feder G, MacPherson H, Little P, Mercer SW, Sharp D. Scoping review of systematic reviews of complementary medicine for musculoskeletal and mental health conditions. BMJ Open. 2018; 8(10):e020222

273. Lu J, Huang L, Wu X, Fu W, Liu Y. Effect of Tai Ji Quan training on self-reported sleep quality in elderly Chinese women with knee osteoarthritis: a randomized controlled trail. Sleep Medicine. 2017; 33:70-75

274. Lu M, Su Y, Zhang Y, Zhang Z, Wang W, He Z et al. Effectiveness of aquatic exercise for treatment of knee osteoarthritis: Systematic review and meta-analysis. Zeitschrift für Rheumatologie. 2015; 74(6):543-552

275. Lue S, Koppikar S, Shaikh K, Mahendira D, Towheed TE. Systematic review of non-surgical therapies for osteoarthritis of the hand: an update. Osteoarthritis and Cartilage. 2017; 25(9):1379-1389

276. Lun V, Marsh A, Bray R, Lindsay D, Wiley P. Efficacy of hip strengthening exercises compared with leg strengthening exercises on knee pain, function, and quality of life in patients with knee osteoarthritis. Clinical Journal of Sport Medicine. 2015; 25(6):509-517

277. Lund H, Weile U, Christensen R, Rostock B, Downey A, Bartels EM et al. A randomized controlled trial of aquatic and land-based exercise in patients with knee osteoarthritis. Journal of Rehabilitation Medicine. 2008; 40(2):137-144

278. Magni NE, McNair PJ, Rice DA. The effects of resistance training on muscle strength, joint pain, and hand function in individuals with hand osteoarthritis: a systematic review and meta-analysis. Arthritis Research & Therapy. 2017; 19(1):131

279. Mangani I, Cesari M, Kritchevsky SB, Maraldi C, Carter CS, Atkinson HH et al. Physical exercise and comorbidity. Results from the Fitness and Arthritis in Seniors Trial (FAST). Aging-Clinical & Experimental Research. 2006; 18(5):374-380

280. Mangione KK, Axen K, Haas F. Mechanical unweighting effects on treadmill exercise and pain in elderly people with osteoarthritis of the knee. Physical Therapy. 1996; 76(4):387-394

281. Mangione KK, McCully K, Gloviak A, Lefebvre I, Hofmann M, Craik R. The effects of high-intensity and low-intensity cycle ergometry in older adults with knee osteoarthritis. Journals of Gerontology Series A-Biological Sciences & Medical Sciences. 1999; 54(4):M184-190

282. Mat S, Tan MP, Kamaruzzaman SB, Ng CT. Physical therapies for improving balance and reducing falls risk in osteoarthritis of the knee: a systematic review. Age and Ageing. 2015; 44(1):16-24

283. Mattos F, Leite N, Pitta A, Bento PC. Effects of aquatic exercise on muscle strength and functional performance of individuals with osteoarthritis: a systematic review. Revista Brasileira de Reumatologia. 2016; 56(6):530-542

284. Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR, Jr. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. Archives of Physical Medicine and Rehabilitation. 1999; 80(10):1293-1299

285. Mazloum V, Rabiei P, Rahnama N, Sabzehparvar E. The comparison of the effectiveness of conventional therapeutic exercises and Pilates on pain and function in patients with knee osteoarthritis. Complementary Therapies in Clinical Practice. 2018; 31:343-348

286. McCaffrey R, Park J, Newman D. Chair yoga: Feasibility and sustainability study with older community-dwelling adults with osteoarthritis. Holistic Nursing Practice. 2017; 31(3):148-157

287. McCaffrey R, Taylor D, Marker C, Park J. A pilot study of the effects of chair yoga and chair-based exercise on biopsychosocial outcomes in older adults with lower extremity osteoarthritis. Holistic Nursing Practice. 2019; 33(6):321-326

288. McCarthy CJ, Mills PM, Pullen R, Roberts C, Silman A, Oldham JA. Supplementing a home exercise programme with a class-based exercise programme is more effective than home exercise alone in the treatment of knee osteoarthritis. Rheumatology. 2004; 43(7):880-886

289. McIlroy S, Sayliss L, Browning P, Bearne LM. Aquatic therapy for people with persistent knee pain: A feasibility study. Musculoskeletal Care. 2017; 15(4):350-355

290. McKnight PE, Kasle S, Going S, Villanueva I, Cornett M, Farr J et al. A comparison of strength training, self-management, and the combination for early osteoarthritis of the knee. Arthritis Care and Research. 2010; 62(1):45-53

291. McVeigh KH, Kannas SN, Ivy CC, Garner HW, Barnes CS, Heckman MG et al. Dynamic stabilization home exercise program for treatment of thumb carpometacarpal osteoarthritis: A prospective randomized control trial. Journal of Hand Therapy. 2021; epub

292. Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. Arthritis and Rheumatism. 2004; 50(5):1501-1510

293. Messier SP, Mihalko SL, Beavers DP, Nicklas BJ, DeVita P, Carr JJ et al. Effect of high-intensity strength training on knee pain and knee joint compressive forces among adults with knee osteoarthritis: The start randomized clinical trial. JAMA. 2021; 325(7):646-657

294. Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. JAMA. 2013; 310(12):1263-1273

295. Messier SP, Resnik AE, Beavers DP, Mihalko SL, Miller GD, Nicklas BJ et al. Intentional weight loss in overweight and obese patients with knee osteoarthritis: Is more better? Arthritis Care and Research. 2018; 70(11):1569-1575

296. Messier SP, Royer TD, Craven TE, O'Toole ML, Burns R, Ettinger WH, Jr. Long-term exercise and its effect on balance in older, osteoarthritic adults: results from the Fitness, Arthritis, and Seniors Trial (FAST). Journal of the American Geriatrics Society. 2000; 48(2):131-138

297. Messier SP, Thompson CD, Ettinger Jr WH. Effects of long-term aerobic or weight training regimens on gait in an older, osteoarthritic population. Journal of Applied Biomechanics. 1997; 13(2):205-225

298. Mihalko SL, Cox P, Beavers DP, Miller GD, Nicklas BJ, Lyles M et al. Effect of intensive diet and exercise on self-efficacy in overweight and obese adults with knee osteoarthritis: The IDEA randomized clinical trial. Translational Behavioral Medicine. 2019; 9(2):227-235

299. Mikesky AE, Mazzuca SA, Brandt KD, Perkins SM, Damush T, Lane KA. Effects of strength training on the incidence and progression of knee osteoarthritis. Arthritis and Rheumatism. 2006; 55(5):690-699

300. Mikkelsen LR, Mechlenburg I, Soballe K, Jorgensen LB, Mikkelsen S, Bandholm T et al. Effect of early supervised progressive resistance training compared tounsupervised home-based exercise after fast-track total hip replacement applied to patients with preoperative functional limitations. A single-blinded randomised controlled trial. Osteoarthritis and Cartilage. 2014; 22(12):2051-2058

301. Miller GD, Rejeski WJ, Williamson JD, Morgan T, Sevick MA, Loeser RF et al. The Arthritis, Diet and Activity Promotion Trial (ADAPT): design, rationale, and baseline results. Controlled Clinical Trials. 2003; 24(4):462-480

302. Minor MA, Brown JD. Exercise maintenance of persons with arthritis after participation in a class experience. Health Education Quarterly. 1993; 20(1):83-95

303. Minor MA, Hewett JE, Webel RR, Anderson SK, Kay DR. Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. Arthritis and Rheumatism. 1989; 32(11):1396-1405

304. Minshull C, Gleeson N. Considerations of the principles of resistance training in exercise studies for the management of knee osteoarthritis: A systematic review. Archives of Physical Medicine and Rehabilitation. 2017; 98(9):1842-1851

305. Monticone M, Frizziero A, Rovere G, Vittadini F, Uliano D, S LAB et al. Hyaluronic acid intra-articular injection and exercise therapy: effects on pain and disability in subjects affected by lower limb joints osteoarthritis. A systematic review by the Italian Society of Physical and Rehabilitation Medicine (SIMFER). European journal of physical & rehabilitation medicine. 2016; 52(3):389-399

306. Moonaz SH, Bingham CO, 3rd, Wissow L, Bartlett SJ. Yoga in sedentary adults with arthritis: Effects of a randomized controlled pragmatic trial. Journal of Rheumatology. 2015; 42(7):1194-1202

307. Moreira VMPS, da Silva Soares F, Hattori WT, Dionisio VC. A comparison of the efficacy of nonweight-bearing and weight-bearing exercise programmes on function and pain pressure thresholds in knee osteoarthritis: a randomised study. European Journal of Physiotherapy. 2021; 23(3):171-178

308. Moseng T, Dagfinrud H, Smedslund G, Osteras N. The importance of dose in land-based supervised exercise for people with hip osteoarthritis. A systematic review and meta-analysis. Osteoarthritis and Cartilage. 2017; 25(10):1563-1576

309. Munukka M, Waller B, Hakkinen A, Nieminen MT, Lammentausta E, Kujala UM et al. Effects of progressive aquatic resistance training on symptoms and quality of life in women with knee osteoarthritis: A secondary analysis. Scandinavian Journal of Medicine and Science in Sports. 2020; 30(6):1064-1072

310. Munukka M, Waller B, Rantalainen T, Hakkinen A, Nieminen MT, Lammentausta E et al. Efficacy of progressive aquatic resistance training for tibiofemoral cartilage in postmenopausal women with mild knee osteoarthritis: a randomised controlled trial. Osteoarthritis and Cartilage. 2016; 24(10):1708-1717

311. Murphy SL, Lyden AK, Smith DM, Dong Q, Koliba JF. Effects of a tailored activity pacing intervention on pain and fatigue for adults with osteoarthritis. American Journal of Occupational Therapy. 2010; 64(6):869-876

312. Myers A. Compliance with exercise therapy in treating seniors with knee osteoarthritis. Clinical Journal of Sport Medicine. 1998; 8(2):148

313. Na YM, Seok H, Park YG, Seo CH, Seong YJ, Park JR. Effects of therapeutic exercise on patients with osteoarthritis of knee. Journal of korean academy of rehabilitation medicine. 2000; 24(5):966-971

314. Nahayatbin M, Ghasemi M, Rahimi A, Khademi-Kalantari K, Naimi SS, Tabatabaee SM et al. The effects of routine physiotherapy alone and in combination with either Tai Chi or closed kinetic chain exercises on knee osteoarthritis: a comparative clinical trial study. Iranian red crescent medical journal. 2018; 20(4)

315. Nambi G, Abdelbasset WK, Alrawail SM, Elnegamy TE, Abodonya AM, Saleh AK. Effects of isokinetic knee muscle training on bone morphogenetic proteins and inflammatory biomarkers in post-traumatic osteoarthritis after anterior cruciate ligament injury: A randomized trial. Journal of Rehabilitation Medicine. 2020; epub

316. Nathani S, Tank KD. Effect of pnf stretching on proprioception and physical function in individual with knee osteoarthritis: An experimental study. Indian journal of public health research and development. 2020; 11(7):779-784

317. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated October 2020]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

318. Neelapala YVR, Bhagat M, Shah P. Hip muscle strengthening for knee osteoarthritis: A systematic review of literature. Journal of Geriatric Physical Therapy. 2020; 43(2):89-98

319. Neelapala YVR, Nayak S, Sivalanka S, Cornelio R, Prajapati M. Influence of isometric exercise on pressure pain sensitivity in knee osteoarthritis. Journal of Pain Management. 2018; 11(4):361-367

320. Nejati P, Farzinmehr A, Moradi-Lakeh M. The effect of exercise therapy on knee osteoarthritis: a randomized clinical trial. Medical Journal of the Islamic Republic of Iran. 2015; 29:186

321. Nelligan R, Hinman R, Kasza J, Crofts S, Bennel K. Effects of self-directed web-based strengthening exercise and physical activity supported by automated sms on pain and function in people with knee osteoarthritis: a randomised controlled trial. Internal Medicine Journal. 2021; 51(Suppl 2):9-10

322. Nery M, Natour J, Jennings F, Fernandes A, Souza MC, Jones A. Effects of a progressive resistance exercise program in patients with hand osteoarthritis: A randomized, controlled trial with a blinded assessor. Clinical Rehabilitation. 2021; 35(12):1757-1767

323. Nery MV, Jones A, Jennings F, Souza MC, Natour J. Effectiveness of a progressive resistance strength program on hand osteoarthritis: a randomised controlled trial. Annals of the Rheumatic Diseases. 2016; 75:60

324. Nery MV, Martinez A, Jennings F, Souza M, Natour J. Effectiveness of a progressive resistence strength programme on hand osteoarthritis: a randomized crontrolled trial. Arthritis and rheumatology. 2015; 67

325. Ng NT, Heesch KC, Brown WJ. Efficacy of a progressive walking program and glucosamine sulphate supplementation on osteoarthritic symptoms of the hip and knee: a feasibility trial. Arthritis Research & Therapy. 2010; 12(1):R25

326. Ni GX, Song L, Yu B, Huang CH, Lin JH. Tai chi improves physical function in older Chinese women with knee osteoarthritis. JCR: Journal of Clinical Rheumatology. 2010; 16(2):64-67

327. Nicklas BJ, Ambrosius W, Messier SP, Miller GD, Penninx BW, Loeser RF et al. Diet-induced weight loss, exercise, and chronic inflammation in older, obese adults: a randomized controlled clinical trial. American Journal of Clinical Nutrition. 2004; 79(4):544-551

328. O'Reilly SC, Muir KR, Doherty M. Effectiveness of home exercise on pain and disability from osteoarthritis of the knee: a randomised controlled trial. Annals of the Rheumatic Diseases. 1999; 58(1):15-19

329. Oiestad BE, Osteras N, Frobell R, Grotle M, Brogger H, Risberg MA. Efficacy of strength and aerobic exercise on patient-reported outcomes and structural changes in patients with knee osteoarthritis: study protocol for a randomized controlled trial. BMC Musculoskeletal Disorders. 2013; 14:266

330. Ojoawo AO, Olaogun MO, Hassan MA. Comparative effects of proprioceptive and isometric exercises on pain intensity and difficulty in patients with knee osteoarthritis: A randomised control study. Technology and Health Care. 2016; 24(6):853-863

331. Olagbegi OM, Adegoke BOA, Odole A. Effectiveness of combined chain exercises on pain and function in patients with knee osteoarthritis. Bangladesh journal of medical science. 2016; 15(2):178-188

332. Oliveira AM, Peccin MS, Silva KN, Teixeira LE, Trevisani VF. Impact of exercise on the functional capacity and pain of patients with knee osteoarthritis: a randomized clinical trial. Revista Brasileira de Reumatologia. 2012; 52(6):876-882

333. Oppong R, Jowett S, Nicholls E, Whitehurst DG, Hill S, Hammond A et al. Joint protection and hand exercises for hand osteoarthritis: an economic evaluation comparing methods for the analysis of factorial trials. Rheumatology. 2014:epub

334. Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2021. Available from: http://www.oecd.org/std/ppp Last accessed: 10/02/2022.

335. Osborne RH, Buchbinder R, Ackerman IN. Can a disease-specific education program augment self-management skills and improve Health-Related Quality of Life in people with hip or knee osteoarthritis? BMC Musculoskeletal Disorders. 2006; 7:90

336. Osteras H, Paulsberg F, Olsen SE, Osteras B, Torstensen TA. Effects of medical exercise therapy in patients with hip osteoarthritis: A randomized controlled trial with six months follow-up. A pilot study. Journal of Bodywork and Movement Therapies. 2017; 21(2):284-289

337. Osteras N, Hagen KB, Grotle M, Sand-Svartrud AL, Mowinckel P, Kjeken I. Limited effects of exercises in people with hand osteoarthritis: results from a randomized controlled trial. Osteoarthritis and Cartilage. 2014; 22(9):1224-1233

338. Østerås N, Kjeken I, Smedslund G, Moe R, Slatkowsky-Christensen B, Uhlig T et al. Exercise for hand osteoarthritis. Cochrane Database of Systematic Reviews 2017, Issue 1. Art. No.: CD010388. DOI: 10.1002/14651858.CD010388.pub2.

339. Osteras N, Kjeken I, Smedslund G, Moe RH, Slatkowsky-Christensen B, Uhlig T et al. Exercise for Hand Osteoarthritis: A Cochrane Systematic Review. Journal of Rheumatology. 2017; 44(12):1850-1858

340. Osugi T, Iwamoto J, Yamazaki M, Takakuwa M. Effect of a combination of whole body vibration exercise and squat training on body balance, muscle power, and walking ability in the elderly. Therapeutics and Clinical Risk Management. 2014; 10:131-138

341. Ozdincler AR, Yeldan I, Kinali P. The effects of closed kinetic chain exercise on pain and functional performance of patients with knee osteoarthritis. Pain clinic. 2005; 17(1):107-115

342. Ozturk O, Bombaci H, Kececi T, Algun ZC. Effects of additional action observation to an exercise program in patients with chronic pain due to knee osteoarthritis: A randomized-controlled trial. Musculoskeletal Science & Practice. 2021; 52:102334

343. Park J, McCaffrey R, Dunn D, Goodman R. Managing osteoarthritis: comparisons of chair yoga, Reiki, and education (pilot study). Holistic Nursing Practice. 2011; 25(6):316-326

344. Park J, McCaffrey R, Newman D, Cheung C, Hagen D. The effect of Sit 'n' Fit Chair Yoga among community-dwelling older adults with osteoarthritis. Holistic Nursing Practice. 2014; 28(4):247-257

345. Park J, McCaffrey R, Newman D, Liehr P, Ouslander JG. A pilot randomized controlled trial of the effects of chair yoga on pain and physical function among community-dwelling older adults with lower extremity osteoarthritis. Journal of the American Geriatrics Society. 2017; 65(3):592-597

346. Park J, Newman D, McCaffrey R, Garrido JJ, Riccio ML, Liehr P. The effect of chair yoga on biopsychosocial changes in english- and spanish-speaking community-dwelling older adults with lower-extremity osteoarthritis. Journal of Gerontological Social Work. 2016; 59(7-8):604-626

347. Park S, Min S, Park SH, Yoo J, Jee YS. Influence of isometric exercise combined with electromyostimulation on inflammatory cytokine levels, muscle strength, and knee joint function in elderly women with early knee osteoarthritis. Frontiers in Physiology. 2021; 12:688260

348. Patrick DL, Ramsey SD, Spencer AC, Kinne S, Belza B, Topolski TD. Economic evaluation of aquatic exercise for persons with osteoarthritis. Medical Care. 2001; 39(5):413-424

349. Pazit L, Jeremy D, Nancy B, Michael B, George E, Hill KD. Safety and feasibility of high speed resistance training with and without balance exercises for knee osteoarthritis: A pilot randomised controlled trial. Physical Therapy in Sport 2018; 34:154-163

350. Peloquin L, Bravo G, Gauthier P, Lacombe G, Billiard JS. Effects of a cross-training exercise program in persons with osteoarthritis of the knee a randomized controlled trial. JCR: Journal of Clinical Rheumatology. 1999; 5(3):126-136

351. Penninx BW, Messier SP, Rejeski WJ, Williamson JD, DiBari M, Cavazzini C et al. Physical exercise and the prevention of disability in activities of daily living in older persons with osteoarthritis. Archives of Internal Medicine. 2001; 161(19):2309-2316

352. Perez-Huerta BD, Diaz-Pulido B, Pecos-Martin D, Beckwee D, Lluch-Girbes E, Fernandez-Matias R et al. Effectiveness of a program combining strengthening, stretching, and aerobic training exercises in a standing versus a sitting position in overweight subjects with knee osteoarthritis: A randomized controlled trial. Journal of Clinical Medicine. 2020; 9(12):20

353. Perez-Marmol JM, Garcia-Rios MC, Ortega-Valdivieso MA, Cano-Deltell EE, Peralta-Ramirez MI, Ickmans K et al. Effectiveness of a fine motor skills rehabilitation program on upper limb disability, manual dexterity, pinch strength, range of fingers motion, performance in activities of daily living, functional independency, and general self-efficacy in hand osteoarthritis: A randomized clinical trial. Journal of Hand Therapy. 2017; 30(3):262-273

354. Petersen SG, Miller BF, Hansen M, Kjaer M, Holm L. Exercise and NSAIDs: effect on muscle protein synthesis in patients with knee osteoarthritis. Medicine and Science in Sports and Exercise. 2011; 43(3):425-431

355. Petrella RJ, Bartha C. Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial. Journal of Rheumatology. 2000; 27(9):2215-2221

356. Pignato M, Arbeeva L, Schwartz TA, Callahan LF, Cooke J, Golightly YM et al. Level of participation in physical therapy or an internet-based exercise training program: associations with outcomes for patients with knee osteoarthritis. BMC Musculoskeletal Disorders. 2018; 19(1):238

357. Pinto D, Robertson MC, Abbott JH, Hansen P, Campbell AJ. Manual therapy, exercise therapy, or both, in addition to usual care, for osteoarthritis of the hip or knee. 2: economic evaluation alongside a randomized controlled trial. Osteoarthritis and Cartilage. 2013; 21(10):1504-1513

358. Pisters MF, Veenhof C, de Bakker DH, Schellevis FG, Dekker J. Behavioural graded activity results in better exercise adherence and more physical activity than usual care in people with osteoarthritis: a cluster-randomised trial. Journal of Physiotherapy. 2010; 56(1):41-47

359. Pisters MF, Veenhof C, Schellevis FG, De Bakker DH, Dekker J. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized controlled trial comparing two different physical therapy interventions. Osteoarthritis and Cartilage. 2010; 18(8):1019-1026

360. Piyakhachornrot N, Aree-Ue S, Putwatana P, Kawinwonggowit V. Impact of an integrated health education and exercise program in middle-aged Thai adults with osteoarthritis of the knee. Orthopaedic nursing / National Association of Orthopaedic Nurses. 2011; 30(2):134-142

361. Praharsitha R, Sowmya MV, Kalaiselvan A, Dev R, Manoj Abraham M. Influence of medial and lateral hamstring strengthening exercises in women with osteoarthritis knee: A randomized trail. Research journal of pharmacy and technology. 2019; 12(4):1721-1725

362. Qi M, Moyle W, Jones C, Weeks B. Tai chi combined with resistance training for adults aged 50 years and older: A systematic review. Journal of Geriatric Physical Therapy. 2020; 43(1):32-41

363. Quicke JG, Foster NE, Thomas MJ, Holden MA. Is long-term physical activity safe for older adults with knee pain?: a systematic review. Osteoarthritis and Cartilage. 2015; 23(9):1445-1456

364. Quicke JG, Runhaar J, van der Windt DA, Healey EL, Foster NE, Holden MA. Moderators of the effects of therapeutic exercise for people with knee and hip osteoarthritis: A systematic review of sub-group analyses from randomised controlled trials. Osteoarthritis and Cartilage Open. 2020; 2(4):100113

365. Quilty B, Tucker M, Campbell R, Dieppe P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. Journal of Rheumatology. 2003; 30(6):1311-1317

366. Raj NB, Saha S, Razali H, Othman NYH, Mahadeva Rao US. Does proprioception of knee improve after various forms of training in osteoarthritis of knee? Research journal of pharmacy and technology. 2019; 12(9):4379-4386

367. Rao A, Evans MF. Does a structured exercise program benefit elderly people with knee osteoarthritis? Canadian family physician médecin de famille canadien. 1998; 44:283-284

368. Rashid SA, Moiz JA, Sharma S, Raza S, Rashid SM, Hussain ME. Comparisons of neuromuscular training versus quadriceps training on gait and womac index in patients with knee osteoarthritis and varus malalignment. Journal of Chiropractic Medicine. 2019; 18(1):1-8

369. Ravaud P, Giraudeau B, Logeart I, Larguier JS, Rolland D, Treves R et al. Management of osteoarthritis (OA) with an unsupervised home based exercise programme and/or patient administered assessment tools. A cluster randomised controlled trial with a 2x2 factorial design. Annals of the Rheumatic Diseases. 2004; 63(6):703-708

370. Regnaux J, Lefevre-Colau M, Trinquart L, Nguyen C, Boutron I, Brosseau L et al. High-intensity versus low-intensity physical activity or exercise in people with hip or knee osteoarthritis. Cochrane Database of Systematic Reviews 2015, Issue 10. Art. No.: CD010203. DOI: 10.1002/14651858.CD010203.pub2.

371. Rejeski WJ, Brawley LR, Ettinger W, Morgan T, Thompson C. Compliance to exercise therapy in older participants with knee osteoarthritis: implications for treating disability. Medicine and Science in Sports and Exercise. 1997; 29(8):977-985

372. Rejeski WJ, Ettinger WH, Jr., Martin K, Morgan T. Treating disability in knee osteoarthritis with exercise therapy: a central role for self-efficacy and pain. Arthritis Care and Research. 1998; 11(2):94-101

373. Rejeski WJ, Focht BC, Messier SP, Morgan T, Pahor M, Penninx B. Obese, older adults with knee osteoarthritis: weight loss, exercise, and quality of life. Health Psychology. 2002; 21(5):419-426

374. Rewald S, Lenssen AFT, Emans PJ, de Bie RA, van Breukelen G, Mesters I. Aquatic cycling improves knee pain and physical functioning in patients with knee osteoarthritis: A randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2020; 101(8):1288-1295

375. Rewald S, Mesters I, Lenssen AF, Emans PJ, Wijnen W, de Bie RA. Effect of aqua-cycling on pain and physical functioning compared with usual care in patients with knee osteoarthritis: study protocol of a randomised controlled trial. BMC Musculoskeletal Disorders. 2016; 17:88

376. Rezasoltani Z, Sanati E, Kazempour Mofrad R, Azizi S, Dadarkhah A, Najafi S. Randomized controlled trial of aquatic cycling for treatment of knee osteoarthritis in elderly people. Topics in Geriatric Rehabilitation. 2020; 36(2):103-109

377. Robbins SR, Alfredo PP, Junior WS, Marques AP. Low-level laser therapy and static stretching exercises for patients with knee osteoarthritis: A randomised controlled trial. Clinical Rehabilitation. 2022; 36(2):204-213

378. Rodriguez-Merchan EC. Conservative treatment of acute knee osteoarthritis: A review of the Cochrane Library. Journal of Acute Disease. 2016; 5(3):190-193

379. Rogers MW, Tamulevicius N, Coetsee MF, Curry BF, Semple SJ. Knee osteoarthritis and the efficacy of kinesthesia, balance & agility exercise training: A pilot study. International Journal of Exercise Science. 2011; 4(2):124-132

380. Rogers MW, Tamulevicius N, Semple SJ, Krkeljas Z. Efficacy of home-based kinesthesia, balance & agility exercise training among persons with symptomatic knee osteoarthritis. Journal of Sports Science & Medicine. 2012; 11(4):751-758

381. Rogers MW, Wilder FV. Exercise and hand osteoarthritis symptomatology: a controlled crossover trial. Journal of Hand Therapy. 2009; 22(1):10-17; discussion 19-20; quiz 18

382. Rogind H, Bibow-Nielsen B, Jensen B, Moller HC, Frimodt-Moller H, Bliddal H. The effects of a physical training program on patients with osteoarthritis of the knees. Archives of Physical Medicine and Rehabilitation. 1998; 79(11):1421-1427

383. Roper JA, Bressel E, Tillman MD. Acute aquatic treadmill exercise improves gait and pain in people with knee osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2013; 94(3):419-425

384. Rosedale R, Rastogi R, May S, Chesworth BM, Filice F, Willis S et al. Efficacy of exercise intervention as determined by the mckenzie system of mechanical diagnosis and therapy for knee osteoarthritis: A randomized controlled trial. Journal of Orthopaedic and Sports Physical Therapy. 2014; 44(3):173-181, A171-176

385. Runhaar J, Deroisy R, van Middelkoop M, Barretta F, Barbetta B, Oei EH et al. The role of diet and exercise and of glucosamine sulfate in the prevention of knee osteoarthritis: Further results from the PRevention of knee Osteoarthritis in Overweight Females (PROOF) study. Seminars in Arthritis and Rheumatism. 2016; 45(Suppl 4):S42-48

386. Saccomanno MF, Donati F, Careri S, Bartoli M, Severini G, Milano G. Efficacy of intra-articular hyaluronic acid injections and exercise-based rehabilitation programme, administered as isolated or integrated therapeutic regimens for the treatment of knee

osteoarthritis. Knee Surgery, Sports Traumatology, Arthroscopy. 2016; 24(5):1686-1694

387. Salacinski AJ, Krohn K, Lewis SF, Holland ML, Ireland K, Marchetti G. The effects of group cycling on gait and pain-related disability in individuals with mild-to-moderate knee osteoarthritis: a randomized controlled trial. Journal of Orthopaedic and Sports Physical Therapy. 2012; 42(12):985-995

388. Salli A, Sahin N, Baskent A, Ugurlu H. The effect of two exercise programs on various functional outcome measures in patients with osteoarthritis of the knee: a randomized controlled clinical trial. Isokinetics and Exercise Science. 2010; 18(4):201-209

389. Sampath KK, Mani R, Miyamori T, Tumilty S. The effects of manual therapy or exercise therapy or both in people with hip osteoarthritis: a systematic review and meta-analysis. Clinical Rehabilitation. 2016; 30(12):1141-1155

390. Samut G, Dincer F, Ozdemir O. The effect of isokinetic and aerobic exercises on serum interleukin-6 and tumor necrosis factor alpha levels, pain, and functional activity in patients with knee osteoarthritis. Modern Rheumatology. 2015; 25(6):919-924

391. Sashika H, Matsuba Y, Watanabe Y. Home program of physical therapy: effect on disabilities of patients with total hip arthroplasty. Archives of Physical Medicine and Rehabilitation. 1996; 77(3):273-277

392. Saw MM, Kruger-Jakins T, Edries N, Parker R. Significant improvements in pain after a six-week physiotherapist-led exercise and education intervention, in patients with osteoarthritis awaiting arthroplasty, in South Africa: a randomised controlled trial. BMC Musculoskeletal Disorders. 2016; 17:236

393. Sayers SP, Gibson K, Cook CR. Effect of high-speed power training on muscle performance, function, and pain in older adults with knee osteoarthritis: a pilot investigation. Arthritis Care and Research. 2012; 64(1):46-53

394. Schencking M, Wilm S, Redaelli M. A comparison of Kneipp hydrotherapy with conventional physiotherapy in the treatment of osteoarthritis: a pilot trial. The Journal of Integrative Medicine. 2013; 11(1):17-25

395. Schepens SL, Braun ME, Murphy SL. Effect of tailored activity pacing on self-perceived joint stiffness in adults with knee or hip osteoarthritis. American Journal of Occupational Therapy. 2012; 66(3):363-367

396. Schilke JM, Johnson GO, Housh TJ, O'Dell JR. Effects of muscle-strength training on the functional status of patients with osteoarthritis of the knee joint. Nursing Research. 1996; 45(2):68-72

397. Schlenk EA, Fitzgerald GK, Rogers JC, Kwoh CK, Sereika SM. Promoting physical activity in older adults with knee osteoarthritis and hypertension: A randomized controlled trial. Journal of Aging & Physical Activity. 2020; 29(2):207-218

398. Schlenk EA, Lias JL, Sereika SM, Dunbar-Jacob J, Kwoh CK. Improving physical activity and function in overweight and obese older adults with osteoarthritis of the knee: a feasibility study. Rehabilitation Nursing Journal. 2011; 36(1):32-42

399. Schmid A, McAlindon T, Schmid CH, Wang C. The influence of tai chi exercise on proprioception in patients with knee osteoarthritis: Results from a pilot randomized controlled trial. International Journal of Integrative Medicine. 2013; 1

400. Sedaghatnezhad P, Shams M, Karimi N, Rahnama L. Uphill treadmill walking plus physical therapy versus physical therapy alone in the management of individuals with

knee osteoarthritis: a randomized clinical trial. Disability and Rehabilitation. 2021; 43(18):2541-2549

401. Segal NA, Glass NA, Teran-Yengle P, Singh B, Wallace RB, Yack HJ. Intensive gait training for older adults with symptomatic knee osteoarthritis. American Journal of Physical Medicine and Rehabilitation. 2015; 94(10 Suppl 1):848-858

402. Seidler A, Luben L, Hegewald J, Bolm-Audorff U, Bergmann A, Liebers F et al. Dose-response relationship between cumulative physical workload and osteoarthritis of the hip - a meta-analysis applying an external reference population for exposure assignment. BMC Musculoskeletal Disorders. 2018; 19(1):182

403. Sekir U, Gur H. A multi-station proprioceptive exercise program in patients with bilateral knee osteoarthrosis: functional capacity, pain and sensoriomotor function. A randomized controlled trial. Journal of Sports Science & Medicine. 2005; 4(4):590-603

404. Sevick MA, Bradham DD, Muender M, Chen GJ, Enarson C, Dailey M et al. Cost-effectiveness of aerobic and resistance exercise in seniors with knee osteoarthritis. Medicine and Science in Sports and Exercise. 2000; 32(9):1534-1540

405. Shahine NF, El Ashri NI, Senna MK, Abd Elhameed SH. Effect of a pedometer based aerobic walking program on pain and function among elderly patients with knee osteoarthritis. European Journal of Molecular and Clinical Medicine. 2020; 7(9):790-799

406. Sharma M, Singh A, Dhillon MS, Kaur S. Comparative impact of nonpharmacological interventions on pain of knee osteoarthritis patients reporting at a tertiary care institution: A randomized controlled trial. Indian Journal of Palliative Care. 2018; 24(4):478-485

407. Shen P, Li L, Song Q, Sun W, Zhang C, Fong D et al. Proprioceptive Neuromuscular Facilitation Improves Symptoms among Older Adults with Knee Osteoarthritis during Stair Ascending-A Randomized Controlled Trial. American Journal of Physical Medicine and Rehabilitation. 2021; 20

408. Silva LE, Valim V, Pessanha AP, Oliveira LM, Myamoto S, Jones A et al. Hydrotherapy versus conventional land-based exercise for the management of patients with osteoarthritis of the knee: a randomized clinical trial. Physical Therapy. 2008; 88(1):12-21

409. Simao AP, Avelar NC, Tossige-Gomes R, Neves CD, Mendonca VA, Miranda AS et al. Functional performance and inflammatory cytokines after squat exercises and whole-body vibration in elderly individuals with knee osteoarthritis. Archives of Physical Medicine and Rehabilitation. 2012; 93(10):1692-1700

410. Sled EA, Khoja L, Deluzio KJ, Olney SJ, Culham EG. Effect of a home program of hip abductor exercises on knee joint loading, strength, function, and pain in people with knee osteoarthritis: a clinical trial. Physical Therapy. 2010; 90(6):895-904

411. Smith C, Kumar S, Pelling N. The effectiveness of self-management educational interventions for osteoarthritis of the knee. JBI Library of Systematic Reviewis. 2009; 7(25):1091-1118

412. Smith TO, King JJ, Hing CB. The effectiveness of proprioceptive-based exercise for osteoarthritis of the knee: a systematic review and meta-analysis. Rheumatology International. 2012; 32(11):3339-3351

413. Somers TJ, Blumenthal JA, Guilak F, Kraus VB, Schmitt DO, Babyak MA et al. Pain coping skills training and lifestyle behavioral weight management in patients with knee osteoarthritis: a randomized controlled study. Pain. 2012; 153(6):1199-1209

414. Song R, Lee EO, Lam P, Bae SC. Effects of a Sun-style Tai Chi exercise on arthritic symptoms, motivation and the performance of health behaviors in women with osteoarthritis. Daehan Ganho Haghoeji. 2007; 37(2):249-256

415. Song R, Lee EO, Lam P, Bae SC. Effects of tai chi exercise on pain, balance, muscle strength, and perceived difficulties in physical functioning in older women with osteoarthritis: a randomized clinical trial. Journal of Rheumatology. 2003; 30(9):2039-2044

416. Song R, Roberts BL, Lee EO, Lam P, Bae SC. A randomized study of the effects of t'ai chi on muscle strength, bone mineral density, and fear of falling in women with osteoarthritis. Journal of Alternative and Complementary Medicine. 2010; 16(3):227-233

417. Sorour AS, Ayoub AS, Abd El Aziz EM. Effectiveness of acupressure versus isometric exercise on pain, stiffness, and physical function in knee osteoarthritis female patients. Journal of Advanced Research. 2014; 5(2):193-200

418. Srikesavan CS, Shay B, Szturm T. Task-oriented training with computer games for people with rheumatoid arthritis or hand osteoarthritis: A feasibility randomized controlled trial. Games for Health Journal. 2016; 5(5):295-303

419. Stamm TA, Machold KP, Smolen JS, Fischer S, Redlich K, Graninger W et al. Joint protection and home hand exercises improve hand function in patients with hand osteoarthritis: a randomized controlled trial. Arthritis and Rheumatism. 2002; 47(1):44-49

420. Steinhilber B, Haupt G, Miller R, Janssen P, Krauss I. Exercise therapy in patients with hip osteoarthritis: Effect on hip muscle strength and safety aspects of exercise-results of a randomized controlled trial. Modern Rheumatology. 2017; 27(3):493-502

421. Stener-Victorin E, Kruse-Smidje C, Jung K. Comparison between electro-acupuncture and hydrotherapy, both in combination with patient education and patient education alone, on the symptomatic treatment of osteoarthritis of the hip. Clinical Journal of Pain. 2004; 20(3):179-185

422. Stensrud S, Risberg MA, Roos EM. Effect of exercise therapy compared with arthroscopic surgery on knee muscle strength and functional performance in middle-aged patients with degenerative meniscus tears: a 3-mo follow-up of a randomized controlled trial. American Journal of Physical Medicine and Rehabilitation. 2015; 94(6):460-473

423. Stoffer-Marx MA, Klinger M, Luschin S, Meriaux-Kratochvila S, Zettel-Tomenendal M, Nell-Duxneuner V et al. Functional consultation and exercises improve grip strength in osteoarthritis of the hand - a randomised controlled trial. Arthritis Research & Therapy. 2018; 20(1):253

424. Sullivan T, Allegrante JP, Peterson MG, Kovar PA, MacKenzie CR. One-year followup of patients with osteoarthritis of the knee who participated in a program of supervised fitness walking and supportive patient education. Arthritis Care and Research. 1998; 11(4):228-233

425. Suzuki Y, Iijima H, Tashiro Y, Kajiwara Y, Zeidan H, Shimoura K et al. Home exercise therapy to improve muscle strength and joint flexibility effectively treats pre-

radiographic knee OA in community-dwelling elderly: a randomized controlled trial. Clinical Rheumatology. 2019; 38(1):133-141

426. Svege I, Fernandes L, Nordsletten L, Holm I, Risberg MA. Long-term effect of exercise therapy and patient education on impairments and activity limitations in people with hip osteoarthritis: Secondary outcome analysis of a randomized clinical trial. Physical Therapy. 2016; 96(6):818-827

427. Svege I, Nordsletten L, Fernandes L, Risberg MA. Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial. Annals of the Rheumatic Diseases. 2015; 74(1):164-169

428. Taglietti M, Facci LM, Trelha CS, de Melo FC, da Silva DW, Sawczuk G et al. Effectiveness of aquatic exercises compared to patient-education on health status in individuals with knee osteoarthritis: a randomized controlled trial. Clinical Rehabilitation. 2018; 32(6):766-776

429. Tak E, Staats P, Van Hespen A, Hopman-Rock M. The effects of an exercise program for older adults with osteoarthritis of the hip. Journal of Rheumatology. 2005; 32(6):1106-1113

430. Takacs J, Krowchuk NM, Garland SJ, Carpenter MG, Hunt MA. Dynamic balance training improves physical function in individuals with knee osteoarthritis: A pilot randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2017; 98(8):1586-1593

431. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. Journal of the American Geriatrics Society. 2003; 51(3):387-392

432. Tamin T, Loekito N. The efficacy of land versus water exercise program on body composition in obese patients with knee osteoarthritis. Annals of Physical and Rehabilitation Medicine. 2018; 61(Supplement):e9

433. Tamin TZ, Loekito N. Aquatic versus land-based exercise for cardiorespiratory endurance and quality of life in obese patients with knee osteoarthritis: a randomized controlled trial. Medical journal of indonesia. 2018; 27(4):284-292

434. Tan SS, Teirlinck CH, Dekker J, Goossens LM, Bohnen AM, Verhaar JA et al. Cost-utility of exercise therapy in patients with hip osteoarthritis in primary care. Osteoarthritis and Cartilage. 2016; 24(4):581-588

435. Tanaka R, Ozawa J, Kito N, Moriyama H. Does exercise therapy improve the health-related quality of life of people with knee osteoarthritis? A systematic review and meta-analysis of randomized controlled trials. Journal of Physical Therapy Science. 2015; 27(10):3309-3314

436. Tanaka R, Ozawa J, Kito N, Moriyama H. Effect of the frequency and duration of land-based therapeutic exercise on pain relief for people with knee osteoarthritis: A systematic review and meta-analysis of randomized controlled trials. Journal of Physical Therapy Science. 2014; 26(7):969-975

437. Tanaka R, Ozawa J, Kito N, Moriyama H. Effects of exercise therapy on walking ability in individuals with knee osteoarthritis: a systematic review and meta-analysis of randomised controlled trials. Clinical Rehabilitation. 2016; 30(1):36-52

438. Tanaka R, Ozawa J, Kito N, Moriyama H. Efficacy of strengthening or aerobic exercise on pain relief in people with knee osteoarthritis: a systematic review and

meta-analysis of randomized controlled trials. Clinical Rehabilitation. 2013; 27(12):1059-1071

439. Teirlinck CH, Luijsterburg PA, Dekker J, Bohnen AM, Verhaar JA, Koopmanschap MA et al. Effectiveness of exercise therapy added to general practitioner care in patients with hip osteoarthritis: a pragmatic randomized controlled trial. Osteoarthritis and Cartilage. 2016; 24(1):82-90

440. Teixeira PE, Piva SR, Fitzgerald GK. Effects of impairment-based exercise on performance of specific self-reported functional tasks in individuals with knee osteoarthritis. Physical Therapy. 2011; 91(12):1752-1765

441. Thomas KS, Miller P, Doherty M, Muir KR, Jones AC, O'Reilly SC. Cost effectiveness of a two-year home exercise program for the treatment of knee pain. Arthritis and Rheumatism (Arthritis Care and Research). 2005; 53(3):388-394

442. Thomas KS, Muir KR, Doherty M, Jones AC, O'Reilly SC, Bassey EJ. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. BMJ. 2002; 325(7367):752

443. Thompson AR, Christopherson Z, Marshall LM, Carlson HL, Carlson NL. A pilot randomized controlled trial for aerobic and strengthening exercises on physical function and pain for hip osteoarthritis. Pm & R. 2020; 12(3):229-237

444. Thorstensson CA, Roos EM, Petersson IF, Ekdahl C. Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis: a randomized controlled trial [ISRCTN20244858]. BMC Musculoskeletal Disorders. 2005; 6:27

445. Topp R, Woolley S, Hornyak J, 3rd, Khuder S, Kahaleh B. The effect of dynamic versus isometric resistance training on pain and functioning among adults with osteoarthritis of the knee. Archives of Physical Medicine and Rehabilitation. 2002; 83(9):1187-1195

446. Tossige-Gomes R, Avelar NCP, Simao AP, Neves CDC, Brito-Melo GEA, Coimbra CC et al. Whole-body vibration decreases the proliferative response of TCD4+ cells in elderly individuals with knee osteoarthritis. Brazilian Journal of Medical and Biological Research. 2012; 45(12):1262-1268

447. Trans T, Aaboe J, Henriksen M, Christensen R, Bliddal H, Lund H. Effect of whole body vibration exercise on muscle strength and proprioception in females with knee osteoarthritis. Knee. 2009; 16(4):256-261

448. Tsai PF, Chang JY, Beck C, Kuo YF, Keefe FJ. A pilot cluster-randomized trial of a 20-week Tai Chi program in elders with cognitive impairment and osteoarthritic knee: effects on pain and other health outcomes. Journal of Pain and Symptom Management. 2013; 45(4):660-669

449. Tsai PF, Chang JY, Beck C, Kuo YF, Keefe FJ, Rosengren K. A supplemental report to a randomized cluster trial of a 20-week Sun-style Tai Chi for osteoarthritic knee pain in elders with cognitive impairment. Complementary Therapies in Medicine. 2015; 23(4):570-576

450. Tsauo JY, Cheng PF, Yang RS. The effects of sensorimotor training on knee proprioception and function for patients with knee osteoarthritis: a preliminary report. Clinical Rehabilitation. 2008; 22(5):448-457

451. Tunay VB, Baltaci G, Atay AO. Hospital-based versus home-based proprioceptive and strengthening exercise programs in knee osteoarthritis. Acta Orthopaedica et Traumatologica Turcica. 2010; 44(4):270-277

452. Tuzun EH, Aytar A, Eker L, Daskapan A. Effectiveness of two different physical therapy programmes in the treatment of knee osteoarthritis. The pain clinic. 2004; 16(4):379-387

453. Ulstein S, Aroen A, Rotterud JH, Loken S, Engebretsen L, Heir S. Microfracture technique versus osteochondral autologous transplantation mosaicplasty in patients with articular chondral lesions of the knee: a prospective randomized trial with long-term follow-up. Knee Surgery, Sports Traumatology, Arthroscopy. 2014; 22(6):1207-1215

454. Unsal S, Caglar-Yagci H, Kaya K, Sahin-Onat S, Ozel S. Comparison of effectiveness of intraarticular sodium hyaluronat and physical therapy applications in patients with knee osteoarthritis: randomized prospective study. Journal of rheumatology and medical rehabilitation. 2008; 19(1):16-22

455. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM et al. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ. 2013; 347:f5555

456. UzunkulaoGlu A, KerIm D, Ay S, ErgIn S. Effects of single-task versus dual-task training on balance performance in elderly patients with knee osteoarthritis. Archives of Rheumatology. 2020; 35(1):35-40

457. Uzunkulaoglu A, Yildirim IB, Gunes Aytekin M, Ay S. Effect of flamingo exercises on balance in patients with balance impairment due to senile osteoarthritis. Archives of Gerontology and Geriatrics. 2019; 81:48-52

458. Vaghela N, Mishra D, Patel J, Dani V. Promoting health and quality of life of patients with osteoarthritis of knee joint through non-pharmacological treatment strategies: A randomized controlled trial. Journal of Education & Health Promotion. 2020; 9:156

459. van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Bijlsma JW. Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up. Annals of the Rheumatic Diseases. 2001; 60(12):1123-1130

460. van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Lemmens JA et al. The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized clinical trial. Journal of Rheumatology. 1998; 25(12):2432-2439

461. Van Ginckel A, Hall M, Dobson F, Calders P. Effects of long-term exercise therapy on knee joint structure in people with knee osteoarthritis: A systematic review and meta-analysis. Seminars in Arthritis and Rheumatism. 2019; 48(6):941-949

462. van Gool CH, Penninx BW, Kempen GI, Rejeski WJ, Miller GD, van Eijk JT et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis and Rheumatism. 2005; 53(1):24-32

463. Veenhof C, Koke AJ, Dekker J, Oostendorp RA, Bijlsma JW, van Tulder MW et al. Effectiveness of behavioral graded activity in patients with osteoarthritis of the hip and/or knee: A randomized clinical trial. Arthritis and Rheumatism. 2006; 55(6):925-934

464. Villadsen A, Overgaard S, Holsgaard-Larsen A, Christensen R, Roos EM. Immediate efficacy of neuromuscular exercise in patients with severe osteoarthritis of the hip or knee: a secondary analysis from a randomized controlled trial. Journal of Rheumatology. 2014; 41(7):1385-1394

465. Vincent KR, Vasilopoulos T, Montero C, Vincent HK. Eccentric and concentric resistance exercise comparison for knee osteoarthritis. Medicine and Science in Sports and Exercise. 2019; 51(10):1977-1986

466. Vincent KR, Vincent HK. Concentric and eccentric resistance training comparison on physical function and functional pain outcomes in knee osteoarthritis: A randomized controlled trial. American Journal of Physical Medicine and Rehabilitation. 2020; 99(10):932-940

467. Waller B, Munukka M, Rantalainen T, Lammentausta E, Nieminen MT, Kiviranta I et al. Effects of high intensity resistance aquatic training on body composition and walking speed in women with mild knee osteoarthritis: a 4-month RCT with 12-month follow-up. Osteoarthritis and Cartilage. 2017; 25(8):1238-1246

468. Waller B, Ogonowska-Slodownik A, Vitor M, Lambeck J, Daly D, Kujala UM et al. Effect of therapeutic aquatic exercise on symptoms and function associated with lower limb osteoarthritis: systematic review with meta-analysis. Physical Therapy. 2014; 94(10):1383-1395

469. Wallis JA, Webster KE, Levinger P, Singh PJ, Fong C, Taylor NF. A walking program for people with severe knee osteoarthritis did not reduce pain but may have benefits for cardiovascular health: a phase II randomised controlled trial. Osteoarthritis and Cartilage. 2017; 25(12):1969-1979

470. Wang C, Iversen MD, McAlindon T, Harvey WF, Wong JB, Fielding RA et al. Assessing the comparative effectiveness of Tai Chi versus physical therapy for knee osteoarthritis: design and rationale for a randomized trial. BMC Complementary and Alternative Medicine. 2014; 14:333

471. Wang C, Schmid CH, Hibberd PL, Kalish R, Roubenoff R, Rones R et al. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. Arthritis and Rheumatism. 2009; 61(11):1545-1553

472. Wang C, Schmid CH, Hibberd PL, Kalish R, Roubenoff R, Rones R et al. Tai Chi for treating knee osteoarthritis: designing a long-term follow up randomized controlled trial. BMC Musculoskeletal Disorders. 2008; 9:108

473. Wang C, Schmid CH, Iversen MD, Harvey WF, Fielding RA, Driban JB et al. Comparative effectiveness of tai chi versus physical therapy for knee osteoarthritis: A randomized trial. Annals of Internal Medicine. 2016; 165(2):77-86

474. Wang H, Ma Y, Guo Y, Pan Y. Effects of exercise therapy for knee osteoarthritis. International Journal of Clinical and Experimental Medicine. 2018; 11(9):10002-10008

475. Wang P, Yang L, Li H, Lei Z, Yang X, Liu C et al. Effects of whole-body vibration training with quadriceps strengthening exercise on functioning and gait parameters in patients with medial compartment knee osteoarthritis: a randomised controlled preliminary study. Physiotherapy. 2016; 102(1):86-92

476. Wang P, Yang L, Liu C, Wei X, Yang X, Zhou Y et al. Effects of whole body vibration exercise associated with quadriceps resistance exercise on functioning and quality of life in patients with knee osteoarthritis: A randomized controlled trial. Clinical Rehabilitation. 2016; 30(11):1074-1087

477. Wang P, Yang X, Yang Y, Yang L, Zhou Y, Liu C et al. Effects of whole body vibration on pain, stiffness and physical functions in patients with knee osteoarthritis: a systematic review and meta-analysis. Clinical Rehabilitation. 2015; 29(10):939-951

478. Wang TJ, Belza B, Elaine Thompson F, Whitney JD, Bennett K. Effects of aquatic exercise on flexibility, strength and aerobic fitness in adults with osteoarthritis of the hip or knee. Journal of Advanced Nursing. 2007; 57(2):141-152

479. Wang TJ, Lee SC, Liang SY, Tung HH, Wu SF, Lin YP. Comparing the efficacy of aquatic exercises and land-based exercises for patients with knee osteoarthritis. Journal of Clinical Nursing. 2011; 20(17-18):2609-2622

480. Wang Y, Lu S, Wang R, Jiang P, Rao F, Wang B et al. Integrative effect of yoga practice in patients with knee arthritis: A PRISMA-compliant meta-analysis. Medicine. 2018; 97(31):e11742

481. Watanabe S, Someya F. Effect of body weight-supported walking on exercise capacity and walking speed in patients with knee osteoarthritis: A randomized controlled trial. Journal of the Japanese Physical Therapy Association. 2013; 16(1):28-35

482. Weng MC, Lee CL, Chen CH, Hsu JJ, Lee WD, Huang MH et al. Effects of different stretching techniques on the outcomes of isokinetic exercise in patients with knee osteoarthritis. Kaohsiung Journal of Medical Sciences. 2009; 25(6):306-315

483. Wetzels R, Trap-Liefers J, Van Weel C, Grol R, Wensing M. Self-management to improve functional status and mobility in older osteoarthritis patients: a randomised trial in general practice. 2005:145-157

484. Wetzels R, van Weel C, Grol R, Wensing M. Family practice nurses supporting self-management in older patients with mild osteoarthritis: a randomized trial. BMC Family Practice. 2008; 9:7

485. Williamson L, Wyatt MR, Yein K, Melton JT. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. Rheumatology. 2007; 46(9):1445-1449

486. Williamson W, Kluzek S, Roberts N, Richards J, Arden N, Leeson P et al. Behavioural physical activity interventions in participants with lower-limb osteoarthritis: a systematic review with meta-analysis. BMJ Open. 2015; 5(8):e007642

487. Witteveen AGH, Kerkhoffs GMMJ, Den Broeder AA, Sierevelt IN, Hofstad CJ. Conservative treatment for osteoarthritis of the ankle. Cochrane Database of Systematic Reviews 2013, Issue 7. Art. No.: CD010643. DOI: http://dx.doi.org/10.1002/14651858.CD010643.

488. Wood LRJ, Blagojevic-Bucknall M, Stynes S, D'Cruz D, Mullis R, Whittle R et al. Impairment-targeted exercises for older adults with knee pain: a proof-of-principle study (TargET-Knee-Pain). BMC Musculoskeletal Disorders. 2016; 17(1)

489. Wortley M, Zhang S, Paquette M, Byrd E, Baumgartner L, Klipple G et al. Effects of resistance and Tai Ji training on mobility and symptoms in knee osteoarthritis patients. Journal of sport and health science. 2013; 2(4):209-214

490. Wyatt FB, Milam S, Manske RC, Deere R. The effects of aquatic and traditional exercise programs on persons with knee osteoarthritis. Journal of Strength and Conditioning Research. 2001; 15(3):337-340

491. Xiao C, Zhuang Y, Kang Y. Effects of Wu Qin xi Qigong exercise on physical functioning in elderly people with knee osteoarthritis: A randomized controlled trial. Geriatrics & gerontology international. 2020; 20(10):899-903

492. Xiao CM, Li JJ, Kang Y, Zhuang YC. Follow-up of a Wuqinxi exercise at home programme to reduce pain and improve function for knee osteoarthritis in older people: a randomised controlled trial. Age and Ageing. 2021; 50(2):570-575

493. Xu T, Yang D, Liu K, Gao Q, Lu H, Qiao Y et al. Efficacy and safety of a self-developed home-based enhanced knee flexion exercise program compared with standard supervised physiotherapy to improve mobility and quality of life after total knee arthroplasty: a randomized control study. Journal of Orthopaedic Surgery. 2021; 16(1):382

494. Yazigi F, Espanha M, Vieira F, Messier SP, Monteiro C, Veloso AP. The PICO project: aquatic exercise for knee osteoarthritis in overweight and obese individuals. BMC Musculoskeletal Disorders. 2013; 14:320

495. Ye J, Cai S, Zhong W, Cai S, Zheng Q. Effects of tai chi for patients with knee osteoarthritis: a systematic review. Journal of Physical Therapy Science. 2014; 26(7):1133-1137

496. Ye J, Simpson MW, Liu Y, Lin W, Zhong W, Cai S et al. The effects of baduanjin qigong on postural stability, proprioception, and symptoms of patients with knee osteoarthritis: A randomized controlled trial. Frontiers in Medicine. 2019; 6:307

497. Ye J, Zheng Q, Zou L, Yu Q, Veronese N, Grabovac I et al. Mindful exercise (baduanjin) as an adjuvant treatment for older adults (60 years old and over) of knee osteoarthritis: A randomized controlled trial. Evidence-Based Complementary & Alternative Medicine: eCAM. 2020; 2020:9869161

498. Yilmaz M, Sahin M, Algun ZC. Comparison of effectiveness of the home exercise program and the home exercise program taught by physiotherapist in knee osteoarthritis. Journal of Back and Musculoskeletal Rehabilitation. 2019; 32(1):161-169

499. You Y, Liu J, Tang M, Wang D, Ma X. Effects of Tai Chi exercise on improving walking function and posture control in elderly patients with knee osteoarthritis: A systematic review and meta-analysis. Medicine. 2021; 100(16):e25655

500. Zafar H, Alghadir A, Anwer S, Al-Eisa E. Therapeutic effects of whole-body vibration training in knee osteoarthritis: a systematic review and meta-analysis. Archives of Physical Medicine and Rehabilitation. 2015; 96(8):1525-1532

501. Zammit G, Menz H, Munteanu S, Landorf K, Gilheany M. Interventions for treating osteoarthritis of the big toe joint. Cochrane Database of Systematic Reviews 2010, Issue 9. Art. No.: CD007809. DOI: 10.1002/14651858.CD007809.pub2.

502. Zgibor JC, Ye L, Boudreau RM, Conroy MB, Vander Bilt J, Rodgers EA et al. Community-based healthy aging interventions for older adults with arthritis and multimorbidity. Journal of Community Health. 2017; 42(2):390-399

503. Zhang Y, Huang L, Su Y, Zhan Z, Li Y, Lai X. The effects of traditional chinese exercise in treating knee osteoarthritis: A systematic review and meta-analysis. PLoS ONE [Electronic Resource]. 2017; 12(1):e0170237

504. Zhang Z, Huang L, Liu Y, Wang L. Effect of tai chi training on plantar loads during walking in individuals with knee osteoarthritis. BioMed Research International. 2020; 2020:3096237

505. Zhu Q, Huang L, Wu X, Wang L, Zhang Y, Fang M et al. Effects of Tai Ji Quan training on gait kinematics in older Chinese women with knee osteoarthritis: A randomized controlled trial. Journal of Sport & Health Science. 2016; 5(3):297-303

Appendices

Appendix A – Review protocols

Review protocol for the clinical and cost-effectiveness of exercise for the management of osteoarthritis

ID	Field	Content
0.	PROSPERO registration number	N/A
1.	Review title	What is the clinical and cost-effectiveness of exercise therapy for the management of osteoarthritis?
2.	Review question	3.1 What is the clinical and cost-effectiveness of exercise therapy for the management of osteoarthritis?
3.	Objective	To assess the clinical and cost-effectiveness of exercise as interventions for the management of osteoarthritis.
4.	Searches	The following databases will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE Searches will be restricted by: • English language • Human studies • Letters and comments are excluded

		Other searches: • Inclusion lists of relevant systematic reviews will be checked by the reviewer. The searches may be re-run 6 weeks before final submission of the review and further studies retrieved for inclusion if relevant. The full search strategies for MEDLINE database will be published in the final review.
5.	Condition or domain being studied	Osteoarthritis (of any joint) in adults (defined as a clinical diagnosis of osteoarthritis with or without imaging)
6.	Population	Inclusion: • Adults (age ≥16 years) with osteoarthritis affecting any joint Exclusion: • Children (age <16 years) • People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, hemochromatosis, haemophilic arthropathy, diseases of childhood that may predispose to osteoarthritis, and malignancy). • Studies with an unclear population (e,g, proportion of participants with osteoarthritis unclear). • Spinal osteoarthritis
7.	Intervention/Exposure/Test	Interventions (minimum duration 1 week):

		• Supervised strength exercise • Supervised aerobic exercise • Other supervised exercise (including flexibility, proprioception)* • Supervised mixed modality exercise (e.g. aerobic and strength exercise combined) • Unsupervised strength exercise • Unsupervised aerobic exercise • Other unsupervised exercise (including flexibility, proprioception)* • Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined) *Subgroup analysis if heterogeneity is present within this group
8.	Comparator/Reference standard/Confounding factors	• Each other • Pharmacological treatment*** • No exercise intervention (including either): ○ Exercise versus no treatment* ○ Exercise plus additional treatment versus additional treatment alone** **No treatment defined as either (1) doing nothing or (2) very low intensity intervention such as advice* ***Inclusion of studies where additional treatment is the same in each arm will be assessed on a case by case basis. Studies including high intensity additional treatment may not be included due to the risk that treatment could have an interaction with the intervention of interest and mask the true treatment effect.* ****Pool classes of pharmacological treatment but conduct subgroup analysis if heterogeneity is present*
9.	Types of study to be included	• Systematic reviews of RCTs • Parallel RCTs
10.	Other exclusion criteria	• Non-English language studies

		• Non-randomised/observational studies • Crossover RCTs Abstracts will be excluded as it is expected there will be sufficient full text published studies available.
11.	Context	N/A
12.	Primary outcomes (critical outcomes)	Stratify by ≤/>3 months (longest time-point in each): • Health-related quality of life [validated patient-reported outcomes, continuous data prioritised] • Pain [validated patient-reported outcomes, continuous data prioritised] • Physical function [validated patient-reported outcomes, continuous data prioritised] The COMET database was searched and several core outcome sets were identified for specific sites of osteoarthritis (including hand, knee and hip). The committee took these into account when defining outcomes: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.22868 https://www.ncbi.nlm.nih.gov/pubmed/26136489 https://www.ncbi.nlm.nih.gov/pubmed/30647185 The committee did not include stiffness or global scores as Delphi discussions by the OMERACT group have found these to not be as important to people with osteoarthritis or clinicians. The outcomes included were universal for all groups allowing for broader comparisons.
13.	Secondary outcomes (important outcomes)	• Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flares [validated patient-reported outcomes, continuous data prioritised]

		• Serious adverse events [dichotomous]
14.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. EviBASE will be used for data extraction. Study investigators may be contacted for missing data where time and resources allow.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual For intervention reviews the following checklists will be used according to the study design being assessed: • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data • a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.

16.	Strategy for data synthesis	• Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5). • GRADEpro will be used to assess the quality of evidence for each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) will be appraised for each outcome. Publication bias is tested for when there are more than 5 studies for an outcome. The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ • Where meta-analysis is not possible, data will be presented and quality assessed individually per outcome. • WinBUGS will be used for network meta-analysis, if possible given the data identified. Heterogeneity between studies in the effect measures will be assessed using the I^2 statistic and visual inspection. We will consider an I^2 value great than 50% as indicative of substantial heterogeneity. If significant heterogeneity is identified during meta-analysis then subgroup analysis, using subgroups predefined by the GC, will take place. If this does not explain the heterogeneity, the results will be presented using a random-effects model.
17.	Analysis of sub-groups	Subgroup analysis to be conducted if heterogeneity in the meta-analysis is present: • Site of osteoarthritis • Diagnosis with or without imaging (indicative of severity)

		• Multimorbidity (high versus low morbidity score; as defined by study, measured by validated instruments e.g. Charlson Comorbidity Index • Age (≤/> 75 years) • For pharmacological treatment comparison: subgroup by class of medicines • For 'other exercise' intervention: subgroup by type of exercise (e.g. proprioception versus flexibility) • Group vs individual interventions		
18.	Type and method of review	☒	Intervention	
		☐	Diagnostic	
		☐	Prognostic	
		☐	Qualitative	
		☐	Epidemiologic	
		☐	Service Delivery	
		☐	Other (please specify)	
19.	Language	English		
20.	Country	England		
21.	Anticipated or actual start date	23/08/2019		
22.	Anticipated completion date	25/08/2021		
23.		Review stage	Started	Completed

	Stage of review at time of this submission	Preliminary searches	☑	☐
		Piloting of the study selection process	☐	☐
		Formal screening of search results against eligibility criteria	☐	☐
		Data extraction	☐	☐
		Risk of bias (quality) assessment	☐	☐
		Data analysis	☐	☐
24.	Named contact	5a. Named contact National Guideline Centre 5b Named contact e-mail [Guideline email]@nice.org.uk [Developer to check with Guideline Coordinator for email address] 5e Organisational affiliation of the review National Institute for Health and Care Excellence (NICE) and the National Guideline Centre		
25.	Review team members	From the National Guideline Centre:		

		Carlos Sharpin [Guideline lead] Julie Neilson [Senior systematic reviewer] George Wood [Systematic reviewer] Margaret Constanti [Senior health economist] Joseph Runicles [Information specialist] Amber Hernaman [Project manager]
26.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10127
29.	Other registration details	
30.	Reference/URL for published protocol	

31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.	
32.	Keywords	Adults; Education; Exercise; Intervention; Non-Pharmacological; Osteoarthritis; Programmes; Weight loss	
33.	Details of existing review of same topic by same authors		
34.	Current review status	☒	Ongoing
		☐	Completed but not published
		☐	Completed and published
		☐	Completed, published and being updated
		☐	Discontinued
35..	Additional information	N/A	
36.	Details of final publication	www.nice.org.uk	

Table 55: Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken for all years using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2005, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Studies published in 2005 or later, that were included in the previous guidelines, will be reassessed for inclusion and may be included or selectively excluded based on their relevance to the questions covered in this update and whether more applicable evidence is also identified. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[317] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included.

Where there is discretion

The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below.

The health economist will be guided by the following hierarchies.

Setting:

- UK NHS (most applicable).
- OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden).
- OECD countries with predominantly private health insurance systems (for example, Switzerland).
- Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations.

Health economic study type:

- Cost–utility analysis (most applicable).
- Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis).
- Comparative cost analysis.
- Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations.

Year of analysis:

- The more recent the study, the more applicable it will be.
- Studies published in 2005 or later (including any such studies included in the previous guidelines) but that depend on unit costs and resource data entirely or predominantly from before 2005 will be rated as 'Not applicable'.
- Studies published before 2005 (including any such studies included in the previous guidelines) will be excluded before being assessed for applicability and methodological limitations.

Quality and relevance of effectiveness data used in the health economic analysis:

- The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B – Literature search strategies

- What is the clinical and cost-effectiveness of exercise therapy for the management of osteoarthritis?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[317]

For more information, please see the Methodology review published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using an Osteoarthritis population. All results were then sifted for each question. Search filters were applied to the search where appropriate.

Table 56: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 17 November 2021	Randomised controlled trials Systematic review studies Exclusions (animals studies, letters, comments)
Embase (OVID)	1974 – 17 November 2021	Randomised controlled trials Systematic review studies Exclusions (animals studies, letters, comments)
The Cochrane Library (Wiley)	Cochrane Reviews to 2021 Issue 11 of 12 CENTRAL to 2021 Issue 11 of 12	None

Medline (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14

16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice or rodent*).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	randomized controlled trial.pt.
28.	controlled clinical trial.pt.
29.	randomi#ed.ti,ab.
30.	placebo.ab.
31.	randomly.ti,ab.
32.	Clinical Trials as topic.sh.
33.	trial.ti.
34.	or/27-33
35.	Meta-Analysis/
36.	exp Meta-Analysis as Topic/
37.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
38.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
39.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
40.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
41.	(search* adj4 literature).ab.
42.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
43.	cochrane.jw.
44.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
45.	or/35-44
46.	26 and (34 or 45)

Embase (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.

12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice or rodent*).ti.
22.	or/14-21
23.	6 not 22
24.	Limit 23 not English language
25.	random*.ti,ab.
26.	factorial*.ti,ab.
27.	(crossover* or cross over*).ti,ab.
28.	((doubl* or singl*) adj blind*).ti,ab.
29.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
30.	crossover procedure/
31.	single blind procedure/
32.	randomized controlled trial/
33.	double blind procedure/
34.	or/25-33
35.	systematic review/
36.	meta-analysis/
37.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
38.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
39.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
40.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
41.	(search* adj4 literature).ab.
42.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
43.	cochrane.jw.
44.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
45.	or/35-44
46.	24 and (34 or 45)

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Osteoarthritis] explode all trees
#2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*):ti,ab
#3.	(degenerative near/2 arthritis):ti,ab
#4.	coxarthrosis:ti,ab
#5.	gonarthrosis:ti,ab

#6.	(or #1-#5)

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to a Gout population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA – this ceased to be updates after March 2018). NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics studies and quality of life studies. Searches for quality of life studies were run for general information.

Table 57: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	1 January 2014 – 17 November 2021	Health economics studies Quality of life studies Exclusions (animals studies, letters, comments)
Embase	1 January 2014 – 17 November 2021	Health economics studies Quality of life studies Exclusions (animals studies, letters, comments)
Centre for Research and Dissemination (CRD)	HTA - Inception – 31 March 2018 NHSEED - Inception to 31 March 2015	None

Medline (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16

18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice or rodent*).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	Economics/
28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/
31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/27-42
44.	quality-adjusted life years/
45.	sickness impact profile/
46.	(quality adj2 (wellbeing or well being)).ti,ab.
47.	sickness impact profile.ti,ab.
48.	disability adjusted life.ti,ab.
49.	(qal* or qtime* or qwb* or daly*).ti,ab.
50.	(euroqol* or eq5d* or eq 5*).ti,ab.
51.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
52.	(hui or hui1 or hui2 or hui3).ti,ab.
53.	(health* year* equivalent* or hye or hyes).ti,ab.
54.	discrete choice*.ti,ab.
55.	rosser.ti,ab.
56.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.

57.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
58.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
59.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
60.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
61.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
62.	or/44-61
63.	26 and (43 or 62)

Embase (Ovid) search terms

1.	exp osteoarthritis/
2.	(osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*).ti,ab.
3.	(degenerative adj2 arthritis).ti,ab.
4.	coxarthrosis.ti,ab.
5.	gonarthrosis.ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice or rodent*).ti.
22.	or/14-21
23.	6 not 22
24.	Limit 23 to English language
25.	health economics/
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/
31.	budget*.ti,ab.

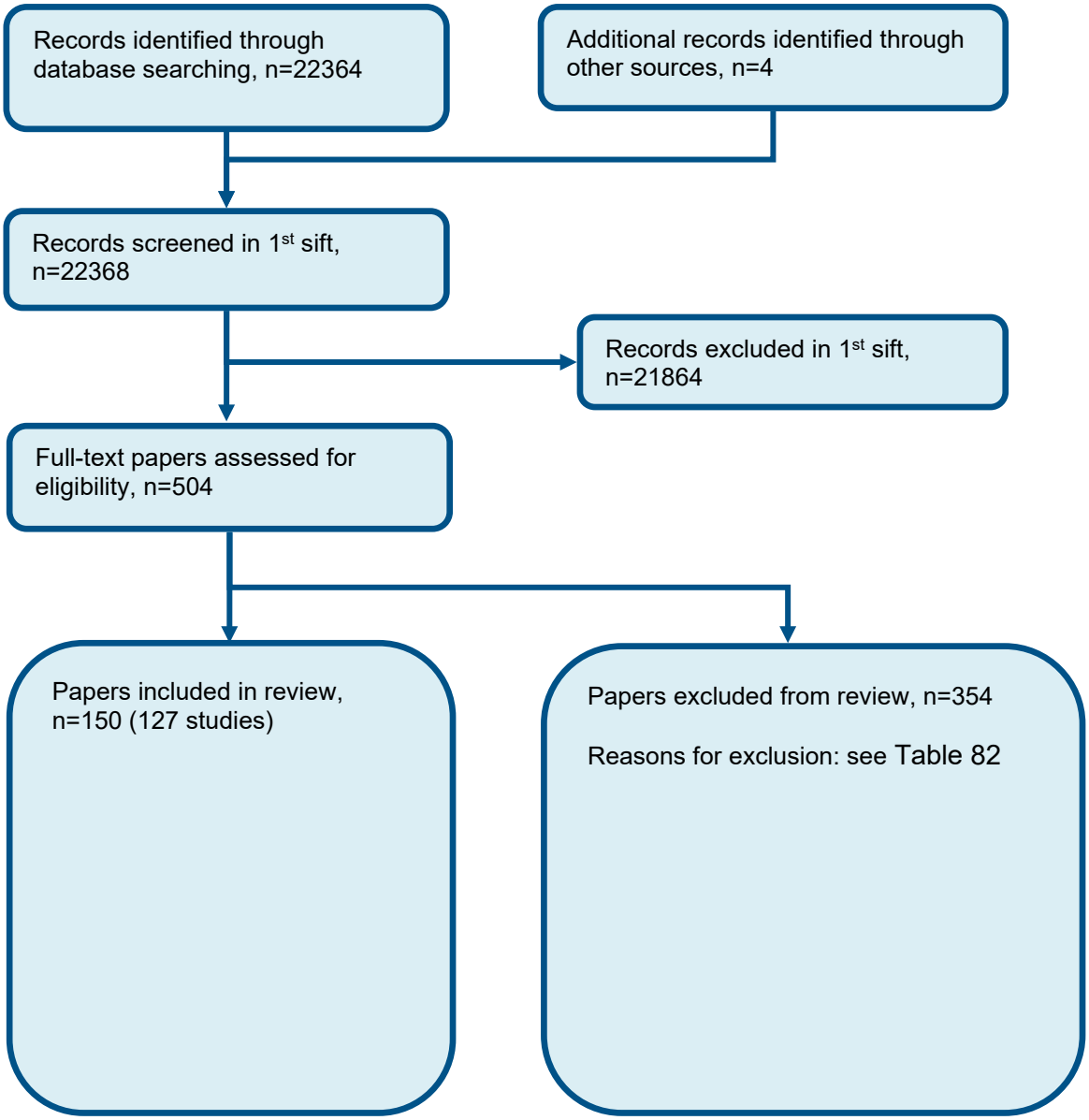
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	quality adjusted life year/
40.	"quality of life index"/
41.	short form 12/ or short form 20/ or short form 36/ or short form 8/
42.	sickness impact profile/
43.	(quality adj2 (wellbeing or well being)).ti,ab.
44.	sickness impact profile.ti,ab.
45.	disability adjusted life.ti,ab.
46.	(qal* or qtime* or qwb* or daly*).ti,ab.
47.	(euroqol* or eq5d* or eq 5*).ti,ab.
48.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
49.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
50.	(hui or hui1 or hui2 or hui3).ti,ab.
51.	(health* year* equivalent* or hye or hyes).ti,ab.
52.	discrete choice*.ti,ab.
53.	rosser.ti,ab.
54.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
55.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
56.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
57.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
58.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
59.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
60.	or/39-59
61.	24 and (38 or 60)

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Osteoarthritis EXPLODE ALL TREES
#2.	((osteoarthriti* or osteo-arthriti* or osteoarthrotic or osteoarthros*))
#3.	((degenerative adj2 arthritis))
#4.	(coxarthrosis)
#5.	(gonarthrosis)
#6.	#1 OR #2 OR #3 OR #4 OR #5
#7.	(#6) IN NHSEED
#8.	(#6) IN HTA

Appendix C – Effectiveness evidence study selection

Figure 1: Flow chart of clinical study selection for the review of the clinical and cost-effectiveness of exercise for people with osteoarthritis



Appendix D – Effectiveness evidence

Study (subsidiary papers)	Abbott 2013[2] (Abbott 2019[3])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=206)
Countries and setting	Conducted in New Zealand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Other: 1 year (including intervention and follow up afterwards)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People referred with hip or knee osteoarthritis of those referred for consideration for hip or knee joint replacement surgery based on the clinical criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People meeting the clinical criteria of osteoarthritis of the hip or knee established by the American College of Rheumatology
Exclusion criteria	Rheumatoid arthritis; previous knee or hip joint replacement surgery of the affected joint; any other surgical procedure on the lower limbs in the previous 6 months; surgical procedure on the lower limbs planned in the next 6 months; initiation of opioid analgesia or corticosteroid or analgesic injection intervention for hip or knee pain within the previous 30 days; physical impairments unrelated to the hip or knee which would prevent safe participation in exercise, manual therapy, walking or stationary cycling; inability to comprehend and complete study assessments or comply with study instructions; or stated inability to attend or complete the proposed course of intervention and follow-up schedule
Recruitment/selection of patients	Recruited by GP referral or referral to orthopaedic clinics for consideration for hip or knee joint replacement
Age, gender and ethnicity	Age - Mean (SD): 66.6 (9.6). Gender (M:F): 92:114. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 2.7 (1.4) years

Indirectness of population	No indirectness
Interventions	(n=51) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Supervised programme of warm-up/aerobic, muscle strengthening, muscle stretching and neuromuscular control exercises. Additional interventions were prescribed individually for each participant on the basis of the physical examination findings. In addition they prescribed a home exercise programme to be completed three times a week.. Duration 1 year. Concurrent medication/care: Usual care was provided for all participants by their GP and other healthcare providers. This was not limited or influenced in any way.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Aerobic, strengthening, neuromodulatory). (n=104) Intervention 2: Other. Multi-modal exercise physiotherapy and manual therapy, and manual therapy alone groups. Duration 1 year. Concurrent medication/care: Usual care was provided for all participants by their GP and other healthcare providers. This was not limited or influenced in any way.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These groups were not included in the analysis (n=51) Intervention 3: No treatment. No trial physiotherapy. Duration 1 year. Concurrent medication/care: Usual care was provided for all participants by their GP and other healthcare providers. This was not limited or influenced in any way.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Pain at > 3 months
- Actual outcome: Pain intensity score (VAS) at 1 year; Group 1: mean -0.96 (SD 2.51); n=51, Group 2: mean -0.06 (SD 2.39); n=51; VAS 0-10 Top=High is poor outcome; Comments: Reports change scores (95% CIs). Reported exercise: -0.96 (-1.65 to -0.27). Reported usual care: -0.06 (-0.71 to 0.60).

Risk of bias: All domain - Low, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: 44 participants of the trial had a hip or knee replacement over the year, unclear how many of the participants in our analysis had surgery; Baseline details: Reports gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up - 1 due to dementia, 1 due to personal reasons; Group 2 Number missing: 4, Reason: 4 lost to follow up - 1 deceased, 2 due to ill health, 1 due to ill health of spouse Protocol outcome 2: Serious adverse events at > 3 months - Actual outcome: Adverse events at 1 year; Group 1: 0/51, Group 2: 1/51; Comments: No treatment: 1 non-trial related death Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: Serious indirectness, Comments: 44 participants of the trial had a hip or knee replacement over the year, unclear how many of the participants in our analysis had surgery; Baseline details: Reports gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 lost to follow up - 1 due to dementia, 1 due to personal reasons; Group 2 Number missing: 4, Reason: 4 lost to follow up - 1 deceased, 2 due to ill health, 1 due to ill health of spouse	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study (subsidiary papers)	Aglamis 2008[6] (Aglamis 2009[5])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=34)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographic grade 2-4 Kellgren Lawrence knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Knee osteoarthritis (radiographic grade 2-4 using Kelgren-Lawrence criteria) being independent in daily activity and being between 50-69 years of age
Exclusion criteria	Having intra-articular injections in their last six months; being involved in regular physical activity and physiotherapy using any assistive equipment; being unable to exercise having a chronic condition.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 55.7 (5.0). Gender (M:F): 9:25. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Low morbidity score (Mean number of comorbidities is less than 1.5). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=17) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Exercise program supervised three times per week by four trainers and one health technician. The programs involved aerobic, strength and flexibility training. Consists of a 10 min warm up and 15 min cool down session with an aerobic training phase consisting a 20 min walk at a comfortable pace,and a functional strengthening exercise performed in a circuit with step-ups, chair-squat, standing hip extension, and knee mid-flexion to end-range extension (in sitting position), utilising body weight as resistance with 12 repetitions in a single set increased by 3 at the third week. Flexibility exercises included a static stretching program with the hip external-internal rotator muscle, hamstring, quadriceps, calf,

	invertor, evertor, plantar fexor and dorsalflexor muscle stretch in sitting and standing positions. The intensity was three repetitions per muscle group and a duration of 20-sec hold.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Other (Aerobic, strengthening, flexibility). (n=17) Intervention 2: No treatment. No treatment control. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical function subscale at 12 weeks; Group 1: mean 87.2 (SD 9.7); n=16, Group 2: mean 36.4 (SD 7.9); n=9; SF-36 physical function subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 65.3 (21.3). Baseline placebo: 45.6 (13.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact
- Actual outcome: SF-36 role-physical subscale at 12 weeks; Group 1: mean 90.6 (SD 25.6); n=16, Group 2: mean 5.6 (SD 11); n=9; SF-36 role-physical subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 43.8 (39.3). Baseline placebo: 27.8 (38.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact
- Actual outcome: SF-36 body pain subscale at 12 weeks; Group 1: mean 67.5 (SD 18.1); n=16, Group 2: mean 20 (SD 17.5); n=9; SF-36 body pain subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 50.3 (22.7). Baseline control: 25.3 (23.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact
- Actual outcome: SF-36 social functioning subscale at 12 weeks; Group 1: mean 96.9 (SD 7.2); n=16, Group 2: mean 38.6 (SD 35.9); n=9; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 67.2 (31.3). Baseline no treatment: 58.3 (34.8).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

- Actual outcome: SF-36 mental health subscale at 12 weeks; Group 1: mean 79.3 (SD 8); n=16, Group 2: mean 46.4 (SD 13.8); n=9; SF-36 mental health subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 57 (21.9). Baseline control: 41.3 (20.1).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

- Actual outcome: SF-36 emotional role subscale at 12 weeks; Group 1: mean 87.5 (SD 26.9); n=16, Group 2: mean 14.7 (SD 33.7); n=9; SF-36 emotional role subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 58.3 (46.4). Baseline control: 40.7 (44.5).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

- Actual outcome: SF-36 vitality subscale at 12 weeks; Group 1: mean 76.3 (SD 9.9); n=16, Group 2: mean 24.4 (SD 25.2); n=9; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 50.3 (21.2). Baseline control: 28.3 (27).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

- Actual outcome: SF-36 general health subscale at 12 weeks; Group 1: mean 77.5 (SD 10.2); n=16, Group 2: mean 40 (SD 20.5); n=9; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 58.4 (20). Baseline control: 38.3 (29.7).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

Protocol outcome 2: Pain at </=3 months

- Actual outcome: Visual analogue scale at 12 weeks; Group 1: mean 0.7 (SD 1); n=16, Group 2: mean 7.7 (SD 2.3); n=9; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.4 (1.9). Baseline control: 6.2 (3.4).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Differences in all outcomes at baseline. Similar for age, number of use analgesics, BMI and number of chronic conditions.; Group 1 Number missing: 1, Reason: 1 reactive arthritis; Group 2 Number missing: 8, Reason: 3 reason not stated, 2 change of city, 1 low back pain, 1 intraarticular injection, 1 no contact

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Allen 2018[12] (Anderson 2019[20], Pignato 2018[356])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=350)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed by a physician, radiographic evidence of knee osteoarthritis or a self-report of physician diagnosis along with items based on the American College of Rheumatology clinical criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Radiographic evidence of knee osteoarthritis, physician diagnosis of knee osteoarthritis in the medical record, or self-report of physician diagnosis along with items based on the American College of Rheumatology clinical criteria; self report of pain, aching or stiffness in one or both knees on most days of the weekend
Exclusion criteria	No regular internet access; currently meeting Department of Health and Human Services Guidelines for physical activity; currently completing series of physiotherapy visits for knee osteoarthritis; diagnosis of gout in the knee, rheumatoid arthritis, fibromyalgia, or other systemic rheumatic disease; severe dementia or other memory loss condition; active diagnosis of psychosis or current uncontrolled substance abuse disorder; on waiting list for arthroplasty; hospitalisation for a strong,, heart attack, heart failure, or had surgery for blocked arteries in the past 3 months; total joint replacement knee surgery, other knee surgery, meniscus tear, or ACL tear in the past 6 months; severely impaired hearing or speech; unable to speak English; serious or terminal illness as indicated by referral to hospice or palliative care; other health problem that would prohibit participation in the study; nursing home residence; current participation in another osteoarthritis intervention; fall history deemed by a study physical therapist co-investigator to impose risk for potential injury with participation in a home-based exercise program study
Recruitment/selection of patients	Two methods: active recruitment of patients with evidence of knee osteoarthritis in the UNC medical record, as well as participants with knee osteoarthritis in the Johnston County Osteoarthritis Project; advertisement within UNC and the surrounding communities.

Age, gender and ethnicity	Age - Mean (SD): 65.3 (11.1). Gender (M:F): 99:251. Ethnicity: White = 255, "Non-white" = 95
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear (Could be diagnosed with imaging or could not be). 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 13.1 (11.7) years
Indirectness of population	No indirectness
Interventions	(n=142) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Internet based exercise therapy, including: tailored exercises (including strengthening, stretching and aerobic activity recommendation); exercise progression recommendations; video display of exercises; automated reminders to engage with the website; progress tracking.. Duration 12 months. Concurrent medication/care: Not stated/unclear. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strengthening, stretching, aerobic). (n=140) Intervention 2: Other. Physiotherapy including exercise, manual therapy, joint protection and devices. This group was not included in the analysis as this was a treatment package and so was considered in another review. Duration 12 months. Concurrent medication/care: Not stated/unclear. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=68) Intervention 3: No treatment. Waiting list control. Duration 12 months. Concurrent medication/care: Not stated/unclear. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: WOMAC physical function subscale at 12 months; Group 1: mean -3.4 (SD 10.4); n=142, Group 2: mean -1.51 (SD 9.5); n=68; WOMAC

physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 21.8 (12.7). Baseline no treatment: 23.9 (13.8). Reports change scores and 95% confidence intervals. Reported exercise: -3.4 (-5.11, -1.7). Reported no treatment: -1.51 (-3.76, 0.74). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, marital status, education, employment, financial status, general health, BMI, joints with symptoms, duration of symptoms and baseline values of outcomes; Group 1 Number missing: 30, Reason: 10 dropped, 14 withdrew, 8 missed visit, 7 lost to follow up (the numbers of their flow diagram don't add up, they state 112 completed); Group 2 Number missing: 5, Reason: 1 dropped, 1 withdrew, 3 missed visit, 3 lost to follow up (the numbers of their flow diagram don't add up, they state 63 completed) Protocol outcome 2: Pain at > 3 months - Actual outcome: WOMAC pain subscale at 12 months; Group 1: mean -1.15 (SD 3.4); n=142, Group 2: mean -0.64 (SD 3.09); n=68; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.0 (3.9) . Baseline no treatment: 6.1 (3.5). Reports change scores and 95% confidence intervals. Reported exercise: -1.15 (-1.71, -0.59). Reported no treatment: -0.64 (-1.38, 0.09). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, marital status, education, employment, financial status, general health, BMI, joints with symptoms, duration of symptoms and baseline values of outcomes; Group 1 Number missing: 30, Reason: 10 dropped, 14 withdrew, 8 missed visit, 7 lost to follow up (the numbers of their flow diagram don't add up, they state 112 completed); Group 2 Number missing: 5, Reason: 1 dropped, 1 withdrew, 3 missed visit, 3 lost to follow up (the numbers of their flow diagram don't add up, they state 63 completed) Protocol outcome 3: Serious adverse events at > 3 months - Actual outcome: Study-related adverse events at 12 months; Group 1: 4/142, Group 2: 0/68; Comments: Exercise: 2 increased knee pain, 1 shoulder pain, 1 ankle pain Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, ethnicity, marital status, education, employment, financial status, general health, BMI, joints with symptoms, duration of symptoms and baseline values of outcomes; Group 1 Number missing: 30, Reason: 10 dropped, 14 withdrew, 8 missed visit, 7 lost to follow up (the numbers of their flow diagram don't add up, they state 112 completed); Group 2 Number missing: 5, Reason: 1 dropped, 1 withdrew, 3 missed visit, 3 lost to follow up (the numbers of their flow diagram don't add up, they state 63 completed)	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	An 2008[18]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=28)
Countries and setting	Conducted in China; Setting: Community
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosed using the clinical criteria for the classification of idiopathic osteoarthritis of the knee developed by the American College of Rheumatology.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Knee pain in the last month and within the current age of greater than 55 years. In addition they had to have had at least three of the following: stiffness <30 minutes, crepitus, bony tenderness, bony enlargement, or absence of palpable warmth. The person had to have symptomatic osteoarthritis in at least one knee for at least 6 months prior to study entry; the person had no current participation in an exercise programe; the person had shown willingness to participate in the study, and to provide a signed informed consent.
Exclusion criteria	Symptoms of locking or instability; a corticosteroid or hyaluronic acid injection in the symptomatic knee within 12 months prior to study entry; a history of any of the following: knee surgery within the last 2 years, a joint replacement at any point, or a priori diagnosis of inflammatory arthritis; people with significant medical complications (e.g. hemiplegia, heart disease, and gout, which might affect the results of the study)
Recruitment/selection of patients	Recruited from the Wuliqiao Community, which is located in the Lu-Wan District, the center of urban Shanghai, People's Republic of China
Age, gender and ethnicity	Age - Mean (SD): 65.0 (7.5). Gender (M:F): 0:28. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=14) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Baduanjin in eight sections repeated 20 times (practice as round, slow and nonstop phases). Delivered by a certified senior instructor during a 30 minute session five times a week.. Duration 8 weeks. Concurrent medication/care: No change in medication for arthritis was permitted during the trial. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (n=14) Intervention 2: No treatment. No treatment. Duration 8 weeks. Concurrent medication/care: No change in medication for arthritis was permitted during the trial. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Other (The study was partially supported by the Wuliqiao Community. We would also like to thank Prof. Nicholas Bellamy for the WOMACTM VA3.1 that he supplied us for free.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 general health subscale at 8 weeks; Group 1: mean 61.2 (SD 17.9); n=11, Group 2: mean 49.1 (SD 25.9); n=10; SF-36 general health 0-100 Top=High is good outcome; Comments: Baseline exercise: 58.8 (13.7). Baseline no treatment: 48.9 (26.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time
- Actual outcome: SF-36 social functioning subscale at 8 weeks; Group 1: mean 75 (SD 28.5); n=11, Group 2: mean 77.5 (SD 24.2); n=10; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 63.6 (25.9). Baseline no treatment: 77.5 (22.7).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time
- Actual outcome: SF-36 mental health subscale at 8 weeks; Group 1: mean 76.4 (SD 15.3); n=11, Group 2: mean 67 (SD 8.2); n=10; SF-36 mental health subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 68.6 (10.5). Baseline control: 67.5 (12.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 347.5 (SD 382.8); n=11, Group 2: mean 511.8 (SD 381.6); n=10; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Baseline exercise: 406.4 (330.8). Baseline control: 296.5 (196.0). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 71.1 (SD 110.1); n=11, Group 2: mean 138.2 (SD 112.6); n=10; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline exercise: 150.0 (99.7). Baseline control: 116.8 (74.5). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Adverse events at 8 weeks; Group 1: 0/14, Group 2: 0/14; Comments: No one had adverse events. People withdrew from the study but this was not due to adverse events Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 social functioning subscale at baseline. Otherwise similar to age and BMI.; Group 1 Number missing: 3, Reason: 3 withdrew - 1 out of the area, 2 had no time; Group 2 Number missing: 4, Reason: 4 withdrew - 1 out of the area, 3 had no time	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Anon 2016[119]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=35)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: ACR grade 1-3 osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with American College of Rheumatology grade 1-3 osteoarthritis
Exclusion criteria	Systemic inflammatory arthritis; secondary knee osteoarthritis and grade 4 osteoarthritis; knee or hip replacement; diabetes mellitus; uncontrolled hypertension; respiratory disease; cognitive or communicative impairments
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 51.1 (6.0). Gender (M:F): 6:24. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: ACR grade 1-3 Duration of symptoms (mean [SD]): 33.9 (36.9) months
Indirectness of population	No indirectness
Interventions	(n=19) Intervention 1: Exercise - Supervised strength exercise. 30 minute exercise regimen aiming to strengthen the quadriceps and hamstring muscles. Five sessions were scheduled each week for 15 days. After completing the 15 sessions, the people continued that regimen at home and were called every 2 weeks to assess their adherence to the program.. Duration 12 weeks. Concurrent medication/care: 30 minutes along with TENS (pulse duration of 150msec, frequency of 120Hz, amplitude of 50mA).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=16) Intervention 2: No treatment. No exercise treatment. Duration 12 weeks. Concurrent medication/care: 30 minutes along with TENS (pulse duration of 150msec,

	frequency of 120Hz, amplitude of 50mA).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was financially supported by the Haydarpasa Training Hospital Research (2012- Project number: 121))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 6.7 (SD 3.7); n=19, Group 2: mean 19.8 (SD 13.8); n=16; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 24.5 (14.0). Baseline no treatment: 27.7 (11.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI, symptom duration and baseline values of outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up (increased knee pain in 2 people, increased blood pressure in 1); Group 2 Number missing: 2, Reason: 2 lost to follow up (ineffectiveness in 1, other disease in the other)

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 3.2 (SD 1.9); n=19, Group 2: mean 7.2 (SD 5.1); n=16; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 9.9 (4.1). Baseline no treatment: 11.1 (4.3).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI, symptom duration and baseline values of outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up (increased knee pain in 2 people, increased blood pressure in 1); Group 2 Number missing: 2, Reason: 2 lost to follow up (ineffectiveness in 1, other disease in the other)

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Anwer 2014[22]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in Saudi Arabia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 5 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Prediagnosed osteoarthritis of the knee as per the American College of Rheumatology and radiological evidence of primary osteoarthritis of grade 3 or less on the Kellgren Lawrence scale
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with prediagnosed cases of knee osteoarthritis; age between 40-65 years; unilateral or bilateral involvement (in the case of bilateral involvement, the more symptomatic knee was included); pain in and around the knee
Exclusion criteria	Any deformity of the knee, hip or back; any central or peripheral nervous system involvement; had received steroids or intra-articular injection within the previous 3 months; uncooperative patients; people who received physiotherapy treatment in the past 6 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 55.5 (7.3). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 3 or less Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=21) Intervention 1: Exercise - Supervised strength exercise. Strength exercise including isometric quadriceps exercise, straight leg raising and isometric hip adduction exercise. Performed over 5 weeks (5 days/week) with 10 repetitions per set, with 1 set twice a day for the first week, 2 sets twice a day until the 3rd week, then 3 sets twice a day until the 5th week.. Duration 5 weeks. Concurrent medication/care: All people received ultrasound therapy as per the patient's requirement with 1.5 watts/cm² for 7 minutes in continuous mode at the tender point around the knee joint prior to

	exercise. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=21) Intervention 2: No treatment. Usual activity. Duration 5 weeks. Concurrent medication/care: All people received ultrasound therapy as per the patient's requirement with 1.5 watts/cm² for 7 minutes in continuous mode at the tender point around the knee joint prior to exercise. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group project No. RGP-VPP-209)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC (function only) at 5 weeks; Group 1: mean -16.66 (SD 1.09); n=21, Group 2: mean -6.47 (SD 0.13); n=21; WOMAC 0-68 Top=High is poor outcome; Comments: Baseline exercise: 24.71 (3.42). Baseline control: 24.52 (4.43). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMi and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Pain at </=3 months - Actual outcome: Numerical rating scale at 5 weeks; Group 1: mean -4.81 (SD 0.1); n=21, Group 2: mean -1.71 (SD 0.23); n=21; NRS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 6.05 (0.86). Baseline no treatment: 5.95 (1.11). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMi and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3

	months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Avelar 2011[29]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=21)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis in at least one knee in accordance with clinical and radiographic criteria of the American College of Rheumatology with a classification of Kellgren and Lawrence grade 1-4
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age at least 60 years; having been diagnosed with osteoarthritis in at least one knee in accordance with the clinical and radiographic criteria of the American College of Rheumatology with a classification of 1, 2, 3 or 4 according to the grading scale established by Kellgren and Lawrence; not having suffered any recent knee injury; not requiring a walking aid; self-report of not having been submitted to any rehabilitation procedure in the previous 3 months; and not having used glucocorticoids for at least 2 months prior to the study.
Exclusion criteria	Any orthopaedic, neurological, respiratory or acute cardiac disease that would preclude the study; if they had any cognitive deficit as determined by the Mini-Mental Status Examination
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 73.1 (5.0). Gender (M:F): 3:18. Ethnicity: Not stated
Further population details	1. Age: Mixed age group 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity (mean [SD]): Kellgren Lawrence grade 3 (1) Duration of symptoms: not stated.
Indirectness of population	No indirectness
Interventions	(n=11) Intervention 1: Exercise – Other supervised exercise (e.g. aerobic and strength exercise combined). Whole body vibration while performing squat exercises. Squats included 3 seconds of isometric flexion of the quadriceps to 60 degrees and 3 seconds of isometric flexion of the quadriceps to 10 degrees in each repetition (with progressive increasing time and repetitions over the duration of the study - increasing

	from 6 reps of 20 seconds up to 8 reps of 40 seconds). Whole body vibration was starting at a frequency of 35Hz-40Hz, amplitude of 4mm, and acceleration that ranged from 2.78G to 3.26G.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not stated / Unclear 2. Group or individual : Individual session 3. Type of exercise: Whole body vibration (and strengthening). (n=10) Intervention 2: Exercise - Supervised strength exercise. Squat exercises only. Squats included 3 seconds of isometric flexion of the quadriceps to 60 degrees and 3 seconds of isometric flexion of the quadriceps to 10 degrees in each repetition (with progressive increasing time and repetitions over the duration of the study - increasing from 6 reps of 20 seconds up to 8 reps of 40 seconds).. Duration 12 weeks. Concurrent medication/care: No additional treatment. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Study funded by industry (This study was supported by FAPEMIG, CNPq e CAPES)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 718 (SD 94); n=11, Group 2: mean 777 (SD 130); n=10; WOMAC physical function subscale 0-1700 Top=High is poor outcome; Comments: Baseline mixed exercise: 970 (96). Baseline strength exercise: 993 (113).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, classification, unilateral/bilateral symptoms, gender, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 dropped out; Group 2 Number missing: 1, Reason: 1 dropped out

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 189 (SD 29); n=11, Group 2: mean 165 (SD 32); n=10; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline mixed exercise: 298 (32). Baseline strength exercise: 165 (32).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, classification, unilateral/bilateral symptoms, gender, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 dropped out; Group 2 Number missing: 1, Reason: 1 dropped out

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Bautch 1997[35]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=34)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Prior diagnosis of osteoarthritis and meeting the American College of Rheumatology clinical and radiographic criteria for primary osteoarthritis of the knee
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee who were at least 58 tears of age, lived independently in the community, were without physical or medical problems for which participation in the exercise program would be contraindicated and were not currently enrolled in a regular exercise program
Exclusion criteria	Receiving intra-articular or systemic steroids within the past 2 years; routine use of any medications with known potential for altering cartilage metabolism
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 67.8 (3.0). Gender (M:F): Unclear, reports only a small number of the population, 3:8. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=17) Intervention 1: Exercise - Supervised strength exercise. 1 hour exercise session 3 times a week involving a walking component (low intensity) with range of motion exercises of the trunk and upper and lower extremities as 3 repetitions of each maneuver, which increased over 4 weeks to 10 repetitions.. Duration 12 weeks. Concurrent medication/care: Both groups received a weekly educational program with content related to health, exercise and arthritis. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group

	session 3. Type of exercise: Not applicable (n=17) Intervention 2: No treatment. No additional information. Duration 12 weeks. Concurrent medication/care: Both groups received a weekly educational program with content related to health, exercise and arthritis. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: AIMS at 12 weeks; Group 1: mean 23.37 (SE 2.48); n=15, Group 2: mean 17.88 (SE 1.85); n=15; AIMS Unclear Top=High is poor outcome; Comments: Baseline exercise: 26.19 (2.01). Baseline control: 21.37 (2.11). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcome is different at baseline (to the point where the change seems minimal, but one has a decrease in pain, one had an increase).; Group 1 Number missing: 2, Reason: No information given; Group 2 Number missing: 2, Reason: No information given Protocol outcome 2: Pain at </=3 months - Actual outcome: Pain (VAS) now at 12 weeks; Group 1: mean 2.19 (SE 0.43); n=15, Group 2: mean 2.08 (SE 0.54); n=15; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 3.49 (0.53). Baseline no treatment: 1.46 (0.43). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcome is different at baseline (to the point where the change seems minimal, but one has a decrease in pain, one had an increase).; Group 1 Number missing: 2, Reason: No information given; Group 2 Number missing: 2, Reason: No information given	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Beckwee 2017[38]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=37)
Countries and setting	Conducted in Belgium; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who fulfilled the criteria defined by the American College of Rheumatology for knee osteoarthritis.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Community-dwelling people aged 50 or older with a painful knee in the last 30 days and radiographic tibiofemoral osteoarthritis
Exclusion criteria	Inability to come to the hospital for assessments and threapy; intra-articular steroid injections in the previous six months; a (systemic) arthritis condition other than osteoarthritis; contraindications for physical exercises; an unstable medical condition.
Recruitment/selection of patients	Recruited through advertisements (posters and local media)
Age, gender and ethnicity	Age - Mean (SD): 61.8 (8.88). Gender (M:F): 17:18. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-4, median grade 1-2 (25) Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=19) Intervention 1: Exercise - Supervised strength exercise. Exercise three times weekly. A total of 54 training sessions over 18 weeks. 18 sessions were supervised at the university hospital and 36 sessions were unsupervised at the participants' homes (included as supervised as this contains supervised sessions rather than purely unsupervised). The strength training sessions lasted 45 minutes each and consisted in seven exercises that focused on strength and functional performance of knee extensors, hamstring, hip abductor and hip adductor muscles. Duration 18 weeks. Concurrent medication/care: Not stated/unclear. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable

	(n=18) Intervention 2: Exercise - Supervised aerobic exercise. Exercise three times weekly. A total of 54 training sessions over 18 weeks. 18 sessions were supervised at the university hospital and 36 sessions were unsupervised at the participants' homes (included as supervised as this contains supervised sessions rather than purely unsupervised). The walking therapy consisted of walking for 40 minutes at an intensity of 14 to 17 on a Borg scale to achieve a heart frequency of 50-80%. Duration 18 weeks. Concurrent medication/care: Not stated/unclear. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Fundingfor this study was obtained from the grant Wetenschappelijk Fonds Willy Gepts of the UZ Brussels which was not involved in the study design and in the collection, analysis and interpretation of the data.
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus SUPERVISED AEROBIC EXERCISE Protocol outcome 1: Pain at >3 months - Actual outcome: Pain (Intermittent and constant osteoarthritis pain total pain) at 18 weeks; Group 1: mean 7.93 (SD 9.77); n=14, Group 2: mean 10.1 (SD 7.27); n=16; ICOPTP 0-20 Top=High is poor outcome; Comments: Values calculated by combining bone marrow lesion and no bone marrow lesion groups. Reported strength exercise bone marrow lesion (n=10): 4.1 (5.15). Reported strength exercise no bone marrow lesion (n=4): 17.5 (11.82). Reported aerobic exercise bone marrow lesion (n=10): 11.0 (8.41). Reported aerobic exercise no bone marrow lesion (n=6): 8.5 (4.37). Baseline strength exercise: 13.1 (9.16). Baseline aerobic exercise: 13.8 (5.94). Risk of bias: All domain - Very high, Selection – High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Overall similar, some differences between bone marrow lesion and no bone marrow lesion group.; Group 1 Number missing: 4, Reason: 1 fall between assessment and start therapy so did not receive intervention, 1 increase of knee pain, related to exercise, 1 no reason, 1 cardiovascular problems, not related to exercise; Group 2 Number missing: 2, Reason: 1 increase in knee pain, not related to exercise, 1 lack of time.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Bennell 2010[45] (Bennell 2007[44])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=89)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis in at least one knee according to the American College of Rheumatology classification criteria with radiographic verification
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Osteoarthritis in at least one knee fulfilling the American College of Rheumatology classification criteria reporting average knee pain on walking >3 on an 11-point scale. To ensure medial tibiofemoral osteoarthritis and varus malalignment, inclusion criteria were medial knee pain and medial compartment osteophytes on medial joint space narrowing and knee alignment no more than 182 degrees on a standardised semiflexed posteroanterior X-ray.
Exclusion criteria	No or doubtful radiographic osteoarthriti; knee surgery or intra-articular corticosteroid injection within 6 months, current or past (within 4 weeks) oral corticosteroid use; systemic inflammatory arthritic conditions; a history of hip or knee joint replacement or tibial osteotomy; intention to start or currently participating in a supervised lower limb strengthening program; body mass index >35 due to difficulty in accurate marker placement for gait analysis; a medial condition that precluded safe participation in an exercise program or unable to ambulate without a gait aid
Recruitment/selection of patients	People were recruited from the community through advertisements in newspapers and local clubs and from our database of research volunteers
Age, gender and ethnicity	Age - Mean (SD): 64.6 (8.4). Gender (M:F): 46:43. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren and Lawrence grades 2-4 Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=45) Intervention 1: Exercise - Unsupervised strength exercise. Home-based hip strengthening exercise five times a week for 12 weeks. Six exercises to strengthen hip abductor and adductor muscles were performed in a sidelying and standing position (three sets of 10 repetitions) with ankle cuff weights or elastic bands. Additionally, people attended a physiotherapy clinic on seven occasions to received appropriate instruction. These sessions lasted 30 minutes initially and 15 minutes subsequently. The exercise intensity was adjusted from 10 repetitions dependent on the participant's ability to complete the activity.. Duration 12 weeks. Concurrent medication/care: Participants were asked to refrain from seeking other forms of treatment during the trial. However, due to ethical considerations, analgesia and non-steroidal anti-inflammatory drugs were permitted as required. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=44) Intervention 2: No treatment. No additional exercise treatment. Duration 12 weeks. Concurrent medication/care: Participants were asked to refrain from seeking other forms of treatment during the trial. However, due to ethical considerations, analgesia and non-steroidal anti-inflammatory drugs were permitted as required. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The trial was funded by the National Health and Medical Research Council (Project #454686).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months

- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean -8.07 (SD 7.7); n=45, Group 2: mean -1.9 (SD 7.7); n=44; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 24.8 (10.9). Baseline no treatment: 23.7 (11.8).

Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, knee affected, unilateral symptoms, BMI, static knee alignment, Kellgren Lawrence grade, activity levels, current medication use, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost. 2 due to unrelated medical condition, 1 unable to contact, 1 personal reasons, 1 increased knee pain unrelated to intervention, 1 adverse event/discontinued intervention. 2 more discontinued intervention, but attended the week 13 assessment.; Group 2 Number missing: 7, Reason: 7 lost. 2 unable to contact, 1 unrelated medical condition, 1 personal reasons, 1 relocated, 1 declined to attend, 1 contralateral arthroscope.

Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean -2.6 (SD 2.6); n=45, Group 2: mean -0.48 (SD 2.7); n=44; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 7.7 (3.0). Baseline no treatment: 6.9 (3.3). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, knee affected, unilateral symptoms, BMI, static knee alignment, Kellgren Lawrence grade, activity levels, current medication use, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost. 2 due to unrelated medical condition, 1 unable to contact, 1 personal reasons, 1 increased knee pain unrelated to intervention, 1 adverse event/discontinued intervention. 2 more discontinued intervention, but attended the week 13 assessment.; Group 2 Number missing: 7, Reason: 7 lost. 2 unable to contact, 1 unrelated medical condition, 1 personal reasons, 1 relocated, 1 declined to attend, 1 contralateral arthroscope. Protocol outcome 3: Serious adverse events at </=3 months - Actual outcome: Adverse events at 12 weeks; Group 1: 5/45, Group 2: 0/44; Comments: Exercise: three reported back pain, one reported back and hip pain, one reported aggravated varicose veins and knee pain for which acupuncture treatment was sought Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, knee affected, unilateral symptoms, BMI, static knee alignment, Kellgren Lawrence grade, activity levels, current medication use, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost. 2 due to unrelated medical condition, 1 unable to contact, 1 personal reasons, 1 increased knee pain unrelated to intervention, 1 adverse event/discontinued intervention. 2 more discontinued intervention, but attended the week 13 assessment.; Group 2 Number missing: 7, Reason: 7 lost. 2 unable to contact, 1 unrelated medical condition, 1 personal reasons, 1 relocated, 1 declined to attend, 1 contralateral arthroscope.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Bennell 2014[46]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee pain with radiographic medial tibiofemoral joint osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Average knee pain over the last week of at least 25 on an 100mm VAS; pain/tenderness predominantly over the medial knee region; radiographic medial tibiofemoral joint osteoarthritis; weight bearing postero-anterior radiograph showing: Kellgren-Lawrence at least grade 2 changes; anatomical axis angle of <181 degrees for females and <183 degrees for males, indicating varus alignment based on mechanical axis values using gender-specific regression equations; medial tibiofemoral joint narrowing grade > lateral tibiofemoral joint narrowing grade; medial compartment osteophyte grade greater than or equal to lateral compartment osteophyte grade
Exclusion criteria	Knee surgery or intra-articular corticosteroid injection within 6 months; current or past (within 4 weeks) oral corticosteroid use; systemic arthritic conditions; prior hip or knee joint replacement or tibial osteotomy surgery; other non-pharmacological treatment within past 6 months; body mass index above 36 kg/m²
Recruitment/selection of patients	People were recruited via advertisements
Age, gender and ethnicity	Age - Mean (SD): 62.5 (7.4). Gender (M:F): 48:52. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: kellgren Lawrence grade no less than 2 Duration of symptoms (median [IQR]): neuromuscular = 60.0 (96.0) months, strength = 84.0 (93.6) months.
Indirectness of population	No indirectness

Interventions	(n=50) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Neuromuscular exercises performed aiming to improve position of the trunk and lower limb joints while dynamically and functionally strengthening the lower limb. 14 visits over 12 weeks lasting 30-40 minutes. All participants were asked to perform home exercises three times per week.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Neuromodulatory (n=50) Intervention 2: Exercise - Supervised strength exercise. Strengthening exercises performed aiming to strengthen the quadriceps with non-weight bearing exercises. The dosage was set to 2-3 sets of 10 repetitions. 14 visits over 12 weeks lasting 30-40 minutes. All participants were asked to perform home exercises three times per week.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This trial was funded by the National Health and Medical Research Council Fellowship)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Assessment of Quality of Life Instrument Version Two at 13 weeks; Group 1: mean 0.78 (SD 0.14); n=38, Group 2: mean 0.78 (SD 0.16); n=44; Assessment of Quality of Life Instrument Version Two -0.04-1.00 Top=High is good outcome; Comments: Baseline neuromuscular: 0.73 (0.14). Baseline strength: 0.73 (0.18).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, symptom duration, height, body mass, BMI, gender, affected knee, unilateral symptoms, dominant side affected, knee alignment, radiographic disease severity, drug use and baseline values of outcomes; Group 1 Number missing: 12, Reason: Unclear. 7 due to increased pain or unanticipated decision to undergo total joint replacement.; Group 2 Number missing: 6, Reason: Unclear. 1 due to increased pain or unanticipated decision to undergo total joint replacement.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 13 weeks; Group 1: mean 18.3 (SD 9.6); n=38, Group 2: mean 20.1 (SD 9.8); n=44; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline neuromuscular: 26.0 (9.1). Baseline strength: 28.2 (9.9).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, symptom duration, height, body mass, BMI, gender, affected knee, unilateral symptoms, dominant side affected, knee alignment, radiographic disease severity, drug use and baseline values of outcomes; Group 1 Number missing: 12, Reason: Unclear. 7 due to increased pain or unanticipated decision to undergo total joint replacement.; Group 2 Number missing: 6, Reason: Unclear. 1 due to increased pain or unanticipated decision to undergo total joint replacement. Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 13 weeks; Group 1: mean 6.4 (SD 3.1); n=38, Group 2: mean 6.4 (SD 2.9); n=44; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline neuromuscular: 8.1 (2.2). Baseline strength: 8.8 (3.3). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, symptom duration, height, body mass, BMI, gender, affected knee, unilateral symptoms, dominant side affected, knee alignment, radiographic disease severity, drug use and baseline values of outcomes; Group 1 Number missing: 12, Reason: Unclear. 7 due to increased pain or unanticipated decision to undergo total joint replacement.; Group 2 Number missing: 6, Reason: Unclear. 1 due to increased pain or unanticipated decision to undergo total joint replacement. Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Adverse events at 13 weeks; Group 1: 13/46, Group 2: 10/44; Comments: Neuromuscular exercise: 10 increased knee pain, 1 back pain, 2 pain in other area, 2 hip pain, 3 swelling/inflammation, 1 stiffness. Strength exercise: 8 increased knee pain, 1 back pain, 1 pain in other area, 1 hip pain, 1 swelling/inflammation. Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, symptom duration, height, body mass, BMI, gender, affected knee, unilateral symptoms, dominant side affected, knee alignment, radiographic disease severity, drug use and baseline values of outcomes; Group 1 Number missing: 12, Reason: Unclear. 7 due to increased pain or unanticipated decision to undergo total joint replacement.; Group 2 Number missing: 6, Reason: Unclear. 1 due to increased pain or unanticipated decision to undergo total joint replacement.	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Bieler 2017[52]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=152)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical hip osteoarthritis according to the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Home dwelling 60+ year old individuals with clinical hip osteoarthritis according to the American College of Rheumatology who were not on a waiting list for hip replacement
Exclusion criteria	Symptomatic osteoarthritis of the knee or the big toe; other types of arthrtis; previous hip or knee replacement; previous hip fracture; comorbidity that prevented exercising; treatment related to hip problems within the last 3 months; inability to use public transportation; performing regular exercise/sports twice or more weekly
Recruitment/selection of patients	Primarily recruited through general practitioners and specialists and advertisements in local newspapers in Greater Copenhagen
Age, gender and ethnicity	Age - Mean (SD): 69.6 (6.1). Gender (M:F): 49:103. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (Cardiovascular disease: 74, Lung disease: 13, Metabolic disease: 12, Prior cancer: 23). 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity (mean [SD]): Kellgren Lawrence grade 2.1 (1.5) Duration of symptoms (mean [SD]): 6.1 (6.3) years.
Indirectness of population	No indirectness
Interventions	(n=50) Intervention 1: Exercise - Supervised strength exercise. Machine based strength training - three mandatory resistance exercises in machines (leg press, seated knee extension, hip extension in a standing position leaned forward 45 degrees with trunk and pelvis resting against an abdominal platform support.. Duration 12 months. Concurrent medication/care: People in the strength exercise and Nordic Walking exercise groups also received individual counseling, a one hour patient education session on the important of exercise and some telephone assisted

	counseling to improve adherence.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=50) Intervention 2: Exercise - Supervised aerobic exercise . Nordic walking with progressive intensity for 1 hour three times weekly as a group session in a local park. Duration 12 months. Concurrent medication/care: People in the strength exercise and Nordic Walking exercise groups also received individual counseling, a one hour patient education session on the important of exercise and some telephone assisted counseling to improve adherence.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=52) Intervention 3: Exercise - Unsupervised strength exercise. Home based strength exercises, including hip range of motion, stretching and strengthening exercises including a chair stand exercise, pelvic lift, isometric hip flexion exercise in the standing position, and gluteus medius muscle exercise in the side lying position. This was progressed with elastic bands as resistance.. Duration 12 months. Concurrent medication/care: No concomitant treatment. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable Comments: This group was not included in the comparison as the difference in concomitant treatment meant that this group was not comparable to the others.
Funding	Academic or government funding (This work was supported by the TrygFonden (1190-09), Nordea foundation (Healthy Ageing grant), Health Foundation (2009B097), Danish Rheumatism Association (R56-Rp2380), Lundbeck Foundation (FP50/2009), School of Physical Therapy in Coperhagen, and The Association of Danish Physiotherapists Research Fund)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus SUPERVISED AEROBIC EXERCISE

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: Arthritis Self-efficacy scale function subscale at 12 months; MD; 7.6 (95%CI 0.7 to 14.4) (P value: 0.0307) Arthritis self efficacy function subscale 0-100 Top=High is poor outcome, Comments: Baseline strength: 86.6 (15.6). Baseline aerobic: 85.0 (14.1). (Originally reported for aerobic vs. strength, but reported as improvement. Direction of change has been flipped).;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, body weight, height, BMI, retired/working, education, home living status, unilateral/bilateral, radiographic score, duration of symptoms, comorbidity, baseline values of outcomes; Group 1 Number missing: 8, Reason: 1 not interested, 1 severe illness, 5 surgery, 2 not interested, 1 severe illness; Group 2 Number missing: 21, Reason: 8 not interested, 2 trauma/illness, 1 no benefit, 1 sick spouse, 2 increased pain, 1 too hard, 1 the weather, 1 severe illness, 3 surgery, 1 personal reason Protocol outcome 2: Pain at > 3 months - Actual outcome: Arthritis Self-efficacy scale pain subscale at 12 months; MD; 11.1 (95%CI 0.1 to 22.2) (P-value: 0.0471) Arthritis Self-efficacy pain subscale 0-100 Top=High is poor outcome, Comments: Baseline strength: 67.8 (19.5). Baseline aerobic: 63.4 (17.9). (Originally reported for aerobic vs. strength, but reported as improvement. Direction of change has been flipped).; Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, body weight, height, BMI, retired/working, education, home living status, unilateral/bilateral, radiographic score, duration of symptoms, comorbidity, baseline values of outcomes; Group 1 Number missing: 8, Reason: 1 not interested, 1 severe illness, 5 surgery, 2 not interested, 1 severe illness; Group 2 Number missing: 21, Reason: 8 not interested, 2 trauma/illness, 1 no benefit, 1 sick spouse, 2 increased pain, 1 too hard, 1 the weather, 1 severe illness, 3 surgery, 1 personal reason	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Bokaeian 2021[56]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=59)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks of intervention, 1 month additional follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with unilateral or bilateral knee osteoarthritis diagnosed radiographically and by symptoms
Stratum	Overall
Subgroup analysis within study	Not applicable

Inclusion criteria	45-76 years of age; knee pain of 30 or greater on the 100-mm visual analog scale, unilateral or bilateral tibiofemoral joint osteoarthritis of grade 2-3 based on the Kellgren-Lawrence grading system; a history of pain for more than a month; and ability to walk without assistive devices.
Exclusion criteria	Systemic arthritis;diabetes; neuromuscular diseases; injection in the lower-extremity joints within the last 6 months; hip or knee replacement; symptomatic hip osteoarthritis; recent trauma to the knee joint; extreme physical weakness; a body mass index >35; a history of lower-extremity surgery in the last 6 months; people with a history of taking oral corticosteroids and physiotherapy within the past 3 months.
Recruitment/selection of patients	People were recruited from local outpatient rheumatology and orthopedic clinics
Age, gender and ethnicity	Age - Mean (SD): 56.1 (5.0). Gender (M:F): 14:45. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not applicable 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grades 2-3 Duration of symptoms: history of pain for > one month IRCT201702222793N4
Indirectness of population	No indirectness
Interventions	(n=22) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). YogaMT group who practiced MT gait for 20 minutes and performed three repetitions of goddess squat and warrior lunge exercises with a minimum rest interval of 40s, supervised by a physiotherapist. For the MT gait, they were trained to walk with slight hip internal rotation and knee flexion (about 20 degrees) at their selected speed on the treadmill. When necessary, they received verbal feedback during training. People wore a pair of comfortable shoes during the training and treatment sessions. The difficulty was adjusted according to the Borg Perceived Exertion Scale. . Duration 4 weeks. Concurrent medication/care: People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification is an essential part of knee osteoarthritis management (and so was offered).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (n=19) Intervention 2: Exercise - Supervised strength exercise. The KMS group that consisted of three 10-repetition sets of resistive knee extension/flexion exercises

	with 2-min interval, using the quadriceps chair. The maximum load that each participant could lift to complete 10-repetition maximum without pain was determined to adjust the amount of load for each exercise weekly. Exercise was performed under supervision of a physiotherapist. If the participant reported pain during exercise they were instructed to perform the exercise only in a pain-free range, and if the pain persisted, the resistive load was reduced.. Duration 4 weeks. Concurrent medication/care: People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification is an essential part of knee osteoarthritis management (and so was offered).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=18) Intervention 3: Exercise - Supervised aerobic exercise . The TMW group that practiced walking on a treadmill at their self-selected speed for 20 minutes, with no gait modification, under the supervision of the physiotherapist. They wore comfortable pairs of shoes during the training and treatment sessions. The treadmill used in this group was the same as the YogaMT group treadmill.. Duration 4 weeks. Concurrent medication/care: People in all groups also received thermotherapy with a hot pack for 20 minutes. They were also instructed to avoid sitting in a cross-legged position, kneeling, prolonged standing, and stair climbing. Patient education on activity and lifestyle modification is an essential part of knee osteoarthritis management (and so was offered).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This work was supported by the Ahvaz Jundishapur University of Medical Sciences (pht-9605))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analog scale (VAS) at 2 months; Group 1: mean 39.8 (SD 36); n=22, Group 2: mean 44.4 (SD 24.6); n=19; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline other supervised exercise: 78.1 (18.4). Baseline supervised strength exercise: 69.3 (13.7).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Comparable for age, sex, BMI, VAS and other

outcomes; Group 1 Number missing: 2, Reason: Other supervised exercise: 2 lost to follow up (tight work schedule).; Group 2 Number missing: 1, Reason: Supervised strength exercise: 1 lost to follow up (tight work schedule). RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED AEROBIC EXERCISE Protocol outcome 1: Pain at </=3 months - Actual outcome: Visual analog scale (VAS) at 2 months; Group 1: mean 39.8 (SD 36); n=22, Group 2: mean 60.3 (SD 26.9); n=18; Visual analog scale 0-100 Top=High is poor outcome; Comments: Baseline other supervised exercise: 78.1 (18.4). Baseline supervised aerobic exercise: 78.9 (16.9). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Comparable for age, sex, BMI, VAS and other outcomes; Group 1 Number missing: 2, Reason: Other supervised exercise: 2 lost to follow up (tight work schedule).; Group 2 Number missing: 0 RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED AEROBIC EXERCISE versus SUPERVISED STRENGTH EXERCISE Protocol outcome 1: Pain at </=3 months - Actual outcome: Visual analog scale (VAS) at 2 months; Group 1: mean 60.3 (SD 26.9); n=18, Group 2: mean 44.4 (SD 24.6); n=19; Visual analog scale 0-100 Top=High is poor outcome; Comments: Baseline supervised strength exercise: 69.3 (13.7). Baseline supervised aerobic exercise: 78.9 (16.9). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Comparable for age, sex, BMI, VAS and other outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: Supervised strength exercise: 1 lost to follow up (tight work schedule).	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Borjesson 1996[57]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=68)
Countries and setting	Conducted in Sweden; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 5 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Medial knee osteoarthrosis grade 1-3 according to the classification based on weight-bearing radiographs (Ahlback, 1968)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People, aged between 55-70 years, with medial knee osteoarthrosis grade 1-3 according to the classification based on weight-bearing radiographs with symptoms for 3-10 years scheduled for surgery, either a high tibial osteotomy or prosthetic replacement for symptoms from their osteoarthrotic knee. The symptoms were unilateral.
Exclusion criteria	No symptoms in the hip or ankle joints.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 64 (4.5). Gender (M:F): 34:34. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Ahlback Osteoarthrosis grade 1-3, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=34) Intervention 1: Exercise - Supervised strength exercise. Physiotherapy conducted as outpatients. The training was aiming at increasing the strength and range of motion in the involved knee, as well as the strength of the whole leg. Exercises were undertaken three times a week for 5 weeks, 15 times altogether. They were also instructed to perform the same exercises at home twice a week. The exercises included: warming up for 10 minutes, knee extension, knee flexion, standing on heel and toes, flexion of the involved knee, hamstrings muscle stretch, hip abduction, hip extension, passive knee extension. Exercises were performed with 2x10 repetitions and each exercise was performed with 10 seconds of isometric hold.

	Resistance was increased according to the 10 RM principle. 3 kilograms was chosen as the maximum weight. The program took 40 minutes to complete.. Duration 5 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=34) Intervention 2: No treatment. No intervention. Duration 5 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Pain at </=3 months - Actual outcome: Pain on walking (NRS) at 5 weeks; Group 1: mean 3 (SD 1.5); n=34, Group 2: mean 3.3 (SD 1.5); n=34; NRS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 3.4 (2.0). Baseline control: 3.3 (1.5). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Randomisation by pulling names out of a hat; Indirectness of outcome: No indirectness ; Baseline details: Reports age, body weight, height, gender, osteoarthrosis grade and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Bossen 2013[59] (Bossen 2013[58])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=199)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Partially adequate method of assessment/diagnosis: Self-reported knee and/or hip osteoarthritis - defined by if they had a painful knee or hip joint and if a doctor or other health care provider had ever told them this was a result of osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 50-75 years; self-reported osteoarthritis in knee and/or hip; self-reported inactivity (<30 minutes of moderate physical activity three or five times or less per week); no face-to-face consultation for osteoarthritis with a health care provider, other than GP in the last 6 months; ability to access the Internet weekly; no contra-indications to exercise without supervision
Exclusion criteria	Not fulfilling the eligibility criteria
Recruitment/selection of patients	People were recruited through advertisements in Dutch newspapers and online on health-related websites
Age, gender and ethnicity	Age - Mean (SD): 62.0 (5.7). Gender (M:F): 70:129. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Low morbidity score (Majority had no comorbidities (125) while 35 had one comorbidity and 39 had two comorbidities.). 4. Site of osteoarthritis: Mixed (Knee and/or hip osteoarthritis).
Extra comments	Severity: Not stated Duration of symptoms: <1 to >7 years - median >3-7 years
Indirectness of population	No indirectness
Interventions	(n=100) Intervention 1: Exercise - Unsupervised aerobic exercise . Joint2move website based exercise program taking activities that a person enjoys (for example: cycling, walking, gardening) and making goals with that in order to make it into a stable physical activity. Activities are stepped up weekly. The website also provided information about osteoarthritis and lifestyle choices.. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=99) Intervention 2: No treatment. Waiting list control. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS quality of life at 3 months; Group 1: mean 49.4 (SD 36); n=85, Group 2: mean 47.3 (SD 35.8); n=80; KOOS quality of life 0-100 Top=High is good outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 49.4 (41.7-57.0). Reported control: 47.3 (39.4-55.1). Baseline exercise: 38 (30.6-45.5), calculated SD (38.0). Baseline control: 40.9 (33.6-48.2), calculated SD (36.9).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 15, Reason: 15 lost to follow up; Group 2 Number missing: 19, Reason: 19 lost to follow up

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: KOOS quality of life at 12 months; Group 1: mean 48.7 (SD 34.9); n=75, Group 2: mean 47.5 (SD 35); n=71; KOOS quality of life 0-100 Top=High is good outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 48.7 (40.8-56.6). Reported control: 47.5 (39.3-55.6). Baseline exercise: 38 (30.6-45.5), calculated SD (38.0). Baseline control: 40.9 (33.6-48.2), calculated SD (36.9).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 25, Reason: 25 lost to follow up; Group 2 Number missing: 28, Reason: 28 lost to follow up

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: KOOS physical functioning at 3 months; Group 1: mean 67.8 (SD 40.2); n=84, Group 2: mean 61.3 (SD 39.2); n=80; KOOS physical functioning 0-100 Top=High is good outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 67.8 (59.2-76.4). Reported control: 61.3 (52.7-69.9). Baseline exercise: 58.8 (51.5-66.0), calculated SD (36.8). Baseline control: 55.2 (47.9-62.5), calculated SD (36.9).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 16, Reason: 165 lost to follow up; Group 2 Number missing: 19, Reason: 19 lost to follow up

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: KOOS physical functioning at 12 months; Group 1: mean 67.9 (SD 38.9); n=75, Group 2: mean 62.9 (SD 38.1); n=72; KOOS physical functioning 0-100 Top=High is good outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 67.9 (59.1-76.7). Reported control: 62.9 (54.1-71.7). Baseline exercise: 58.8 (51.5-66.0), calculated SD (36.8). Baseline control: 55.2 (47.9-62.5), calculated SD (36.9).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 25, Reason: 25 lost to follow up; Group 2 Number missing: 27, Reason: 27 lost to follow up

Protocol outcome 5: Pain at </=3 months
- Actual outcome: VAS pain at 3 months; Group 1: mean 3.5 (SD 4.9); n=85, Group 2: mean 4.5 (SD 5.3); n=81; VAS 0-10 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 3.5 (2.5-4.6). Reported control: 4.5 (3.4-5.7). Baseline exercise: 5.4 (4.2-6.5), calculated SD (5.9). Baseline control: 4.9 (3.7-6.1), calculated SD (6.1).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 15, Reason: 15 lost to follow up; Group 2 Number missing: 18, Reason: 18 lost to follow up

Protocol outcome 6: Pain at > 3 months
- Actual outcome: VAS pain at 12 months; Group 1: mean 3.5 (SD 4.7); n=76, Group 2: mean 3.8 (SD 4.7); n=71; VAS 0-10 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 3.5 (2.4-4.5). Reported control: 3.8 (2.7-4.9). Baseline exercise: 5.4 (4.2-6.5), calculated SD (5.9). Baseline control: 4.9 (3.7-6.1), calculated SD (6.1).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 24, Reason: 24 lost to follow up; Group 2 Number missing: 28, Reason: 28 lost to follow up

Protocol outcome 7: Psychological distress at </=3 months
- Actual outcome: HADS anxiety subscale at 3 months; Group 1: mean 3.5 (SD 4.7); n=85, Group 2: mean 4.2 (SD 4.8); n=79; HADS anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 3.5 (2.5-4.5). Reported control: 4.2 (3.1-5.2). Baseline exercise: 4 (2.5-5.6), calculated SD (7.9). Baseline control: 4.2 (2.6-5.9), calculated SD (8.4).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 15, Reason: 15 lost to follow up; Group 2 Number missing: 20, Reason: 20 lost to follow up
- Actual outcome: HADS depression subscale at 3 months; Group 1: mean 2.6 (SD 5.2); n=85, Group 2: mean 3.2 (SD 5); n=79; HADS depression subscale 0-21 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 2.6 (1.5-3.7). Reported control: 3.2 (2.1-4.3).

Baseline exercise: 4 (2.5-5.6), calculated SD (7.9). Baseline control: 4.2 (2.6-5.9), calculated SD (8.3). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 15, Reason: 15 lost to follow up; Group 2 Number missing: 21, Reason: 21 lost to follow up Protocol outcome 8: Psychological distress at > 3 months - Actual outcome: HADS anxiety subscale at 12 months; Group 1: mean 3.1 (SD 5.1); n=75, Group 2: mean 4.1 (SD 5); n=72; HADS anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 3.1 (2.0-4.3). Reported control: 4.1 (2.9-5.2). Baseline exercise: 4 (2.5-5.6), calculated SD (7.9). Baseline control: 4.2 (2.6-5.9), calculated SD (8.4). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 25, Reason: 25 lost to follow up; Group 2 Number missing: 27, Reason: 27 lost to follow up - Actual outcome: HADS depression subscale at 12 months; Group 1: mean 2.4 (SD 5.1); n=75, Group 2: mean 3 (SD 5); n=72; HADS depression subscale 0-21 Top=High is poor outcome; Comments: Reports mean final values and 95% CIs. Reported exercise: 2.4 (1.3-3.6). Reported control: 3 (1.9-4.2). Baseline exercise: 4 (2.5-5.6), calculated SD (7.9). Baseline control: 4.2 (2.6-5.9), calculated SD (8.3). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, location of osteoarthritis, duration of symptoms, education, comorbidity and baseline values of outcomes; Blinding details: Caregiver was a computer, so not blind but not able to be influenced.; Group 1 Number missing: 25, Reason: 25 lost to follow up; Group 2 Number missing: 27, Reason: 27 lost to follow up	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Brosseau 2012[66]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=222)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Confirmed diagnosis with mild to moderate unilateral or bilateral osteoarthritis according to the American College of Rheumatology clinical and radiographic/magnetic resonance imaging criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a confirmed diagnosis of mild to moderate unilateral or bilateral osteoarthritis according to the American College of Rheumatology clinical and radiographic/magnetic resonance imagery criteria; reported pain for at least 3 months; expected their medication to change during the study period; demonstrated an ability to ambulate for a minimum of 20 minutes, at their own pace with minimal reports of pain (no less than 3 out of 10 on a visual analogue pain rating scores); were able to be treated as an outpatient; were available three times a week over a period of 12 months
Exclusion criteria	Participated in regular physical or aerobic sports at least 2 times per week for more than 20 minutes per session during the previous 6 months; severe osteoarthritis of the knee or other weight bearing joints of the lower extremity; no written consent from their physician to participate in the study; pain at rest or at night; received rehabilitation treatment, corticosteroids injection, or any other pain-related treatment besides medication for arthritis within the last 12 months; uncontrolled hypertension (systolic blood pressure >160mmHg confirmed by the screening initial VO2 max test at the Ottawa Heart Institute); other illnesses, such as rheumatoid arthritis (judged by the patient or study physician to make participation in this study inadvisable); significant cognitive deficit resulting in an inability to understand or comply with instructions; surgery planned in the next year; intention to move away from Ottawa region in the next year; an inability to communicate in English or French; an unwillingness to sign informed consent
Recruitment/selection of patients	No additional information

Age, gender and ethnicity	Age - Mean (SD): 63.4 (8.6). Gender (M:F): 69:153. Ethnicity: White = 197, Black = 5, Hispanic = 8. Asian or Pacific Islander = 10, American Indian or Alaskan native = 1, Other = 1
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Mild to moderate Duration of osteoarthritis (mean [SD]): 10.3 (9.26).
Indirectness of population	No indirectness
Interventions	(n=79) Intervention 1: Exercise - Supervised aerobic exercise . Three weekly walking sessions over a 12 month period. Every walking session began with a 10-minute warm-up, before engaging in the 45 minute aerobic walking phase. This ended with a 10 minute cool down. The aim was to achieve an intensity of 50-75% based on the subjects' pre-determined maximum heart rate. Divided into two stages: a progressive aerobic phase and a maintenance aerobic phase.. Duration 12 months (with an additional 6 months follow up). Concurrent medication/care: Everyone received an educational pamphlet describing the benefits of walking, and a pedometer. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=74) Intervention 2: No treatment. No walking intervention. Duration 12 months (with an additional 6 months follow up). Concurrent medication/care: Everyone received an educational pamphlet describing the benefits of walking, and a pedometer. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=69) Intervention 3: Other. Walking program and a behavioral intervention. Duration 12 months (with an additional 6 months follow up). Concurrent medication/care: Everyone received an educational pamphlet describing the benefits of walking, and a pedometer. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group was not included in the final analysis as it did not meet the criteria for this review

Funding	Academic or government funding (This study was completed with the support of a research grant obtained from the Canadian Institutes of Health Research (CIHR) (Grant#MCT82367); University Research Chair (salary support for research staff) and the Ministry of Human Resources (summer student program) (Canada).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: SF-36 physical component summary at 18 months; Group 1: mean 42.82 (SD 9.24); n=44, Group 2: mean 45.149 (SD 8.93); n=36; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 40.516 (8.598). Baseline no treatment: 41.996 (9.656).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, affected knee, duration of osteoarthritis, weight, BMI, walking aid use, racial background, marital status, level of education, medication use and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 dropped out by 18 months; Group 2 Number missing: 38, Reason: 38 dropped out by 18 months
- Actual outcome: SF-36 mental component summary at 18 months; Group 1: mean 51.993 (SD 11); n=44, Group 2: mean 53.101 (SD 9.914); n=36; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 52.914 (10.845). Baseline no treatment: 53.556 (8.995).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, affected knee, duration of osteoarthritis, weight, BMI, walking aid use, racial background, marital status, level of education, medication use and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 dropped out by 18 months; Group 2 Number missing: 38, Reason: 38 dropped out by 18 months

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: WOMAC physical function subscale at 18 months; Group 1: mean 18.2 (SD 14.63); n=43, Group 2: mean 19.4 (SD 17.08); n=35; WOMAC physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline exercise: 29.70 (14.09). Baseline no treatment: 28.95 (15.28).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, affected knee, duration of osteoarthritis, weight, BMI, walking aid use, racial background, marital status, level of education, medication use and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 dropped out by 18 months; Group 2 Number missing: 38, Reason: 38 dropped out by 18 months

Protocol outcome 3: Pain at > 3 months
- Actual outcome: WOMAC pain subscale at 18 months; Group 1: mean 23.6 (SD 15.09); n=43, Group 2: mean 23.5 (SD 17.78); n=35; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline exercise: 31.15 (14.29). Baseline no treatment: 30.30 (16.47).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, affected knee, duration of osteoarthritis, weight, BMI, walking aid use, racial background, marital status, level of education, medication use and baseline values of outcomes; Group 1 Number missing: 35, Reason: 35 dropped out by 18 months; Group 2 Number missing: 38, Reason: 38 dropped out by 18 months

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Bruce-brand 2012[67]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=26)
Countries and setting	Conducted in Irish Republic; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 14 weeks (6 weeks of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic moderate to severe knee osteoarthritis confirmed radiographically as Kellgren Lawrence grade 3-4
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 55-75 years with symptomatic, moderate to severe knee osteoarthritis; arthroscopically grade 3-4 osteoarthritis on the Outerbridge scale within the last 2 years, or placed within the last 6 months on the waiting list for knee replacement surgery, confirmed radiographically with Kellgren-Lawrence severity grades 3-4
Exclusion criteria	Medical co-morbidities precluding participation in an exercise program; implanted electrical devices; neurological disorders; inflammatory arthritis; non-ambulatory status; significant cognitive impairment; participation in an exercise program within the last 6 months; involvement in a previous similar study; anticoagulant therapy; and recent or imminent surgery (within 3 months)
Recruitment/selection of patients	People were recruited from the arthroscopy database and knee arthroplasty waiting list from Cappagh National Orthopaedic Hospital
Age, gender and ethnicity	Age - Mean (SD): 64.0 (5.4). Gender (M:F): 11:15. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 3-4 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=14) Intervention 1: Exercise - Supervised strength exercise. Resistance training groups for 3 home based training sessions per week for 6 weeks. Each session was approximately 30 minutes in duration and was separated by a minimum of 36 hours. Two of the three weekly sessions were supervised by an exercise specialist. Exercises included: knee presses, bottle knee presses, extended leg raises, leg

	extensions, wall squats and hamstring curls. This comprised 3 sets of 10 repetitions for each of the 6 exercises. Each set was performed bilaterally starting with the less affected limb.. Duration 6 weeks of treatment, with 14 weeks follow up in total. Concurrent medication/care: Standard care was available to all including osteoarthritis education, weight loss, pharmacologic therapy and physical therapy. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=13) Intervention 2: No treatment. No treatment. Duration 14 weeks. Concurrent medication/care: Standard care was available to all including osteoarthritis education, weight loss, pharmacologic therapy and physical therapy. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=14) Intervention 3: Other. Neuromuscular electrical stimulation treatment. Duration 14 weeks. Concurrent medication/care: Standard care was available to all including osteoarthritis education, weight loss, pharmacologic therapy and physical therapy. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: Not included in this review as this is not a comparison in this protocol
Funding	Study funded by industry (This study was supported by a grant from the Cappagh Hospital Trust. The Kneehab stimulators were provided by Bio-Medical Research Ltd, Galway, Ireland. Neither sponsor had any involvement in the design of the study, in the collection, analysis and interpretation of data; in the writing of the manuscript; or in the decision to submit the manuscript for publication)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: SF-36 physical health at 14 weeks; Group 1: mean 53.2 (SD 25.09); n=10, Group 2: mean 67.83 (SD 21.71); n=6; SF-36 physical health 0-100 Top=High is good outcome; Comments: Baseline exercise: 39.73 (16.51). Baseline no treatment: 51.78 (24.34) Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 outcomes in baseline. Other outcomes similar between groups.; Group 1 Number missing: 4, Reason: 2 unwell, 1 work commitments, 1 underwent total knee replacement; Group 2	

Number missing: 7, Reason: 3 unwell, 1 personal reasons, 2 underwent total knee replacement, 1 away abroad - Actual outcome: SF-36 mental health at 14 weeks; Group 1: mean 65.3 (SD 24.91); n=10, Group 2: mean 70.5 (SD 22.4); n=6; SF-36 mental health 0-100 Top=High is good outcome; Comments: Baseline exercise: 56.36 (21.91). Baseline control: 62.00 (25.41). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 outcomes in baseline. Other outcomes similar between groups.; Group 1 Number missing: 4, Reason: 2 unwell, 1 work commitments, 1 underwent total knee replacement; Group 2 Number missing: 7, Reason: 3 unwell, 1 personal reasons, 2 underwent total knee replacement, 1 away abroad Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC physical function at 14 weeks; Group 1: mean 31.5 (SD 14.4); n=10, Group 2: mean 21.67 (SD 18.9); n=6; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline exercise: 31.68 (12.92). Baseline control: 31.67 (17.95). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 outcomes in baseline. Other outcomes similar between groups.; Group 1 Number missing: 4, Reason: 2 unwell, 1 work commitments, 1 underwent total knee replacement; Group 2 Number missing: 7, Reason: 3 unwell, 1 personal reasons, 2 underwent total knee replacement, 1 away abroad Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 14 weeks; Group 1: mean 9.6 (SD 4.14); n=10, Group 2: mean 8.33 (SD 4.08); n=6; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 11.05 (3.02). Baseline control: 9.00 (3.65). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in SF-36 outcomes in baseline. Other outcomes similar between groups.; Group 1 Number missing: 4, Reason: 2 unwell, 1 work commitments, 1 underwent total knee replacement; Group 2 Number missing: 7, Reason: 3 unwell, 1 personal reasons, 2 underwent total knee replacement, 1 away abroad	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Cantero-tellez 2021[72]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=12)
Countries and setting	Conducted in Spain; Setting: Outpatient follow up

Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months (12 weeks)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Thumb carpometacarpal joint osteoarthritis diagnosed as grade I or II by the Eaton Classification Stage
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	A diagnosis of grade I or II thumb CMC joint osteoarthritis according to the Eaton Classification Stage in their dominant hand; a minimum pain rating of 4/10 on the Visual Analogue Scale during activities of daily living at initial evaluation; the ability to read and understand the patient information sheets and exercises, and the ability to sign a consent form.
Exclusion criteria	A neurological disorder affecting the upper limb; received other conservative treatments in the last 6 months for thumb CMC joint osteoarthritis; fractures; tenosynovitis; other significant injuries to the thumb, hand or wrist; had a diagnosis of Dupuytren disease.
Recruitment/selection of patients	Recruited from Tecan Hand Center located in Malaga, Spain, where they were seeking hand therapy treatment for symptoms related to thumb CMC joint osteoarthritis.
Age, gender and ethnicity	Age - Range of means: 65.33-67.17. Gender (M:F): Not stated/unclear. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hand osteoarthritis
Extra comments	Severity: Eaton Classification Stage grades I-II Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=6) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). A supervised proprioceptive training program, divided into three phases. Each phase is performed for 2 consecutive weeks and are as follows: phase 1, threshold to detection of passive motion; phase 2, reproduction of passive and active joint position; and phase 3, active movement extent discrimination assessment. Phase one involved the therapist passively moving the patient's thumb MCP or interphalangeal joint (while the patient is blindfolded) and asking the patient to identify the direction in which the thumb was moved. Phase 2 involves passive angle repositioning and active angle repositioning. Phase 3 involves a variety of devices used for everyday tasks that incorporate different strengths, textures and weight of the objects. The person is instructed to use them to introduce dynamic proprioception using 'real-life' movements with everyday objects.. Duration 12 weeks. Concurrent medication/care: Conservative treatments including a short opponens orthosis for night-time wear, self passive traction of the thumb CMC joint, self-massage to the thumb muscles, active resistance of the FDI muscle, and instruction for functional incorporation of the thumb for activities of daily living. The exercise routine was performed on a home program basis 2 times per day (3 sets of 8-10 repetitions) and seen twice a week in the clinic to monitor and provide feedback for proper performance of the exercise routine.. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Proprioception (n=6) Intervention 2: Exercise - Supervised strength exercise. No additional treatment. Duration 12 weeks. Concurrent medication/care: Conservative treatments including a short opponens orthosis for night-time wear, self passive traction of the thumb CMC joint, self-massage to the thumb muscles, active resistance of the FDI muscle, and instruction for functional incorporation of the thumb for activities of daily living. The exercise routine was performed on a home program basis 2 times per day (3 sets of 8-10 repetitions) and seen twice a week in the clinic to monitor and provide feedback for proper performance of the exercise routine.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	No funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE Protocol outcome 1: Pain at </=3 months - Actual outcome: Pain)visual analogue scale) at 12 weeks; Group 1: mean 3.67 (SD 0.81); n=6, Group 2: mean 3.5 (SD 1.05); n=6; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline supervised mixed modality: 5.67 (1.21). Baseline supervised strength: 5.83 (1.17). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age and baseline symptoms; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Chaipinyo 2009[77]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=48)
Countries and setting	Conducted in Thailand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis as per the American College of Rheumatology clinical criteria achieving at last five of the following: age at least 50 years, morning joint stiffness that usually resolved within 30 minutes, crepitus with active motion of the knee, bony tenderness, bony enlargement, or no palpable warmth of the synovium
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Volunteers at least 50 years of age who met the American College of Rheumatology clinical criteria for knee osteoarthritis
Exclusion criteria	If they had a history of cardiovascular disease; Parkinsonism; osteoporosis; limitations in knee motion that prevented them from comfortably positioning their knee for knee strength measurement; were unable to walk for 15 metres; had been receiving intra-articular injections or physiotherapy intervention for their knee during the preceding six months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 66 (7.2). Gender (M:F): 11:37. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=24) Intervention 1: Exercise - Other unsupervised exercise (including flexibility, proprioception). Balance exercise performed as 30 repetitions of stepping forward and backward then sideways for each leg, 5 days a week. Then 30 repetitions of a bilateral mini squat within pain free range in order to strengthen the quadriceps muscle in standing.. Duration 4 weeks. Concurrent medication/care: No additional information.

	Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=24) Intervention 2: Exercise - Unsupervised strength exercise. People in the strength group performed 30 repetitions of isometric knee extension in sitting for each leg, 5 days a week. To start the knee was flexed to 90 degrees, then it was maximally extended and a maximum isometric contraction was held for 5 seconds. People were instructed to contract their knee muscles as hard as they could without pain. They performed 10 repetitions/set for 3 sets and took a rest between each set as long as necessary before starting the next set.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This project was supported by the Srinakharinwirot University Research Fund)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER UNSUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS quality of life at 4 weeks; Group 1: mean 6 (SD 16); n=24, Group 2: mean 23 (SD 20); n=18; KOOS quality of life 0-100 Top=High is good outcome; Comments: Baseline balance: 64 (19). Baseline strength: 39 (19).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in KOOS physical function and quality of life subscales at baseline.; Group 1 Number missing: 0; Group 2 Number missing: 6, Reason: 4 lost for other illness, 1 for personal reason, 1 uncontactable

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS function in daily living at 4 weeks; Group 1: mean 7 (SD 14); n=24, Group 2: mean 13 (SD 12); n=18; KOOS function 0-100 Top=High is good outcome; Comments: Baseline balance: 82 (16). Baseline strength: 69 (16).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in KOOS physical function and quality of life subscales at baseline.; Group 1 Number missing: 0; Group 2 Number missing: 6, Reason: 4 lost for other illness, 1 for personal reason, 1 uncontactable

Protocol outcome 3: Pain at </=3 months

- Actual outcome: KOOS pain at 4 weeks; Group 1: mean 8 (SD 8); n=24, Group 2: mean 11 (SD 17); n=18; KOOS pain subscale 0-100 Top=High is good outcome; Comments: Baseline balance: 79 (13). Baseline strength: 71 (16). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in KOOS physical function and quality of life subscales at baseline.; Group 1 Number missing: 0; Group 2 Number missing: 6, Reason: 4 lost for other illness, 1 for personal reason, 1 uncontactable	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Chang 2012[79]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with knee osteoarthritis and no less than Kellgren-Lawrence grade 3
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with unilateral or bilateral knee osteoarthritis based on the Altman diagnosis standard, no less than 3 on the Kellgren-Lawrence grading scale, and showing clinical manifestations.
Exclusion criteria	Those who had undergone knee or hip joint surgery; those with chronic diseases (such as severe cardiovascular disease or rheumatoid arthritis); a lower extremity fracture; lower extremity weakness caused by nervous system disease; people using steroids or hyaluronic acid injected into the knee joints within the past 2 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 67.4 (8.9). Gender (M:F): All female. Ethnicity: Not stated

Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grades 2-3 (majority grade 3) Duration of symptoms (mean [SD]): 9.0 (7.6) months
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise - Unsupervised strength exercise. Elastic bands leg press exercises including 10 repetitions per set with 3 sets per day for 2 sessions per week. Advancement to higher intensity took place in 2 weekly intervals. Duration 8 weeks. Concurrent medication/care: Conventional modality treatments available to everyone included shortwave diathermy, hot packs, transcutaneous electrical nerve stimulation, interferential current and "so on".. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=30) Intervention 2: No treatment. Conventional modality treatments only. Duration 8 weeks. Concurrent medication/care: Conventional modality treatments available to everyone included shortwave diathermy, hot packs, transcutaneous electrical nerve stimulation, interferential current and "so on".. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean -10.7 (SD 5.9); n=24, Group 2: mean -4.5 (SD 4.4); n=17; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 20.0 (8.9). Baseline no treatment: 22.0 (8.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, BMI, knee osteoarthritis diagnosis side and duration, severity, exercise habit, medication use, and baseline values of outcomes; Group 1 Number missing: 6, Reason: Improvement in symptoms = 5, going abroad = 1; Group 2 Number missing: 13, Reason: Improvement in symptoms = 5, changing places = 5, no obvious improvement = 1, absent = 2

Protocol outcome 2: Pain at </=3 months

- Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean -2.3 (SD 1.3); n=24, Group 2: mean -0.9 (SD 1.5); n=17; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 4.2 (1.7). Baseline no treatment: 4.5 (1.7). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight, BMI, knee osteoarthritis diagnosis side and duration, severity, exercise habit, medication use, and baseline values of outcomes; Group 1 Number missing: 6, Reason: Improvement in symptoms = 5, going abroad = 1; Group 2 Number missing: 13, Reason: Improvement in symptoms = 5, changing places = 5, no obvious improvement = 1, absent = 2	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Chao 2020[80]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=185)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 5 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of knee osteoarthritis, Kellgren Lawrence grades I to III with obvious symptoms
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 50 to 70 years; diagnosis of knee osteoarthritis; classified as Kellgren-Lawrence grades I to III with obvious symptoms; patients who provided signed informed consent for inclusion into the clinical trial and agreed to comply with the protocol requirements of the study
Exclusion criteria	Rheumatoid arthritis; previous joint replacement; severe organ failure, specifically patients with cardiovascular diseases, classified as New York Heart Association (NYHA) class III or

	IV; chronic kidney disease, classified as stage at least 3; liver disease, with a Model of End-stage Liver Disease (MELD) score of at least 20; patients with severe mental illness
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 56.3 (10.1). Gender (M:F): 42:124. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade I to III Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=105) Intervention 1: Exercise - Supervised strength exercise. Systematic exercise rehabilitation program mainly including lower limb static, dynamic and flexibility exercises; exercises targeting the gluteus muscles; and core strength training for 20 minutes per day. . Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=80) Intervention 2: Pharmacological treatment - NSAIDs. Administration of NSAIDs and COX-2 inhibitors. In this trial, naproxen and diclofenac were administrated to patients, respectively (27 had diclofenac, 28 had naproxen, 19 had celecoxib). All people received the same drug dosage.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Oral treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	No funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NSAIDS Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: SF-36 at 12 weeks; Group 1: mean 105.4 (SD 21.5); n=92, Group 2: mean 83.4 (SD 4.2); n=74; SF-36 Unclear Top=High is good outcome; Comments: Combined the three types of NSAIDs together to get the final value. Diclofenac = 83.7 (5.0). Naproxen = 82.5 (3.7). Celecoxib = 84.2 (3.5). Baseline exercise = 78.1 (1.2). Baseline diclofenac = 77.9 (1.2). Baseline naproxen = 76.8 (3.2). Baseline celecoxib = 78.9 (1.9). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI and baseline values of outcomes; Group 1 Number missing: 13, Reason: 13 lost to follow up; Group 2 Number missing: 6, Reason: 6 lost to follow up	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Pain at > 3 months; Osteoarthritis flares at

	</=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Chen 2019[81]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=171)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Previously diagnosed with knee osteoarthritis with knee pain
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 60 years of age or older; experiencing knee pain on most days of the past month; average knee pain in the last week between 3 and 7 on an 11-point numeric rating scale; having intact cognitive functioning, as indicated by a Short Portable Mental Status Questionnaire score of 8-10.
Exclusion criteria	Joint replacement surgery or arthroscopic surgery on the affected side of the knee; other surgery on lower limbs within the past 6 months; severe deformity of lower limbs; having health problems that can easily induce adverse events during home exercise, a myocardial infarction, cerebral infarction, unstable angina, arrhythmia, severe vision problems, or neurological dysfunction.
Recruitment/selection of patients	People were recruited from four community centers in Beijing via print and social media advertisements
Age, gender and ethnicity	Age - Other: Mean: 68.9. Gender (M:F): 22:119. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean): 6.4 years
Indirectness of population	No indirectness
Interventions	(n=84) Intervention 1: Exercise - Unsupervised strength exercise. Home based exercise over 12 weeks with 4 weeks of physiotherapy training in exercise and health education. People were recommended an exercise prescription of 30-40 minutes per day at least 3 days a week including: isometric contractions of the quadriceps, supine straight leg lifts, leg lifts in the prone position, resistance knee extension, resistance

	knee flexion, passive knee flexion, passive knee extension, and shifting the center of gravity.. Duration 12 weeks. Concurrent medication/care: Health education was available to both groups (with the control group not receiving any education regarding exercise). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=87) Intervention 2: No treatment. Health education without any reference to exercise. Duration 12 weeks. Concurrent medication/care: Health education was available to both groups (with the control group not receiving any education regarding exercise). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	No funding (This research did not receive a specific grant from funding agencies in the public, commercial, or not-for-profit sectors)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: AIMS2-SF total at 12 weeks; Group 1: mean 82 (SD 9.96); n=71, Group 2: mean 77.9 (SD 9.52); n=70; Arthritis Impact Measurement Scales 2 - Short form 19-95 Top=High is good outcome; Comments: Baseline exercise: 75.06 (10.00). Baseline control: 76.57 (10.62).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, symptom duration, level of education, marital status, number of affected knees, uses a walker, comorbid conditions, current drug use and baseline values of outcomes; Group 1 Number missing: 13, Reason: Lost to follow up = 13. Immigration = 3, intra-articular injection therapy = 2, surgical treatment = 1, quit due to busy = 6, other illness = 1; Group 2 Number missing: 17, Reason: Lost to follow up = 17. unable to contact = 3, immigration = 1, intra-articular injection therapy = 2, acupuncture = 1, quit due to busy = 7, surgical treatment = 2, go on a holiday = 1

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 4.28 (SD 3.3); n=71, Group 2: mean 5.73 (SD 3.54); n=70; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 7.34 (3.36). Baseline no treatment: 7.19 (4.48).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, symptom duration, level of education, marital status, number of affected knees, uses a walker, comorbid conditions, current drug use and baseline values of outcomes; Group 1 Number missing: 13, Reason: Lost to follow up = 13. Immigration = 3, intra-articular injection therapy = 2, surgical treatment = 1, quit due to busy = 6, other illness = 1; Group 2 Number missing: 17, Reason: Lost to follow up = 17. unable to contact = 3, immigration = 1, intra-articular injection therapy = 2,

acupuncture = 1, quit due to busy = 7, surgical treatment = 2, go on a holiday = 1	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Chen 2021[83]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=32)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed by the American College of Rheumatology clinical criteria enrolled from outpatients of the hospital (including radiological evidence)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed with knee osteoarthritis by the American College of Rheumatology clinical criteria; age from 50 to 75 years; Kellgren Lawrence grade at least 1 in one or both knees; no balance training experience, such as Tai Chi, Baduanjin and Yoga, prior to six months beofre enrollment; an ability to stand independently on the platform for 30 seconds without any assistive device for static stability test and depict 5 circles within 120 seconds for the proprioception assessment.
Exclusion criteria	Presence of any known inflammatory rheumatic disease/arthritis; concomitant neurologic diseases, such as stroke, Parkinson's disease, severe cardiovascular respiratory, spinal cord injury, or other musculoskeletal diseases; presence of acute joint effusion in knees; use of any medications that could affect the musculoskeletal system or postural stability; history of ankle diseases and lower extremity fracture/surgery.

Recruitment/selection of patients	People recruited from outpatient services
Age, gender and ethnicity	Age - Mean (SD): 60.6 (7.4). Gender (M:F): 6:26. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade at least 1 in one or both knees Duration of symptoms (SD): 37.3 (36.5) months
Indirectness of population	No indirectness
Interventions	(n=16) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Backwards walking training in addition to conventional training. Backwards walking for 10 minutes with 5 minutes of warm-up and cool-down sessions 3 days a week for 4 weeks at their comfortable walking speed. Participants took the training session in the hospital for the first time under the supervision of another therapist, and then continued the practice at home and gradually increased their walking time to 30 minutes over the 4 week period as long as they did not experience increased pain.. Duration 4 weeks. Concurrent medication/care: Conventional treatment comprising acupotomy, medications and routine exercise, once a week for 4 weeks. Based on the previous method, the subjects in both groups were treated with needle-knife therapy at the dominant inserted points of Neixiyan and Waixiyan, as well as the conjugate points Dubi (ST35) and Xuehai (SP10). All were prescribed with an oral medication, Celebrex capsules (0.2g/d, once a day) for the first 6 days, while no extra painkillers were used in the next 3 weeks. Additionally, straight leg raising, as a routine exercise, was prescribed to practice at home for both legs, 1 set of 10 repetitions, twice a day, and gradually increase exercise time to 3 sets over the 4-week period, according to their pain intensity.. Indirectness: No indirectness Further details: 1. Class of medicine: Not stated / Unclear 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and aerobic). (n=16) Intervention 2: Exercise - Unsupervised strength exercise. No additional treatment. Duration 4 weeks. Concurrent medication/care: Conventional treatment comprising acupotomy, medications and routine exercise, once a week for 4 weeks. Based on the previous method, the subjects in both groups were treated with needle-knife therapy at the dominant inserted points of Neixiyan and Waixiyan, as well as the conjugate points Dubi (ST35) and Xuehai (SP10). All were prescribed with an oral medication, Celebrex capsules (0.2g/d, once a day) for the first 6 days, while no extra painkillers were used in the next 3 weeks. Additionally, straight leg raising, as a routine exercise, was prescribed to practice at home for both legs, 1 set of 10 repetitions, twice a day, and gradually increase exercise time to 3 sets over the 4-week period, according to their pain intensity.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by the Scientific Research Project of the Traditional Chinese Medicine Bureau of Guangdong Province (no.20194002), Soft Science Research Program of Guangdong Province (no.2018B020207009), and Guangdong Science and Technology Innovation Strategy Special Fund (no.2021b1111610007).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 4 weeks; Group 1: mean -6.44 (SD 3.69); n=16, Group 2: mean -2.88 (SD 1.78); n=16; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline unsupervised mixed modality exercise: 14.63 (3.56). Baseline unsupervised strength exercise: 15.00 (3.31).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, Kellgren Lawrence scale, duration of symptoms and baseline values of symptoms.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 4 weeks; Group 1: mean -3 (SD 1.67); n=16, Group 2: mean -1.88 (SD 1.03); n=16; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline unsupervised mixed modality exercise: 5.63 (1.93). Baseline unsupervised strength exercise: 5.19 (1.56).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, Kellgren Lawrence scale, duration of symptoms and baseline values of symptoms.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Cheung 2014[85]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=36)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis of the knee diagnosed at least 6 months prior. Symptoms classified under the American College of Rheumatology criteria, which does not require any radiographic evidence
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Community-dwelling women between the ages of 65 and 90 years; had symptomatic osteoarthritis of the knee diagnosed for at least 6 months; had no previous training in any form of yoga; were not currently participating in a supervised exercise program
Exclusion criteria	Short portal mental status questionnaire score less than 8 (indicating moderate/severe cognitive impairment); symptoms of join locking; instability indicated by chronic use of a knee brace, cane, walker, or wheelchair; a corticosteroid injection in the symptomatic joint within three months of study entry; a hyaluronicc acid injection in the symptomatic joint within 6 months of study entry; a history of knee surgery within the last two years or a joint replacement at any point; individuals reporting significant medical comorbidities that might preclude exercise participation including: uncontrolled high blood pressure or existing heart condition; other comorbid condition with overlapping symptoms (i.e. fibromyalgia, rheumatoid arthritis)
Recruitment/selection of patients	Advertisements to various senior centers, distributing press release to the university's Alumnae Monthly Newsletter, local and senior newspapers, accessing the database and mailing invitation letters out to patients meeting demographic and diagnostic criteria from the University of Minnesota Physician Practice
Age, gender and ethnicity	Age - Other: Mean (95% CIs): yoga = 71.9 (69.3, 74.6), waiting list = 71.9 (69.0, 75.0). Gender (M:F): All participants were women. Ethnicity: Predominantly white (86%)
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: High morbidity score (Yoga mean (95% CI): 2.8 (1.7, 3.9). Waiting list control: 1.4 (0.8, 2.0).). 4. Site of osteoarthritis: Knee osteoarthritis

Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=18) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Hatha yoga classes, lasting 60 minutes per week for eight weeks. Sessions included asanas (poses) in the seated, supine, and standing positions; pranas (beathing); and meditation. The class size was 9 people per day. All classes were taught by the same yoga practitioners. In addition to attending classes, people were instructed to practice 30-minute yoga four times a week at home. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Yoga). (n=18) Intervention 2: No treatment. Waiting list control. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by grants from the John A. Hartford Foundation, Atlantic Philanthropies, Midwest Nursing Research Society Joanne Stevenson Seed Grant, and St. Catherine University)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical component summary at 8 weeks; Group 1: mean 38 (SD 4.2); n=18, Group 2: mean 38.7 (SD 4.2); n=18; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reports mean (SE). Reported yoga: 38.0 (0.98). Reported control: 38.7 (1.0). Baseline yoga: 39.5 (6.2). Baseline control: 33.9 (4.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC physical function subscale difference. Otherwise ok. Reported age, BMI, education and baseline values of outcomes; Group 1 Number missing: 0, Reason: No lost participants; Group 2 Number missing: 2, Reason: 1 withdrew due to family obligation, 1 disqualified due to receiving cortisone injection
- Actual outcome: SF-12 mental component summary at 8 weeks; Group 1: mean 49.7 (SD 5.1); n=18, Group 2: mean 51.7 (SD 5.1); n=18; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reports mean (SE). Reported yoga: 49.7 (1.2). Reported control: 51.7 (1.2). Baseline yoga: 51.0 (5.7). Baseline control: 53.4 (4.8).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC physical function subscale difference. Otherwise ok. Reported age, BMI, education and baseline values of outcomes; Group 1 Number missing: 0, Reason: No lost participants; Group 2 Number missing: 2, Reason: 1 withdrew due to family obligation, 1 disqualified due to receiving cortisone injection Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 22 (SD 9.8); n=18, Group 2: mean 26.2 (SD 9.8); n=18; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean (SE). Reported yoga: 22.0 (2.3). Reported control: 26.2 (2.3). Baseline yoga: 35.0 (11.8). Baseline control: 27.1 (15.2)). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC physical function subscale difference. Otherwise ok. Reported age, BMI, education and baseline values of outcomes; Group 1 Number missing: 0, Reason: No lost participants; Group 2 Number missing: 2, Reason: 1 withdrew due to family obligation, 1 disqualified due to receiving cortisone injection Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 5.8 (SD 2.8); n=18, Group 2: mean 8.3 (SD 2.8); n=18; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean (SE). Reported yoga: 5.8 (0.67). Reported control: 8.3 (0.67). Baseline yoga: 9.3 (4.0). Baseline control: 7.7 (4.2). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC physical function subscale difference. Otherwise ok. Reported age, BMI, education and baseline values of outcomes; Group 1 Number missing: 0, Reason: No lost participants; Group 2 Number missing: 2, Reason: 1 withdrew due to family obligation, 1 disqualified due to receiving cortisone injection Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Adverse events at 8 weeks; Group 1: 0/18, Group 2: 0/18; Comments: No one reported any yoga practice related adverse events/injuries Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC physical function subscale difference. Otherwise ok. Reported age, BMI, education and baseline values of outcomes; Group 1 Number missing: 0, Reason: No lost participants; Group 2 Number missing: 2, Reason: 1 withdrew due to family obligation, 1 disqualified due to receiving cortisone injection	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Cheung 2017[84]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=83)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Partially adequate method of assessment/diagnosis: A self-reported medical diagnosis of osteoarthritis of the knee for at least 6 months
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Community-dwelling adults aged 60 years or over; a self-reported medical diagnosis of symptomatic osteoarthritis of the knee for at least 6 months; have not practiced any form of yoga for 2 months because physiological changes induced by regular exercise training are generally lost after 4-8 weeks of detraining; not currently participating in a supervised exercise program more than 2 times a week.
Exclusion criteria	Symptoms of joint locking to a degree that affects the individual's balance and makes participating in a group exercise program unsafe; chronic use of assistive devices; corticosteroid injections within 3 months of study entry; hyaluronic acid injection within 6 months of study start date; history of knee surgery within the last 2 years; knee joint replacement; self-reported comorbidities including uncontrolled hypertension, unstable heart conditions, or comorbidities with overlapping symptoms (i.e. rheumatoid arthritis).
Recruitment/selection of patients	People were recruited through osteoarthritis related presentations at various community and senior centers, senior programs, flyers, press releases, and community newsletters
Age, gender and ethnicity	Age - Mean (SD): 71.6 (8.1). Gender (M:F): 14:70. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Low morbidity score (1.5 (1.5)). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=32) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Hatha yoga program designed by a group of expert yoga teachers composed of one 45-minute class per week for eight weeks and additional 30 min/day, four times/week of yoga practice at home during the intervention period. Sessions included poses in the seated, supine, prone, and standing positions; breathing exercises, and relaxation/mindfulness training. Key yoga poses included @easy@ seated pose, reclining bound angle, half locust variation, head to knee pose, bridge, standing forward fold, chair pose, mountain pose, warrior I and II, tree pose variation, reclining hamstring stretch with hip opener with strap, reclining twist, and relaxation pose. A progressive series of poses with props such as yoga mats, blocks, straps, blankets, and chairs were used during class, and poses were modified when needed based on the participants' physical abilities to increase confidence and the ability to remain in pose and achieve benefits. Each class consisted of approximately 8-10 yoga poses and with 2-3 new, variable poses were introduced at each session.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (n=28) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Aerobic-strength exercise consisting of eight weekly group-based classes that involved 15 minutes of mild aerobic exercise that served as a full body warm up, and 30 minutes of strengthening exercises including both isometric (without moving the joints) and isotonic (moving the joints) exercises. Additionally, participants were asked to practice the aerobic portion of the program for 15-30 minutes/day, four times/week and the strengthening exercise 30min/day, two times/week on non-consecutive days at home. The program was progressive in nature. It was based on the current Arthritis Foundation recommendations and taught by a certified arthritis exercise instructor who taught all the classes. The specific types of exercises included head rotations, shoulder flexion/extension, torso 360 degree rotation (circles in both direction), shoulder circles, marching in place, heel and toe raises, overhead arm reaches, side bends, torso twist (gentle 30 degrees), seated side steps alternating sides, and ankle circles. Props such as elastic bands and chairs were used during the class.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength, aerobic).

	(n=23) Intervention 3: No treatment. Preprinted education brochures from the Arthritis Foundation on how to manage osteoarthritis pain, and physical activity and exercise for osteoarthritis. Each participant in the control group received weekly telephone calls during the 8 week intervention period. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by the University of Iowa Hartford Center Geriatric Nursing Excellent Pilot Grant, and Deborah E. Powell Center of mature Women's Health and Research Grants. It is also supported in part by the National Center for Advancing Translational Sciences Award UL1TR000114. The study sponsors played no role in study design, methods, participant recruitment, data collection, data analysis or development of this manuscript)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED)

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical component summary at 8 weeks; Group 1: mean 41.5 (SD 8.5); n=32, Group 2: mean 38.8 (SD 9); n=28; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 41.5 (38.6, 44.5). Reported aerobic/strength exercise: 38.8 (35.4, 42.1). Calculated SD yoga: 8.5. Calculated SD aerobic/strength exercise: 9.0. Baseline yoga: 38.9 (9.7). Baseline aerobic/strength exercise: 27.6 (11.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).
- Actual outcome: SF-12 mental component summary at 8 weeks; Group 1: mean 55.2 (SD 8.7); n=32, Group 2: mean 53.8 (SD 9.2); n=28; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 55.2 (52.2, 58.2). Reported aerobic/strength exercise: 53.8 (50.4, 57.2). Calculated SD yoga: 8.7. Calculated SD aerobic/strength exercise: 9.2. Baseline yoga: 53.5 (9.8). Baseline aerobic/strength exercise: 52.4 (10.2).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 18.2 (SD 8.4); n=32, Group 2: mean 25.8 (SD 8.4); n=28; WOMAC

physical function 0-68 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 18.2 (15.3, 21.1). Reported aerobic/strength exercise: 25.8 (22.7, 28.9). Calculated SD yoga: 8.4. Calculated SD aerobic/strength exercise: 8.4. Baseline yoga: 27.1 (13.2). Baseline aerobic/strength exercise: 29.9 (15.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 5.1 (SD 2.7); n=32, Group 2: mean 6.5 (SD 2.6); n=28; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 5.1 (4.1, 6.0). Reported aerobic/strength exercise: 6.5 (5.5, 7.4). Calculated SD yoga: 2.7. Calculated SD aerobic/strength exercise: 2.6. Baseline yoga: 7.9 (2.8). Baseline aerobic/strength exercise: 7.7 (4.4).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).

Protocol outcome 4: Psychological distress at </=3 months
- Actual outcome: HADS-anxiety subscale at 8 weeks; Group 1: mean 3.8 (SD 2.6); n=32, Group 2: mean 5.2 (SD 2.7); n=28; HADS-anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 3.8 (2.9, 4.7). Reported aerobic/strength exercise: 5.2 (4.2, 6.2). Calculated SD yoga: 2.6. Calculated SD aerobic/strength exercise: 2.7. Baseline yoga: 5.5 (3.3). Baseline aerobic/strength exercise: 4.9 (3.7).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).
- Actual outcome: HADS-depression subscale at 8 weeks; Group 1: mean 3.8 (SD 2); n=32, Group 2: mean 4.2 (SD 2); n=28; HADS-depression subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 3.8 (3.1, 4.5). Reported aerobic/strength exercise: 4.2 (3.5, 5.0). Calculated SD yoga: 2.0. Calculated SD aerobic/strength exercise: 2.0. Baseline yoga: 4.0 (3.0). Baseline aerobic/strength exercise: 4.4 (2.4).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months

- Actual outcome: SF-12 physical component summary at 8 weeks; Group 1: mean 41.5 (SD 8.5); n=32, Group 2: mean 39 (SD 8.4); n=23; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 41.5 (38.6, 44.5). Reported control: 39.0 (35.5, 42.4). Calculated SD yoga: 8.5. Calculated SD control: 8.4. Baseline yoga: 38.9 (9.7). Baseline control: 39.1 (10.5).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact
- Actual outcome: SF-12 mental component summary at 8 weeks; Group 1: mean 55.2 (SD 8.7); n=32, Group 2: mean 52.8 (SD 8.8); n=23; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 55.2 (52.2, 58.2). Reported control: 52.8 (49.2, 56.4). Calculated SD yoga: 8.7. Calculated SD control: 8.8. Baseline yoga: 53.5 (9.8). Baseline control: 58.1 (7.4).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 18.2 (SD 8.4); n=32, Group 2: mean 25.2 (SD 8.4); n=23; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 18.2 (15.3, 21.1). Reported control: 25.2 (21.8, 28.7). Calculated SD yoga: 8.4. Calculated SD control: 8.4. Baseline yoga: 27.1 (13.2). Baseline control: 24.3 (11.3).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 5.1 (SD 2.7); n=32, Group 2: mean 6.5 (SD 2.7); n=23; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 5.1 (4.1, 6.0). Reported control: 6.5 (5.4, 7.6). Calculated SD yoga: 2.7. Calculated SD control: 2.7. Baseline yoga: 7.9 (2.8). Baseline control: 6.3 (3.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact

Protocol outcome 4: Psychological distress at </=3 months
- Actual outcome: HADS-anxiety subscale at 8 weeks; Group 1: mean 3.8 (SD 2.6); n=32, Group 2: mean 4.4 (SD 2.8); n=23; HADS-anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 3.8 (2.9, 4.7). Reported control: 4.4 (3.3, 5.6). Calculated SD yoga: 2.6. Calculated SD control: 2.8. Baseline yoga: 5.5 (3.3). Baseline control: 4.4 (3.8).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact
- Actual outcome: HADS-depression subscale at 8 weeks; Group 1: mean 3.8 (SD 2); n=32, Group 2: mean 3.7 (SD 2.1); n=23; HADS-depression subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported yoga: 3.8 (3.1, 4.5). Reported aerobic/strength exercise: 4.2 (3.5, 5.0). Reported control: 3.7 (2.8, 4.5). Calculated SD yoga: 2.0. Calculated SD aerobic/strength exercise: 2.0. Calculated SD control: 2.1. Baseline yoga: 4.0 (3.0). Baseline aerobic/strength exercise: 4.4 (2.4). Baseline control: 3.3 (1.8).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 4, Reason: Yoga: 4 withdrew (due to health issue, too busy, or work schedule changed).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical component summary at 8 weeks; Group 1: mean 38.8 (SD 9); n=28, Group 2: mean 39 (SD 8.4); n=23; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 38.8 (35.4, 42.1). Reported control: 39.0 (35.5, 42.4). Calculated SD aerobic/strength exercise: 9.0. Calculated SD control: 8.4. Baseline aerobic/strength exercise: 27.6 (11.1). Baseline control: 39.1 (10.5).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact
- Actual outcome: SF-12 mental component summary at 8 weeks; Group 1: mean 53.8 (SD 9.2); n=28, Group 2: mean 52.8 (SD 8.8); n=23; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 53.8 (50.4, 57.2). Reported control: 52.8 (49.2, 56.4). Calculated SD aerobic/strength exercise: 9.2. Calculated SD control: 8.8. Baseline aerobic/strength exercise: 52.4 (10.2). Baseline control: 58.1 (7.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 25.8 (SD 8.4); n=28, Group 2: mean 25.2 (SD 8.4); n=23; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 25.8 (22.7, 28.9). Reported control: 25.2 (21.8, 28.7). Calculated SD aerobic/strength exercise: 8.4. Calculated SD control: 8.4. Baseline aerobic/strength exercise: 29.9 (15.9). Baseline control: 24.3 (11.3).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 6.5 (SD 2.6); n=28, Group 2: mean 6.5 (SD 2.7); n=23; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 6.5 (5.5, 7.4). Reported control: 6.5 (5.4, 7.6). Calculated SD aerobic/strength exercise: 2.6. Calculated SD control: 2.7. Baseline aerobic/strength exercise: 7.7 (4.4). Baseline control: 6.3 (3.1). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact Protocol outcome 4: Psychological distress at </=3 months - Actual outcome: HADS-anxiety subscale at 8 weeks; Group 1: mean 5.2 (SD 2.7); n=28, Group 2: mean 4.4 (SD 2.8); n=23; HADS-anxiety subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 5.2 (4.2, 6.2). Reported control: 4.4 (3.3, 5.6). Calculated SD aerobic/strength exercise: 2.7. Calculated SD control: 2.8. Baseline aerobic/strength exercise: 4.9 (3.7). Baseline control: 4.4 (3.8). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact - Actual outcome: HADS-depression subscale at 8 weeks; Group 1: mean 4.2 (SD 2); n=28, Group 2: mean 3.7 (SD 2.1); n=23; HADS-depression subscale 0-21 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported aerobic/strength exercise: 4.2 (3.5, 5.0). Reported control: 3.7 (2.8, 4.5). Calculated SD aerobic/strength exercise: 2.0. Calculated SD control: 2.1. Baseline aerobic/strength exercise: 4.4 (2.4). Baseline control: 3.3 (1.8). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, comorbidities, and baseline values of outcomes; Group 1 Number missing: 5, Reason: Mixed modality exercise: 5 withdrew (due to lost contact, falls, leg pain, back pain and too busy).; Group 2 Number missing: 1, Reason: No treatment: 1 lost due to lost to contact	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Christensen 2015[89]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=192)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 68 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis confirmed by clinical symptoms, including pain, and on standing radiographs in at least 1 joint compartment
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged at least 50 years with confirmed knee osteoarthritis based on clinical symptoms, including pain, and on standing radiographs in at least 1 joint compartment. All people were obese, as defined by a BMI of at least 30.
Exclusion criteria	Lack of motivation to lose weight; inability to speak Danish; planned antiobesity surgery; total knee alloplasty; receiving pharmacological therapy for obesity.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.5 (6.4). Gender (M:F): 37:155. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade I-IV, median grade III Duration of symptoms (median {IQR}): Control = 8.0 (4.5-13.0), diet = 8.0 (3.8-10.0), exercise = 9.5 (4.8-15.0)..
Indirectness of population	No indirectness
Interventions	(n=64) Intervention 1: Exercise - Supervised aerobic exercise . 3 day/week exercise program consisting of a warm-up phase (10 minutes), a circuit-training phase (45 minutes), and a cool down/stretching phase (5 minutes). The exercise was divided into 4 periods of 12 weeks and 1 period of 4 weeks, with the idea being to gradually translate the exercise from supervised to unsupervised.. Duration 52 weeks. Concurrent medication/care: All participants had a 12 week period prior to the studies where they had intensive weight loss before being assigned to the groups. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated

	/ Unclear 3. Type of exercise: Not applicable (n=64) Intervention 2: Other. Weight loss therapy. Duration 52 weeks. Concurrent medication/care: All participants had a 12 week period prior to the studies where they had intensive weight loss before being assigned to the groups. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group was not included in the analysis due to it not meeting the protocol (n=64) Intervention 3: No treatment. No attention control with the offer to enter another program at the end of the intervention period. Duration 52 weeks. Concurrent medication/care: All participants had a 12 week period prior to the studies where they had intensive weight loss before being assigned to the groups. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Study funded by industry (Sponsored by the Cambridge Weight Plan)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: SF-36 physical component summary at 68 weeks; Group 1: mean 3.8 (SD 7.8); n=64, Group 2: mean 4.4 (SD 8); n=64; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Reports mean change score (95% confidence interval). Reported exercise: 3.8 (1.9, 5.7). Reported control: 4.4 (2.5, 6.4). Baseline exercise: 34.1 (9.1). Baseline control: 33.5 (9.0).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, duration of symptoms, radiographic stage, joint space width, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 withdrew. 4 non-compliance, 1 adverse event, 1 total knee arthroplasty in target knee; Group 2 Number missing: 9, Reason: 9 withdrew. 5 withdrew consent, 2 adverse events, 2 total knee arthroplasty in target knee
- Actual outcome: SF-36 mental component summary at 68 weeks; Group 1: mean 0.1 (SD 7.6); n=64, Group 2: mean 1.3 (SD 7.6); n=64; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Reports mean change score (95% confidence interval). Reported exercise: 0.1 (-1.7, 2.0). Reported control: 1.3 (-0.5, 3.2). Baseline exercise: 53.2 (10.4). Baseline control: 52.4 (13.3).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, duration of symptoms, radiographic stage, joint space width, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 withdrew. 4 non-compliance, 1 adverse event,

1 total knee arthroplasty in target knee; Group 2 Number missing: 9, Reason: 9 withdrew. 5 withdrew consent, 2 adverse events, 2 total knee arthroplasty in target knee Protocol outcome 2: Physical function at > 3 months - Actual outcome: KOOS function in daily living at 68 weeks; Group 1: mean 8.4 (SD 14.5); n=64, Group 2: mean 6.2 (SD 14.5); n=64; KOOS function in daily living 0-100 Top=High is good outcome; Comments: Reports mean change score (95% confidence interval). Reported exercise: 8.4 (4.8, 11.9). Reported control: 6.2 (2.7, 9.8). Baseline exercise: 60.5 (17.0). Baseline control: 58.3 (16.6). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, duration of symptoms, radiographic stage, joint space width, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 withdrew. 4 non-compliance, 1 adverse event, 1 total knee arthroplasty in target knee; Group 2 Number missing: 9, Reason: 9 withdrew. 5 withdrew consent, 2 adverse events, 2 total knee arthroplasty in target knee Protocol outcome 3: Pain at > 3 months - Actual outcome: KOOS pain at 68 weeks; Group 1: mean 6.8 (SD 15.1); n=64, Group 2: mean 8.7 (SD 15.3); n=64; KOOS pain 0-100 Top=High is good outcome; Comments: Reports mean change score (95% confidence interval). Reported exercise: 6.8 (3.1, 10.5). Reported control: 8.7 (4.9, 12.4). Baseline exercise: 58.5 (16.3). Baseline control: 54.3 (16.2). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, duration of symptoms, radiographic stage, joint space width, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 withdrew. 4 non-compliance, 1 adverse event, 1 total knee arthroplasty in target knee; Group 2 Number missing: 9, Reason: 9 withdrew. 5 withdrew consent, 2 adverse events, 2 total knee arthroplasty in target knee	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Cochrane 2005[91]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=312)
Countries and setting	Conducted in United Kingdom; Setting: Recruited from general practices. Exercises were performed in public swimming baths located in four inner-city communities
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 year intervention period with a further 6 months of follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis confirmed by the person's general practitioner and confirmed by a member of the research team
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged over 60 years who responded 'yes' to two questions: 'Do you have pain in the affected joint on most days of the month?' and 'Is the affected joint stiff first thing in the morning or after a period of sitting?' who were assessed as having osteoarthritis by their general practitioner and a member of the research team
Exclusion criteria	Currently on a waiting list for joint replacement or other surgery; currently receiving hydrotherapy or regularly participating in exercise (defined as more than once per week for 20 minutes or more); having a medical condition that precluded water-based exercise (heart attack in the past 3 months, hip/knee replacement in the past 6 months, stroke in the past 2 months, angina, urinary infection or incontinence, open wounds or skin disease, advanced chronic obstructive pulmonary disease, paralysis or dementia).
Recruitment/selection of patients	Recruited from general practices in the North Staffordshire areas. Recruitment directly from general practice databases.
Age, gender and ethnicity	Age - Mean (SD): 69.75 (6.54). Gender (M:F): 116:196. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging (Explanation appears to be driven by symptoms with no mention of imaging). 3. Multimorbidity: High morbidity score (147 had obesity, 50 had cardiovascular comorbidities, 35 has gastrointestinal comorbidities, 26 had other musculoskeletal comorbidities, 24 had cancer, 14 had opthalmological comorbidities, 14 had obstetric, gynaecological or urinary comorbidities, 12 had endocarine comorbidities, 9 had ENT comorbidities, 7 had skin comorbidities, 6 had respiratory comorbidities, 3 had nutrition and blood comorbidities). 4. Site of osteoarthritis: Mixed (Lower limb osteoarthritis - hip and/or knee).

Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=153) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic exercise therapy, including exercises for flexibility (static stretch), strength (resistance and isometric), isotonic and endurance (aerobic) exertion. Sessions were delivered by specially train instructors, each session lasting for approximately 1 hour. Progression in activities was added every 6-8 weeks by increasing the number of repetitions and/or making the exercises more advanced, for instance by using floats to increase resistance. Participants were asked to attend at least two sessions per week throughout the year (allowing for holidays this averaged at 84 sessions).. Duration 1 year. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=159) Intervention 2: No treatment. No exercise control. Duration 1 year. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The research was funded by the National Coordinating Centre for Health Technology Assessment acting on behalf of the NHS Executive (Project No. 96/32/99).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: EQ-5D visual analogue scale at 18 months; Group 1: mean 62 (SD 19); n=150, Group 2: mean 60 (SD 19); n=157; EQ-5D 0-100 Top=High is good outcome; Comments: Baseline exercise: 60.00 (19.01). Baseline control: 61.67 (17.05).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study discusses that during the randomisation process there were three cases where husband and wife pairs entered the study. In these cases they were both allocated the same treatment, which resulted in a change to the randomisation sequence.; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, comorbidities, baseline values of outcomes and baseline biomechanical values; Group 1 Number missing: 42, Reason: Reasons not given; Group 2 Number missing: 39, Reason: Reasons not given

Protocol outcome 2: Physical function at > 3 months - Actual outcome: WOMAC physical function score at 18 months; Group 1: mean 29.73 (SD 14.62); n=150, Group 2: mean 31.15 (SD 12.73); n=156; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 30.06 (13.13). Baseline control: 31.05 (11.24). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study discusses that during the randomisation process there were three cases where husband and wife pairs entered the study. In these cases they were both allocated the same treatment, which resulted in a change to the randomisation sequence.; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, comorbidities, baseline values of outcomes and baseline biomechanical values; Group 1 Number missing: 42, Reason: Reasons not given; Group 2 Number missing: 39, Reason: Reasons not given Protocol outcome 3: Pain at > 3 months - Actual outcome: WOMAC pain score at 18 months; Group 1: mean 8.49 (SD 3.94); n=152, Group 2: mean 8.88 (SD 3.45); n=158; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 8.72 (3.62). Baseline control: 9.10 (3.14). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - The study discusses that during the randomisation process there were three cases where husband and wife pairs entered the study. In these cases they were both allocated the same treatment, which resulted in a change to the randomisation sequence.; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, comorbidities, baseline values of outcomes and baseline biomechanical values; Group 1 Number missing: 42, Reason: Reasons not given; Group 2 Number missing: 39, Reason: Reasons not given	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	De Matos Brunelli Braghin 2019[100]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a radiographic diagnosis of knee osteoarthritis, grade 1-3 according to the Kellgren and Lawrence classification.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a radiographic diagnosis of knee osteoarthritis, grade 1-3 according to the Kellgren and Lawrence classification.
Exclusion criteria	Presence of cardiovascular, neurological or musculoskeletal disease that disabled the volunteers for the performance of the exercises; uncontrolled diabetes mellitus; dizziness; evidence of secondary, inflammatory or metabolic disease; osteonecrosis and previous intra-articular injection; surgery within the 3 months prior to the study; use of continued anti-inflammatory drugs or participation in exercise therapy within the last 12 months, or use of drug treatment that could potentially have an effect on the results of the study
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.5 (8.0). Gender (M:F): 13:47. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Three stages of 15 supervised group exercise sessions, with a progression of exercises in each phase (4 or 5 sessions). This included: warm up (5 min); strengthening exercises for the lower limbs (20 min): 3 sets of 15 repetitions including flexion, straight leg raise, abduction SLR, and extension SLR; standing knee flexion; quadriceps isometrics, 10 repetitions of 5s, 0 and 30

	degrees; aerobic exercise on a stationary bike for 20 min, starting at 65-70% of maximum heart rate to reach 85-90% of maximum heart rate in the 5th week; and stretching (5 min).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength and aerobic). (n=15) Intervention 2: No treatment. No treatment. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=30) Intervention 3: Other. Exercise and laser therapy or laser therapy alone. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group was not included in the final analysis as the groups did not fulfill the inclusion criteria
Funding	Academic or government funding (This work was supported by the grant: Sao Paulo Research Foundation (FAPESP - process number 2013/18319-3); Coordination for the Improvement of High Education Personnel (CAPES) for financial support and FAPESP (process number 2012/01770-1) for buying GAITRite Platinum 26' Portable Walkway System)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 6.57 (SD 8.28); n=15, Group 2: mean 15.2 (SD 21.73); n=15; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline exercise: 19.22 (19.14). Baseline control: 15.39 (26.69).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 52 people dropped out - 8 did not complete re-evaluations (personal reasons, problems with transportation, family care or acute disease), 44 did not complete treatment (personal reasons, problems with transportation, family care, or acute disease) . The original cohort had 112 people, therefore a significant loss.; Group 2 Number missing: -, Reason: Overall 52 people dropped out - 8 did not complete re-evaluations (personal reasons, problems with transportation, family care or acute disease), 44 did not complete treatment (personal reasons,

problems with transportation, family care, or acute disease) . The original cohort had 112 people, therefore a significant loss. Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 8 (SD 10.99); n=15, Group 2: mean 22.33 (SD 23.59); n=15; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline exercise: 24.64 (26.18). Baseline control: 16.67 (29.07). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 52 people dropped out - 8 did not complete re-evaluations (personal reasons, problems with transportation, family care or acute disease), 44 did not complete treatment (personal reasons, problems with transportation, family care, or acute disease) . The original cohort had 112 people, therefore a significant loss.; Group 2 Number missing: -, Reason: Overall 52 people dropped out - 8 did not complete re-evaluations (personal reasons, problems with transportation, family care or acute disease), 44 did not complete treatment (personal reasons, problems with transportation, family care, or acute disease) . The original cohort had 112 people, therefore a significant loss.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	De rooij 2017[101]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=126)
Countries and setting	Conducted in Netherlands; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 20 weeks and additional 12 weeks of follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the clinical criteria of the American College of Rheumatology
Stratum	Overall:
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of knee osteoarthritis according to the clinical criteria of the American College of Rheumatology and presence of a tleast 1 of the target comorbidities (coronary disease, heart failure, type 2 diabetes mellitus, COPD, or obesity [body mass index (BMI) at least 30kg/m²]), all diagnosed by a medical specialist, with severity score at least 2 for the comorbidity on the Cumulative Illness Rating Scale, indicating that the comorbidity has an impact on daily activities and the person was receiving regular care for their comorbid disease. Also the primary treatment goal was related to knee osteoarthritis (instead of comorbidity related).
Exclusion criteria	Absolute contraindication for exercise therapy (e.g. myocardial infarction within last 3 months); total knee arthroplasty or planned total knee arthroplasty in the near future; participation in exercise therapy for knee osteoarthritis within the preceding 3 months; insufficient comprehension of the Dutch language; psychological distress necessitating treatment; dementia (mini-mental state examination score >24), significant physical limitations that would prohibit the participant from following exercise therapy; an expectation of being lost for follow up (e.g. because of a planned change of residency); refusal to sign informed consent
Recruitment/selection of patients	People were recruited through regular referral by general health practitioners, rheumatologists, rehabilitation physicians, and orthopaedic surgeons, or from advertisements in local newspapers
Age, gender and ethnicity	Age - Mean (SD): 63.6 (10.6). Gender (M:F): 31:95. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (Inclusion criteria required a score of at least 2 on a comorbidity scale). 4. Site of osteoarthritis: Knee osteoarthritis

Extra comments	Severity: Kellgren Lawrence grade 0-4, median grade 2. Duration of symptoms (mean [SD]): 9.0 (9.0) years
Indirectness of population	No indirectness
Interventions	(n=63) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Individualised tailored knee osteoarthritis exercise program, with 2x30-60 minutes sessions including muscle strength training of the lower extremity, aerobic training and training of daily activities. Exercise therapy was adapted by changing frequency, intensity, timing and type factors of the exercises or by adding educational or coaching strategies. People were also advised to perform exercises at home at least 5 times a week.. Duration 20 weeks. Concurrent medication/care: People continued their current medical care for knee osteoarthritis and comorbid disease. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Mixed strength, aerobic and activity based). (n=63) Intervention 2: No treatment. Waiting list control. Duration 20 weeks. Concurrent medication/care: People continued their current medical care for knee osteoarthritis and comorbid disease. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Study funded by industry (The study was funded by Merck Sharp & Dohme and the Royal Dutch Society for PHysical Therapy)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 10 weeks; Group 1: mean 30.4 (SD 11.6); n=60, Group 2: mean 32.9 (SD 11.2); n=55; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 35.1 (11.9). Baseline control: 31.0 (12.3).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, education, gender, duration of symptoms, BMI, clinical diagnosis of knee osteoarthritis, radiographic severity, total number of comorbidities, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 person dropped out due to lack of time; Group 2 Number missing: 3, Reason: 3 people dropped out due to disatsifaction with waiting list period, complications after meniscectomy, and death due to cardiac disease

Protocol outcome 2: Physical function at > 3 months

- Actual outcome: WOMAC physical function subscale at 32 weeks; Group 1: mean 23.5 (SD 13.1); n=51, Group 2: mean 31.4 (SD 12.6); n=56; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 35.1 (11.9). Baseline control: 31.0 (12.3).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, education, gender, duration of symptoms, BMI, clinical diagnosis of knee osteoarthritis, radiographic severity, total number of comorbidities, and baseline values of outcomes; Group 1 Number missing: 9, Reason: 9 people dropped out for vertebrae fracture after fall (not treatment related), acute low back pain, total knee arthroplasty (2), total hip arthroplasty, severe knee pain (2), withdrawal due to lack of time, anxiety disorder; Group 2 Number missing: 5, Reason: 5 people dropped out for total hip arthroplasty, severe knee pain (2), deceased partner, other reason

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 10 weeks; Group 1: mean 8.4 (SD 3); n=60, Group 2: mean 9.1 (SD 3.6); n=55; WOMAC pain subscale 0-17 Top=High is poor outcome; Comments: Baseline exercise: 10.1 (3.4). Baseline control: 9.4 (3.5).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, education, gender, duration of symptoms, BMI, clinical diagnosis of knee osteoarthritis, radiographic severity, total number of comorbidities, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 person dropped out due to lack of time; Group 2 Number missing: 3, Reason: 3 people dropped out due to disatsifaction with waiting list period, complications after meniscectomy, and death due to cardiac disease

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain subscale at 32 weeks; Group 1: mean 6.6 (SD 3.6); n=51, Group 2: mean 8.6 (SD 3.6); n=56; WOMAC pain subscale 0-17 Top=High is poor outcome; Comments: Baseline exercise: 10.1 (3.4). Baseline control: 9.4 (3.5).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, education, gender, duration of symptoms, BMI, clinical diagnosis of knee osteoarthritis, radiographic severity, total number of comorbidities, and baseline values of outcomes; Group 1 Number missing: 9, Reason: 9 people dropped out for vertebrae fracture after fall (not treatment related), acute low back pain, total knee arthroplasty (2), total hip arthroplasty, severe knee pain (2), withdrawal due to lack of time, anxiety disorder; Group 2 Number missing: 5, Reason: 5 people dropped out for total hip arthroplasty, severe knee pain (2), deceased partner, other reason

Protocol outcome 5: Psychological distress at > 3 months
- Actual outcome: Hospital Anxiety and Depression scale score at 32 weeks; Group 1: mean 9.6 (SD 6.5); n=51, Group 2: mean 8 (SD 6.7); n=55; Hospital Anxiety and Depression scale 0-21 Top=High is poor outcome; Comments: Baseline exercise: 11.3 (6.6). Baseline control: 10.0 (6.8).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, education, gender, duration of symptoms, BMI, clinical diagnosis of knee osteoarthritis, radiographic severity, total number of comorbidities, and baseline values of outcomes; Group 1 Number missing: 9, Reason: 9 people dropped out for vertebrae fracture after fall (not treatment related), acute low back pain, total knee arthroplasty (2), total hip arthroplasty, severe knee pain (2), withdrawal due to lack of time, anxiety disorder; Group 2 Number missing: 5, Reason: 5 people dropped out for total hip arthroplasty, severe knee pain (2), deceased partner, other reason

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Diracoglu 2005[106]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=66)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis according to the criteria of American College of Rheumatology with radiological stage 1-2 bilateral knee osteoarthritis according to the Kellgren and Lawrence scale
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People between the ages of 35 and 65 years who were diagnosed as having "primary osteoarthritis" according to the criteria of American College of Rheumatology who had scores equal to or greater than "7" according to the Lequesne Index, and who had radiologically stage 1 and 2 bilateral knee osteoarthritis according to the Kellgren and Lawrence scale.
Exclusion criteria	People diagnosed as having secondary osteoarthritis; people with active synovitis; people who had serious knee trauma surgical interventio; or intraarticular knee injection in the last 6 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Range: 35 to 65 years. Gender (M:F): 0:66. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren and Lawrence grade 1-2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=33) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Kinesthesia and balance exercises with strengthening exercises including modified Romberg exercise, retrowalking, walking on heels/toes/with eye closed, standing on one extremity for 30 seconds, leaning, exercises with balance boards, mini trampoline exercises, plyometric exercise and isometric and isotonic strength exercises. Completed 3 days a week in groups of 5

	people under the supervision of a physiotherapist.. Duration 8 weeks. Concurrent medication/care: Paracetamol was given as an escape medicine for pain control. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength, proprioception). (n=33) Intervention 2: Exercise - Supervised strength exercise. Strengthening exercises only - isometric and isotonic strength exercises. Completed 3 days a week in groups of 5 people under the supervision of a physiotherapist.. Duration 8 weeks. Concurrent medication/care: Paracetamol was given as an escape medicine for pain control. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical function at 8 weeks; Group 1: mean 69.33 (SD 17.8); n=30, Group 2: mean 56.25 (SD 16.7); n=30; SF-36 physical function 0-100 Top=High is good outcome; Comments: Baseline mixed: 48.22 (15.2). Baseline strength: 49.1 (13.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: SF-36 role limitations subscale was different at baseline; Group 1 Number missing: 3, Reason: No additional information; Group 2 Number missing: 3, Reason: No additional information
- Actual outcome: SF-36 role limitation (physical) at 8 weeks; Group 1: mean 77.5 (SD 34.9); n=30, Group 2: mean 57.14 (SD 45); n=30; SF-36 role physical 0-100 Top=High is good outcome; Comments: Baseline mixed: 43.54 (14.2). Baseline strength: 37.1 (12.4).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: SF-36 role limitations subscale was different at baseline; Group 1 Number missing: 3, Reason: No additional information; Group 2 Number missing: 3, Reason: No additional information
- Actual outcome: SF-36 vitality at 8 weeks; Group 1: mean 54 (SD 19.5); n=30, Group 2: mean 43.5 (SD 18.3); n=30; SF-36 vitality 0-100 Top=High is good outcome; Comments: Baseline mixed: 45.0 (16.5). Baseline strength: 42.14 (17.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: SF-36 role limitations subscale was different at baseline; Group 1 Number missing: 3, Reason: No additional information; Group 2 Number missing: 3, Reason: No additional information

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 2 (SD 1.6); n=30, Group 2: mean 2.7 (SD 1.4); n=30; WOMAC physical function 0-

10 Top=High is poor outcome; Comments: Baseline mixed: 4.30 (1.6). Baseline strength: 4.34 (1.1). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: SF-36 role limitations subscale was different at baseline; Group 1 Number missing: 3, Reason: No additional information; Group 2 Number missing: 3, Reason: No additional information Protocol outcome 3: Serious adverse events at </=3 months - Actual outcome: Adverse events at 8 weeks; Group 1: 0/30, Group 2: 0/30 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: SF-36 role limitations subscale was different at baseline; Group 1 Number missing: 3, Reason: No additional information; Group 2 Number missing: 3, Reason: No additional information	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at </=3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Diracoglud 2008[107]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=66)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks, with 1 year total follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis fulfilling the clinical and radiological criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Females diagnosed as having primary osteoarthritis according to the clinical and radiological diagnostic criteria of the American College of Rheumatology with radiological grade I and II, unilateral or bilateral, knee osteoarthritis according to the Kellgren and Lawrence classification
Exclusion criteria	Active synovitis; history of severe trauma; surgical intervention or intraarticular knee injection in the last six months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 50.5 (7.2). Gender (M:F): 0:66. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren and Lawrence grade 1-2, median grade 1 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=33) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Active range of motion exercises and active stretching and isometric strengthening exercises for hamstring and quadriceps muscles. Strengthening exercises were initiated by 40% of 1-repetition maximum, fast contraction velocity and 2-3 minute of rest between sets. Each exercise was applied as 3 sets of 8 repetitions. The program was maintained with 10% increase in 1-RM every week. After reaching 70% of 1-RM at the end of the 3rd week, the program was maintained with this loading level, and a period of 8 weeks was thus completed. In both the groups, one session of exercise lasted for a mean duration of 20 minutes

	during the 1st week. The exercises lasted 30 minutes in week 2, and finally the exercises remained unchanged until week 8 as 40 minutes. People performed the exercises in a clinical setting three days a week for eight weeks in groups of five and under the supervision of a physiotherapist. After this people people were encouraged to continue their exercises at home. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength, range of motion). (n=33) Intervention 2: Exercise - Supervised strength exercise. Strength exercise component only - no balance exercises. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 vitality subscale at 8 weeks; Group 1: mean 54 (SD 19.58); n=33, Group 2: mean 45.54 (SD 18.33); n=33; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Baseline mixed modality exercise: 45.00 (16.53). Baseline strength exercise: 42.14 (17.66).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: SF-36 vitality subscale at 1 year; Group 1: mean 51.07 (SD 15.54); n=33, Group 2: mean 43.82 (SD 11.93); n=33; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Baseline mixed modality exercise: 45.00 (16.53). Baseline strength exercise: 42.14 (17.66).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 1.27 (SD 0.63); n=33, Group 2: mean 1.6 (SD 0.49); n=33; WOMAC physical function subscale 0-4 Top=High is poor outcome; Comments: Baseline mixed modality exercise: 2.03 (0.42). Baseline strength exercise: 2.06 (0.30).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: WOMAC physical function subscale at 1 year; Group 1: mean 1.3 (SD 0.56); n=33, Group 2: mean 1.76 (SD 0.6); n=33; WOMAC physical function subscale 0-4 Top=High is poor outcome; Comments: Baseline mixed modality exercise: 2.03 (0.42). Baseline strength exercise: 2.06 (0.30).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcome 5: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 1.33 (SD 0.58); n=33, Group 2: mean 1.5 (SD 0.57); n=33; WOMAC pain subscale 0-4 Top=High is poor outcome; Comments: Baseline mixed modality exercise: 2.13 (0.39). Baseline strength exercise: 2.20 (0.37).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcome 6: Pain at > 3 months
- Actual outcome: WOMAC pain subscale at 1 year; Group 1: mean 1.26 (SD 0.62); n=33, Group 2: mean 1.7 (SD 0.69); n=33; WOMAC pain subscale 0-4 Top=High is poor outcome; Comments: Baseline mixed modality exercise: 2.13 (0.39). Baseline strength exercise: 2.20 (0.37).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, education, Kellgren and Lawrence grade, side and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 lost to follow up - 1 was not found in their address, 3 did not come to study visit; Group 2 Number missing: 10, Reason: 10 lost to follow up - 5 did not come to study visit, 5 could not be contacted due to change in address and phone number

Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Duman 2012[111]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=54)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology criteria with grade 3 or higher Kellgren Lawrence scale radiographic changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with knee osteoarthritis according to the American College of Rheumatology criteria with grade 3 or higher Kellgren Lawrence radiographic changes
Exclusion criteria	People with a history of surgery of the knee joint; other pathologies that might potentially impair the balance (cerebellar problems, Parkinson disease, vertigo, etc.), dementia, or serious renal or hepatic problems
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 64 (3.7). Gender (M:F): 5:49. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 3-4, median grade 3 Duration of symptoms (mean [SD]): 7.9 (1.7) years
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Proprioceptive exercise program including strengthening of the quadriceps, ankle extensors and hip abductors, bicycling, walking by making 45 degrees corner at every two steps, walking forward by heel-to-toe and then backward by toe-to-heel and walking sidelong to the right and then to the left. Therapy was continued for 3 weeks, 5 days a week for one session a day.. Duration 3 weeks. Concurrent medication/care: All people received non-steroidal anti-inflammatory drugs (meloxicam 15mg/day) and physical therapy (infrared and short wave therapy). Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Proprioception (n=24) Intervention 2: No treatment. No exercise treatment. Duration 3 weeks. Concurrent medication/care: All people received non-steroidal anti-inflammatory drugs (meloxicam 15mg/day) and physical therapy (infrared and short wave therapy). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC function subscale at 3 weeks; Group 1: mean 1307 (SD 286.1); n=30, Group 2: mean 1274 (SD 384); n=24; WOMAC function subscale 0-1800 Top=High is poor outcome; Comments: Baseline exercise: 1356 (249.2). Baseline no treatment: 1282 (380.2). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in WOMAC physical function subscale; Group 1 Number missing: ; Group 2 Number missing: Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 3 weeks; Group 1: mean 316 (SD 84.84); n=30, Group 2: mean 323 (SD 64.9); n=24; WOMAC pain subscale 0-500 Top=High is poor outcome; Comments: Baseline exercise: 316 (84.84). Baseline no treatment: 326 (65.08). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in WOMAC physical function subscale; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Dziedzic 2015[114] (Dziedzic 2011[115], Oopong 2014 [333])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=257)
Countries and setting	Conducted in United Kingdom; Setting: Primary care
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People meeting the criteria for hand osteoarthritis according to the American College of Rheumatology criteria or had unilateral or bilateral thumb base osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 50 years or older who gave consent to further contact; reported hand pain in the last 12 months; reported hand pain, aching or stiffness on 'some days', 'most days' or 'all days' in the last month; had an Austrialian Canadian Hand Osteoarthritis Outcomes Index pain score of at least 5 or an AUSCAN function score of at least 9; reported that they had not seen an occupational therapist or physiotherapist for their hand problem in the last 6 months; had not had a hand operation, injection nor injured their hands badly enough to see a doctor in the previous 6 months; had no other member of their household participating in the trial.
Exclusion criteria	Did not have an alternative clinical diagnosis, such as inflammatory arthritis; were able to attend for the trial interventions at participating occupational therapy departments
Recruitment/selection of patients	People registered with five general practices in Central Cheshire and North Staffordshire, UK, were mailed a health survey. Responders were invited for an assessment at a research clinic to check eligibility for the trial if they met the inclusion criteria
Age, gender and ethnicity	Age - Mean (SD): 65.8 (9.1). Gender (M:F): 87:170. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hand osteoarthritis (Hand, including base of thumb).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=65) Intervention 1: Exercise - Unsupervised strength exercise. Strengthening and strengthening exercises including: wrist flexion and extension, pronation and supination

	exercises, tendon gliding, radial finger walking, making an 'O' with the thumb and index finger, thumb extension, abduction and opposition to the base of the fifth finger, using an elastic band to provide resistance to thumb extension, thumb abduction and finger extension, using Play-Doh rolling and forming into a ring to provide resistance to thumb and finger extension, squeezing it into a ball, and pinching off pieces between the thumb and index fingers; holding a 0.5-0.75kg weight while doing wrist flexion and extension exercises in pronation then supination. People were guided to start with three repetitions of each exercise, gradually building up to 10 repetitions of each exercise daily (or most days) and to perform the exercises within their limit of discomfort. Exercises could be spread over several exercise sessions during the day and performed more than once per day.. Duration 12 months. Concurrent medication/care: Leaflet and advice - all participants were given standardised written information on self-management approaches for hand osteoarthritis including general information only looking after hand joints, and using analgesia. Participants were advised to continue with any self-management approaches they were currently using, and were given advice to consult their general practitioner if symptoms continued to be troublesome.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=65) Intervention 2: No treatment. No additional intervention. Duration 12 months. Concurrent medication/care: Leaflet and advice - all participants were given standardised written information on self-management approaches for hand osteoarthritis including general information only looking after hand joints, and using analgesia. Participants were advised to continue with any self-management approaches they were currently using, and were given advice to consult their general practitioner if symptoms continued to be troublesome.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=127) Intervention 3: Other. Joint protection advice and a treatment package including joint protection and hand exercises. Duration 12 months. Concurrent medication/care: Leaflet and advice - all participants were given standardised written information on self-management approaches for hand osteoarthritis including general information only looking after hand joints, and using analgesia. Participants were advised to continue with any self-management approaches they were currently using, and were given advice to consult their general practitioner if symptoms continued to be troublesome.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable

	3. Type of exercise: Not applicable Comments: These groups were not included in the analysis as they did not meet the inclusion criteria
Funding	Academic or government funding (This study was supported financially by a Project Grant awarded by Arthritis Research UK, Grant Code: 17958 and by Support for Science Funding secured by North Staffordshire Primary Care Research Consortium for NHS service support costs.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: EQ-5D at 3 months; Group 1: mean 0.66 (SD 0.22); n=65, Group 2: mean 0.665 (SD 0.24); n=65; EQ-5D 0-1 Top=High is good outcome; Comments: Baseline exercise: 0.645 (0.21). Baseline no treatment: 0.623 (0.26). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, occupation, age leaving school, further education, BMI, baseline values of outcomes and American College of Rheumatology criteria met; Group 1 Number missing: 5, Reason: 3 no response, 2 declined to complete; Group 2 Number missing: 3, Reason: 3 no response Protocol outcome 2: Health related quality of life at > 3 months - Actual outcome: EQ-5D at 12 months; Group 1: mean 0.708 (SD 0.18); n=65, Group 2: mean 0.634 (SD 0.22); n=65; EQ-5D 0-1 Top=High is good outcome; Comments: Baseline exercise: 0.645 (0.21). Baseline no treatment: 0.623 (0.26). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, occupation, age leaving school, further education, BMI, baseline values of outcomes and American College of Rheumatology criteria met; Group 1 Number missing: 2, Reason: 1 addressee unknown, 1 died (Ca lung); Group 2 Number missing: 7, Reason: 4 no response, 2 declined to complete, 1 died (Ca lung)	
Protocol outcomes not reported by the study	Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Ebnezar 2011[116] (Ebnezar 2012[117], Ebnezar 2012[118])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=250)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People fulfilling the American College of Rheumatology criteria for the diagnosis of osteoarthritis of the knee
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Persistent pain for 3 months prior to recruitment; moderate-to-severe pain on walking; Kellgren and Lawrence radiologic grading of 2-4 in X-rays taken within 6 months prior to entry; those fully ambulant, literate and willing to participate in the study
Exclusion criteria	Grade 1 changes on X-ray; acute knee pain; secondary osteoarthritis due to rheumatoid arthritis, gout, septic arthritis, tuberculosis, tumour, trauma or haemophilia; those with major medical or psychiatric disorders
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 59.5 (9.5). Gender (M:F): 76:174. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (38 had Diabetes, 49 had Hypertension, 171 were Overweight/obese, 145 had Osteoporosis and 56 has Other diseases). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated explicitly. Kellgren Lawrence grade 2-4. Duration of symptoms: Median length 1-2 years
Indirectness of population	No indirectness
Interventions	(n=125) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Integrated yoga including shithilikaranavyayama (loosening and strengthening), asanas, relaxation techniques, pranayama, meditation and didactic lectures on yama, niyama, jnana yoga, bhakti yoga, and karma yoga for a healthy lifestyle change. Duration 12 weeks. Concurrent medication/care: All participants received 20 minutes of physiotherapy with transcutaneous electrical stimulation and ultrasound for 2 weeks. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group

	session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Yoga). (n=125) Intervention 2: Exercise - Supervised strength exercise. Therapeutic exercises included loosening and strengthening practices for all the joints of the upper and lower limbs, brief period of rest, specific knee practices, and supine rest followed by light music. Later people were advised to continue the therapeutic exercise practice of 40 minutes at home for the next 12 weeks. Duration 12 weeks. Concurrent medication/care: All participants received 20 minutes of physiotherapy with transcutaneous electrical stimulation and ultrasound for 2 weeks. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not stated / Unclear
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical functioning subscale at 12 weeks; Group 1: mean 67.5 (SD 9.09); n=125, Group 2: mean 50.94 (SD 14.76); n=125; SF-36 physical functioning subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 12.03 (9.94). Baseline strength: 12.82 (10.81).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 role physical subscale at 12 weeks; Group 1: mean 86.44 (SD 16.55); n=125, Group 2: mean 58.33 (SD 44.52); n=125; SF-36 role physical subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: Not reported. Baseline strength: 0.21 (2.31)
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 role emotion subscale (emotional problems) at 12 weeks; Group 1: mean 86.41 (SD 17.59); n=125, Group 2: mean 58.75 (SD 38.94); n=125; SF-36 role emotional subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: not reported. Baseline strength: 0.56 (6.15).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention

group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 vitality subscale at 12 weeks; Group 1: mean 36.35 (SD 6.08); n=125, Group 2: mean 53.2 (SD 6.86); n=125; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 66.36 (5.66). Baseline strength: 64.91 (5.41).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 mental health subscale (emotional well-being) at 12 weeks; Group 1: mean 34.33 (SD 5.46); n=125, Group 2: mean 52.27 (SD 5.91); n=125; SF-36 mental health subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 63.10 (7.17). Baseline strength: 62.46 (6.61)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 social functioning subscale at 12 weeks; Group 1: mean 64.04 (SD 8.92); n=125, Group 2: mean 57.15 (SD 10.42); n=125; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 50.50 (6.82). Baseline strength: 51.92 (9.37).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 pain subscale at 12 weeks; Group 1: mean 73.77 (SD 12.67); n=125, Group 2: mean 46.93 (SD 11.22); n=125; SF-36 pain subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 11.54 (11.55). Baseline strength: 11.68 (9.11).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment
- Actual outcome: SF-36 general health subscale at 12 weeks; Group 1: mean 77.47 (SD 20.91); n=125, Group 2: mean 60.12 (SD 12.57); n=125; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Baseline yoga: 36.91 (6.94). Baseline strength: 36.99 (11.08).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment

Protocol outcome 2: Pain at </=3 months - Actual outcome: Resting pain (VAS) at 12 weeks; Group 1: mean 1.94 (SD 1.11); n=125, Group 2: mean 4.17 (SD 1.51); n=125; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline yoga: 6.89 (0.69). Baseline strength: 6.68 (0.70). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Does not report baseline value for intervention group for SF-36 role physical and role emotional; Group 1 Number missing: 7, Reason: 7 drop outs - 2 due to getting relief on the 10th day and discontinuing, 3 discontinued due to emergencies at home, 2 due to office calls; Group 2 Number missing: 8, Reason: 8 drop outs - 3 respiratory tract infections, 2 had relief and discontinued, 1 due to emergencies at home, 2 pain became severe and could not continue the treatment	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Evcik 2002[122]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=81)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis with Kellgren and Lawrence grade 1-3 changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with knee osteoarthritis aged between 48 and 71 years with Kellgren Lawrence grade 1-3 changes
Exclusion criteria	Quadriceps exercise program during the last 6 months; effusion on knees; previous knee replacement; severe cardiovascular diseases; grade 4 osteoarthritis according to Kellgren and Lawrence criteria
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 56.4 (6.5). Gender (M:F): 28:53. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 8.0 (3.3)
Indirectness of population	No indirectness
Interventions	(n=27) Intervention 1: Exercise - Unsupervised strength exercise. dose/quantity: Home exercise program including isometric and isotonic quadriceps exercises including: isometric straight leg lifts, isometric quadriceps contraction and isotonic quadriceps exercise. Each exercise was applied at least ten times, twice a day at home. They were taught by a physiotherapist at the hospital.. Duration 12 weeks. Concurrent medication/care: All people were allowed to take analgesic drugs (paracetamol). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable

	(n=28) Intervention 2: Exercise - Unsupervised aerobic exercise . Regular walking program - started for 10 minutes, three times per week. Gradually increased up to half an hour.. Duration 12 weeks. Concurrent medication/care: All people were allowed to take analgesic drugs (paracetamol). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=26) Intervention 3: No treatment. Continue their normal daily activities, no extra exercise or regular walking programs. Duration 12 weeks. Concurrent medication/care: All people were allowed to take analgesic drugs (paracetamol). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus UNSUPERVISED AEROBIC EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Nottingham Health Profile pain subscale at 12 weeks; Group 1: mean 9.8 (SD 3.1); n=27, Group 2: mean 9 (SD 3.3); n=28; Nottingham Health Profile pain subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 40.0 (4.0). Baseline aerobic: 41.3 (4.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile energy subscale at 12 weeks; Group 1: mean 33.4 (SD 2.1); n=27, Group 2: mean 14.6 (SD 1.3); n=28; Nottingham Health Profile energy subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 53.2 (0.3). Baseline aerobic: 50.7 (1.6).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile physical mobility subscale at 12 weeks; Group 1: mean 29.5 (SD 4.8); n=27, Group 2: mean 8.6 (SD 4); n=28; Nottingham Health Profile 0-100 Top=High is poor outcome; Comments: Baseline strength: 40.2 (6.9). Baseline aerobic: 41.3 (7.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.

- Actual outcome: Nottingham Health Profile emotional reactions subscale at 12 weeks; Group 1: mean 19.1 (SD 2.2); n=27, Group 2: mean 6.9 (SD 3); n=28; Nottingham Health Profile emotional reactions subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 28.9 (4.5). Baseline aerobic: 28.6 (4.4). Baseline control: 28.0 (3.7).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile sleep subscale at 12 weeks; Group 1: mean 31.9 (SD 4.9); n=27, Group 2: mean 19.6 (SD 4); n=28; Nottingham Health Profile sleep subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 45.7 (5.2). Baseline aerobic: 44.9 (4.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile social isolation subscale at 12 weeks; Group 1: mean 17.1 (SD 4.1); n=27, Group 2: mean 17.3 (SD 3.9); n=28; Nottingham Health Profile social isolation subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 19.5 (4.3). Baseline aerobic: 21.0 (3.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 10.8 (SD 1.8); n=27, Group 2: mean 10.2 (SD 2.4); n=28; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline strength: 25.4 (5.3). Baseline aerobic: 23.9 (6.3). Baseline control: 25.2 (5.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 3 (SD 1.7); n=27, Group 2: mean 3.4 (SD 1.3); n=28; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline strength: 6.6 (3.2). Baseline aerobic: 6.9 (3.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Nottingham Health Profile pain subscale at 12 weeks; Group 1: mean 9.8 (SD 3.1); n=27, Group 2: mean 20.4 (SD 3.2); n=26; Nottingham Health Profile pain subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 40.0 (4.0). Baseline control: 40.9 (4.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile energy subscale at 12 weeks; Group 1: mean 33.4 (SD 2.1); n=27, Group 2: mean 49.3 (SD 1.7); n=26; Nottingham Health Profile energy subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 53.2 (0.3). Baseline control: 52.9 (2.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile physical mobility subscale at 12 weeks; Group 1: mean 29.5 (SD 4.8); n=27, Group 2: mean 36.6 (SD 6.1); n=26; Nottingham Health Profile physical mobility subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 40.2 (6.9). Baseline control: 44.1 (7.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile emotional reactions subscale at 12 weeks; Group 1: mean 19.1 (SD 2.2); n=27, Group 2: mean 27.9 (SD 4.5); n=26; Nottingham Health Profile emotional reactions subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 28.9 (4.5). Baseline control: 28.0 (3.7).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile sleep subscale at 12 weeks; Group 1: mean 31.9 (SD 4.9); n=27, Group 2: mean 35.3 (SD 4.4); n=26; Nottingham Health Profile sleep subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 45.7 (5.2). Baseline control: 45.0 (5.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile social isolation subscale at 12 weeks; Group 1: mean 17.3 (SD 3.9); n=27, Group 2: mean 19.2 (SD 4.7); n=26; Nottingham Health Profile social isolation subscale 0-100 Top=High is poor outcome; Comments: Baseline strength: 19.5 (4.3). Baseline control: 21.1 (4.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 10.8 (SD 1.8); n=27, Group 2: mean 20.7 (SD 4.4); n=26; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline strength: 25.4 (5.3). Baseline control: 25.2 (5.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 3 (SD 1.7); n=27, Group 2: mean 6 (SD 3.3); n=26; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline strength: 6.6 (3.2). Baseline control: 6.6 (3.5).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: Strength: 3 out of contact. ; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Nottingham Health Profile pain subscale at 12 weeks; Group 1: mean 9 (SD 3.3); n=28, Group 2: mean 20.4 (SD 3.2); n=26; Nottingham Health Profile pain subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 41.3 (4.1). Baseline control: 40.9 (4.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile energy subscale at 12 weeks; Group 1: mean 14.6 (SD 1.3); n=28, Group 2: mean 49.3 (SD 1.7); n=26; Nottingham Health Profile energy subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 50.7 (1.6). Baseline control: 52.9 (2.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.
- Actual outcome: Nottingham Health Profile physical mobility subscale at 12 weeks; Group 1: mean 8.6 (SD 4); n=28, Group 2: mean 36.6 (SD 6.1); n=26; Nottingham Health Profile physical mobility subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 41.3 (7.1). Baseline control: 44.1 (7.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.

- Actual outcome: Nottingham Health Profile emotional reactions subscale at 12 weeks; Group 1: mean 6.9 (SD 3); n=28, Group 2: mean 27.9 (SD 4.5); n=26; Nottingham Health Profile emotional reactions subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 28.6 (4.4). Baseline control: 28.0 (3.7). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication. - Actual outcome: Nottingham Health Profile sleep subscale at 12 weeks; Group 1: mean 19.6 (SD 4); n=28, Group 2: mean 35.3 (SD 4.4); n=26; Nottingham Health Profile sleep subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 44.9 (4.9). Baseline control: 45.0 (5.1). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication. - Actual outcome: Nottingham Health Profile social isolation subscale at 12 weeks; Group 1: mean 17.3 (SD 3.9); n=28, Group 2: mean 19.2 (SD 4.7); n=26; Nottingham Health Profile social isolation subscale 0-100 Top=High is poor outcome; Comments: Baseline aerobic: 21.0 (3.9). Baseline control: 21.1 (4.0). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication. Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 10.2 (SD 2.4); n=28, Group 2: mean 20.7 (SD 4.4); n=26; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline aerobic: 23.9 (6.3). Baseline control: 25.2 (5.9). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication. Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 3.4 (SD 1.3); n=28, Group 2: mean 6 (SD 3.3); n=26; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline aerobic: 6.9 (3.1). Baseline control: 6.6 (3.5). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, disease duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 2, Reason: Aerobic: 2 unwilling to participate because of taking extra medication.; Group 2 Number missing: 4, Reason: Control: 4 unwilling to participate because of taking extra medication.	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months;

	Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Fitzgerald 2011[129] (Teixeira 2011[440])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=183)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis meeting the 1986 American College of Rheumatology clinical criteria with grade 2 or greater radiographic changes in the tibiofemoral joint
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 50 years or older; morning stiffness of less than 30 minutes' duration; crepitus with active motion of the knee (eg, when squatting while weight bearing); tenderness on palpation of the bony margins of the joint; bony enlargement; and no palpable warmth. Individuals with patellofemoral joint radiographic changes were included provided that they had tibiofemoral radiographic changes as well
Exclusion criteria	Use of an assistive device for ambulation; reported a history of 2 or more falls within the previous year; were unable to ambulate a distance of 30.5m without an assistive device or need of a rest period; or reported severe visual problems; if they had undergone total knee arthroplasty; exhibited uncontrolled hypertension; had a history of cardiovascular disease; had neurological disorders that affected lower extremity function
Recruitment/selection of patients	Recruited through physician offices, community flyers, newspaper advertisements, and the University of Pittsburgh Arthritis Institute Registry
Age, gender and ethnicity	Age - Mean (SD): 64.0 (8.7). Gender (M:F): 61:122. Ethnicity: White = 161, Black = 17, Hispanic = 0, Asian = 2, Native American = 3
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (Hypertension = 86, back pain = 88, cancer = 27, congestive heart failure = 1, diabetes = 11, depression = 32, heart disease = 15, previous hip fracture = 2, kidney disease = 2, liver disease = 2, lung disease = 11, memory problems = 15, memory problems = 15, stomach ulcer = 14, stroke = 3). 4. Site of osteoarthritis: Knee osteoarthritis

Extra comments	Severity: Not stated explicitly. Kellgren and Lawrence grade 2 or more. Duration of symptoms: Median 5-10 years. Wasn't able to extract a 12 month physical function outcome due to typos on the standard and agility group 95% confidence intervals
Indirectness of population	No indirectness
Interventions	(n=91) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Lower extremity muscle stretching (quadriceps femoris, hamstring and calf muscle stretching) and strengthening (quad sets, supine straight leg raises, prone hip extensions, seated isometric knee extensions, single-leg leg presses, standing hamstring curls, and standing heel raises), long-sitting knee flexion and extension range of motion, and treadmill walking. All lower-extremity exercises were performed bilaterally. In addition, agility training techniques including: side stepping, braiding (lateral stepping combined with forward and backward crossover steps), front crossover steps during forward ambulation, back crossover steps during backward ambulation, shuttle walking (forward and backward walking to and from designated markers), and a drill requiring multiple changes in direction in which the therapist provided hand signals at random to prompt the individual to change direction during walking. The perturbation techniques incorporated the use of foam surfaces, tiltboards, and rollerboards to expose the individual's lower limbs and body to potentially destabilising forces. The participants attempted to maintain balance and control over the exercised lower extremity during the perturbations.. Duration 12 months. Concurrent medication/care: All participants also were instructed to continue a walking program of at least 30 minutes per day for at least 3 days a week for the home program. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Agility and perturbation). (n=92) Intervention 2: Exercise - Unsupervised strength exercise. Lower extremity muscle stretching (quadriceps femoris, hamstring and calf muscle stretching) and strengthening (quad sets, supine straight leg raises, prone hip extensions, seated isometric knee extensions, single-leg leg presses, standing hamstring curls, and standing heel raises), long-sitting knee flexion and extension range of motion, and treadmill walking. All lower-extremity exercises were performed bilaterally.All participants also were instructed to continue a walking program of at least 30 minutes per day for at least 3 days a week for the home program.. Duration 12 months. Concurrent medication/care: All participants also were instructed to continue a walking program of at least 30 minutes per day for at least 3 days a week for the home

	program. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant 1-R01-AR048760))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 2 months; Group 1: mean 15.2 (SD 11.5); n=84, Group 2: mean 12.8 (SD 11.1); n=75; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Reports mean (95% confidence intervals). Reported mixed modality: 12.8 (10.3-15.3). Reported strength: 15.2 (12.7-17.6). Baseline mixed modality: 19.5 (12.3). Baseline strength: 19.9 (11.9).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, body mass index, prior history of knee injury, years with arthritis, medication at baseline, comorbidities and baseline values of outcomes ; Group 1 Number missing: 16, Reason: 2 lost to follow up, 1 total hip arthroplasty, 1 refused further participation, 12 missed 2 month testing visit; Group 2 Number missing: 7, Reason: 1 lost for total knee arthroplasty, 1 due to illness, 5 missed 2 month testing visit

Protocol outcome 2: Pain at </=3 months
- Actual outcome: Knee pain numeric rating scale at 2 months; Group 1: mean 3.5 (SD 2.4); n=75, Group 2: mean 4.1 (SD 2.6); n=84; Numeric rating scale 0-10 Top=High is poor outcome; Comments: Reports mean (95% confidence intervals). Reported mixed modality: 3.5 (3.0-4.1). Reported strength: 4.1 (3.5-4.6). Baseline mixed modality: 4.7 (2.6). Baseline strength: 4.4 (2.4).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, body mass index, prior history of knee injury, years with arthritis, medication at baseline, comorbidities and baseline values of outcomes ; Group 1 Number missing: 16, Reason: 2 lost to follow up, 1 total hip arthroplasty, 1 refused further participation, 12 missed 2 month testing visit; Group 2 Number missing: 7, Reason: 1 lost for total knee arthroplasty, 1 due to illness, 5 missed 2 month testing visit

Protocol outcome 3: Pain at > 3 months
- Actual outcome: Knee pain numeric rating scale at 12 months; Group 1: mean 3.6 (SD 2.7); n=66, Group 2: mean 3.5 (SD 3.1); n=76; Numeric rating scale 0-10 Top=High is poor outcome; Comments: Reports mean (95% confidence intervals). Reported mixed modality: 3.6 (3.0-4.3). Reported strength: 3.5 (2.8-4.2). Baseline mixed modality: 4.7 (2.6). Baseline strength: 4.4 (2.4).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, race, weight, height, body mass index, prior history of knee injury, years with arthritis, medication at baseline, comorbidities and baseline values of outcomes ; Group 1 Number missing: 15, Reason: 2 months: 2 lost to follow up, 1 total hip arthroplasty, 1 refused further participation, 12 missed 2 month testing visit. 6 months: 5 lost to follow up, 1

unilcompartmental knee arthroplasty, 1 total hip arthroplasty, 2 illness, 3 missed 6 month testing visit. 12 months: 1 lost to follow up, 1 missed 12 month testing visit (numbers unclear, people appear to have multiple reasons for exclusion reported); Group 2 Number missing: 23, Reason: 2 months: 1 lost for total knee arthroplasty, 1 due to illness, 5 missed 2 month testing visit. 6 months: 3 lost to follow up, 2 unicompartmental knee arthroplasty, 1 illness, 2 refused further participation, 3 missed 6 month testing visit. 12 months: 3 lost to follow up, 3 unicompartmental knee arthroplasty, 2 total knee arthroplasty, 3 illness, 1 death, 3 refused further participation (numbers unclear, people appear to have multiple reasons for exclusion reported)	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Foley 2003[134]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=105)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiological diagnosis of osteoarthritis of the hip or knee, or both
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Community living adults aged 50 years and over with radiological diagnosis of osteoarthritis of the hip or knee, or both. Had to be able to read, write and speak English, give informed consent, provide their own transport, and attend the Repatriation General Hospital three times a week for six consecutive weeks
Exclusion criteria	Received physiotherapy or hydrotherapy in the past 6 weeks; were attending community exercise classes; had joint replacement surgery within the past 12 months or it was scheduled within the next 12 weeks; if there was an indication of cognitive impairment
Recruitment/selection of patients	Recruited from physiotherapy, orthopaedic, and rheumatology departments of the Repatriation General Hospital, the orthopaedic department of the Flinders Medical Centre, and from the community through an advertisement
Age, gender and ethnicity	Age - Mean (SD): 70.9 (8.8). Gender (M:F): 53:52. Ethnicity: No additional information
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee osteoarthritis).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=35) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Hydrotherapy - walking forwards, sideways and backwards through the water. The strengthening exercises included hip flexion and extension, hip adduction and abduction, knee flexion and extension, and knee cycling. One set of 10 repetitions was increased to three sets of 10 repetitions for each exercise, usually

	within the first week. Once three sets of 15 repetitions could be performed, weighted gaiters were fastened around the ankles to provide additional resistance. At this point, repetitions were dropped back to 10 and then increased to 15 as tolerated.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=35) Intervention 2: Exercise - Supervised strength exercise. Warm up with four minutes of stationary cycling. The strengthening exercises included seated bench press, hip adduction and abduction, knee extension, and double leg press. Participants started the programme working at either their 10 repetition maximum or just below their maximum, depending upon pain experienced during and after the initial exercise session. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=35) Intervention 3: No treatment. Fortnightly telephone calls to record any changes in their condition, drug use, or injuries and free exercise treatment at the hospital at the end of the study period. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical score at 12 weeks; Group 1: mean 37.1 (SD 12.7); n=35, Group 2: mean 31.4 (SD 12.7); n=35; SF-12 physical score 0-100 Top=High is good outcome; Comments: Baseline hydro: 31.4 (7.9). Baseline gym: 30.7 (11.2).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.

- Actual outcome: SF-12 mental score at 12 weeks; Group 1: mean 53.3 (SD 15.5); n=35, Group 2: mean 57.9 (SD 19.5); n=35; SF-12 mental score 0-100 Top=High is good outcome; Comments: Baseline hydro: 53.4 (15.7). Baseline gym: 51.8 (21.2).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 33 (SD 17); n=35, Group 2: mean 27 (SD 12); n=35; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline hydro: 34.0 (16.0). Baseline gym: 28.0 (13.0).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain score at 12 weeks; Group 1: mean 10 (SD 4); n=35, Group 2: mean 8 (SD 5); n=35; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline hydro: 10.0 (3.0). Baseline gym: 8.0 (4.0).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical score at 12 weeks; Group 1: mean 37.1 (SD 12.7); n=35, Group 2: mean 28.8 (SD 11); n=35; SF-12 physical score 0-100 Top=High is good outcome; Comments: Baseline hydro: 31.4 (7.9). Baseline control: 30.9 (11.2).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.
- Actual outcome: SF-12 mental score at 12 weeks; Group 1: mean 53.3 (SD 15.5); n=35, Group 2: mean 50.5 (SD 14); n=35; SF-12 mental score 0-100 Top=High is good outcome; Comments: Baseline hydro: 53.4 (15.7). Baseline control: 50.5 (16.9).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 33 (SD 17); n=35, Group 2: mean 37 (SD 13); n=35; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline hydro: 34.0 (16.0). Baseline control: 37.0 (17.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain score at 12 weeks; Group 1: mean 10 (SD 4); n=35, Group 2: mean 10 (SD 4); n=35; WOMAC pain score 0-20 Top=High is poor outcome; Comments: Baseline hydro: 10.0 (3.0). Baseline control: 10.0 (4.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMac pain and physical function. Otherwise similar.; Group 1 Number missing: 7, Reason: Hydro: 1 joint replacement surgery, 2 increased pain, 2 doctor advised to cease, 1 joint replacement surgery. 1 not coping with exercise.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical score at 12 weeks; Group 1: mean 31.4 (SD 12.7); n=35, Group 2: mean 28.8 (SD 11); n=35; SF-12 physical score 0-100 Top=High is good outcome; Comments: Baseline gym: 30.7 (11.2). Baseline control: 30.9 (11.2).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMAC pain and physical function. Otherwise similar.; Group 1 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.
- Actual outcome: SF-12 mental score at 12 weeks; Group 1: mean 57.9 (SD 19.5); n=35, Group 2: mean 50.5 (SD 14); n=35; SF-12 mental score 0-100 Top=High is good outcome; Comments: Baseline gym: 51.8 (21.2). Baseline control: 50.5 (16.9).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMAC pain and physical function. Otherwise similar.; Group 1 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean 27 (SD 12); n=35, Group 2: mean 37 (SD 13); n=35; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline gym: 28.0 (13.0). Baseline control: 37.0 (17.0).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMAC pain and physical function. Otherwise similar.; Group 1 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain score at 12 weeks; Group 1: mean 8 (SD 5); n=35, Group 2: mean 10 (SD 4); n=35; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline gym: 8.0 (4.0). Baseline control: 10.0 (4.0).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in gym baseline values for WOMAC pain and physical function. Otherwise similar.; Group 1 Number missing: 9, Reason: Gym: 1 joint replacement surgery, 2 withdrew consent after randomisation, 2 increased pain, 1 doctor advised to cease, 1 sick partner, 1 unrelated surgery, 1 increased blood pressure.; Group 2 Number missing: 3, Reason: Control: 2 joint replacement surgery, 1 illness.

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Fransen 2007[145]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=152)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of osteoarthritis involving the hip or knee as per American College of Rheumatology criteria and current and chronic hip or knee pain (at least 1 year)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 59-85 years, a diagnosis of osteoarthritis involving the hip or knee as per American College of Rheumatology criteria and current and chronic hip or knee pain
Exclusion criteria	Current participation in recreational physical activity more than twice per week; inability to walk indoors without a walking aide; unstable cardiac conditions or severe pulmonary disease; incontinence, fear of water, or uncontrolled epilepsy; low back pain referred to the lower limbs; joint replacement surgery in the previous year; arthroscopic surgery or intraarticular injections within previous 3 months; and current participation in Tai Chi or hydrotherapy
Recruitment/selection of patients	People were recruited via advertisements in local newspapers, through presentations at local social clubs for older persons, and through referral from local general practitioners and rheumatologists
Age, gender and ethnicity	Age - Mean (SD): 70.2 (6.3). Gender (M:F): 40:112. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: High morbidity score (Mean (SD): 4.7 (2.7)). 4. Site of osteoarthritis: Mixed (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: median 6-10 years
Indirectness of population	No indirectness
Interventions	(n=111) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Tai Chi or hydrotherapy (two groups pooled together). Exercises were performed for 1 hour, twice a week for 12 weeks. The Tai Chi program was a modification of 24 forms of the Sun style of Tai Chi and includes a preliminary 10-

	minute warm-up session. Participants were able to purchase, if they desired, a Tai Chi video to assist with home practice. Hydrotherapy was conducted under a standardized protocol.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Tai Chi in one group, hydrotherapy in another group). Comments: The two groups were combined in the same group for analysis as agreed in the protocol (n=41) Intervention 2: No treatment. Waiting list control. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not stated / Unclear 2. Group or individual : Not stated / Unclear 3. Type of exercise: Not stated / Unclear
Funding	Academic or government funding (Supported by a National Arthritis and Musculoskeletal conditions Improvements grant)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical component summary at 12 weeks; Group 1: mean 36.7 (SD 10.6); n=111, Group 2: mean 33.1 (SD 10.6); n=41; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Reported hydrotherapy: 35.7 (9.8). Reported tai chi: 37.6 (11.2). Reported control: 33.1 (10.6). Baseline hydrotherapy: 31.9 (8.5). Baseline tai chi: 35.6 (9.6). Baseline control: 33.2 (10.1).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections.
Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction
- Actual outcome: SF-12 mental component summary at 12 weeks; Group 1: mean 52.7 (SD 10); n=111, Group 2: mean 48 (SD 11.4); n=41; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Reported hydrotherapy: 54.6 (8.9). Reported tai chi: 50.9 (10.7). Reported control: 48.0 (11.4). Baseline hydrotherapy: 53.4 (11.1). Baseline tai chi: 50.9 (11.4). Baseline control: 47.7 (12.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections.

Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 35.7 (SD 22.4); n=111, Group 2: mean 49.9 (SD 19); n=41; WOMAC function 0-100 Top=High is poor outcome; Comments: Reported hydrotherapy: 34.8 (23.7). Reported tai chi: 36.6 (20.9). Reported control: 49.9 (19.0). Baseline hydrotherapy: 46.3 (20.4). Baseline tai chi: 47.2 (20.6). Baseline control: 50.8 (19.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections.
Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 29 (SD 18.9); n=111, Group 2: mean 40 (SD 16.2); n=41; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Reported hydrotherapy: 27.3 (18.7). Reported tai chi: 30.7 (18.9). Reported control: 40.0 (16.2). Baseline hydrotherapy: 38.2 (17.4). Baseline tai chi: 40.3 (19.0). Baseline control: 44.4 (17.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections.
Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction

Protocol outcome 4: Psychological distress at </=3 months
- Actual outcome: Depression, Anxiety and Stress Scale 21 - Depression subscale at 12 weeks; Group 1: mean 5.9 (SD 7.4); n=111, Group 2: mean 9 (SD 11); n=41; Depression, Anxiety and Stress Scale 21 - Depression subscale 0-42 Top=High is poor outcome; Comments: Reported hydrotherapy: 4.7 (6.1). Reported tai chi: 7.0 (8.3). Reported control: 9.0 (11.0). Baseline hydrotherapy: 6.8 (6.8). Baseline tai chi: 7.4 (8.5). Baseline control: 9.5 (10.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections.
Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction
- Actual outcome: Depression, Anxiety and Stress Scale 21 - Anxiety subscale at 12 weeks; Group 1: mean 4.9 (SD 5.6); n=111, Group 2: mean 7.3 (SD 7.8); n=41; Depression, Anxiety and Stress Scale 21 - Anxiety subscale 0-42 Top=High is poor outcome; Comments: Reported hydrotherapy: 4.6 (5.2). Reported tai chi: 5.1 (6.0). Reported control: 7.3 (7.8). Baseline hydrotherapy: 4.9 (6.3). Baseline tai chi: 5.5 (5.7). Baseline control: 6.9 (7.7).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections. Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction - Actual outcome: Depression, Anxiety and Stress Scale 21 - Stress subscale at 12 weeks; Group 1: mean 7.6 (SD 8.3); n=111, Group 2: mean 12.6 (SD 10.9); n=41; Depression, Anxiety and Stress Scale 21 - Stress subscale 0-42 Top=High is poor outcome; Comments: Reported hydrotherapy: 7.1 (8.0). Reported tai chi: 8.1 (8.6). Reported control: 12.6 (10.9). Baseline hydrotherapy: 9.5 (8.2). Baseline tai chi: 9.3 (8.4). Baseline control: 13.7 (9.7). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcomes are different at baseline, which is not significant, but the final values don't change too much from the baseline values and so amplifies the effect; Group 1 Number missing: 10, Reason: Hydrotherapy: 1 withdrew, 1 total knee replacement, 2 intraarticular injections. Tai Chi: 3 withdrew, 1 knee surgery, 1 total knee replacement, 1 unknown; Group 2 Number missing: 9, Reason: 7 withdrew, 1 knee surgery, 1 myocardial infarction	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	French 2013[146] (French 2009[147])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=131)
Countries and setting	Conducted in Irish Republic; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip according to the American College of Rheumatology clinical and radiographic criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the hip fulfilling the American College of Rheumatology clinical ad radiographic criteria aged 40-80 years
Exclusion criteria	Previous hip arthroplasty; congenital or adolescent hip disease; clinical signs of lumbar spine disease; physiotherapy in the previous 6 months for hip symptoms; pregnancy; hip fracture; contraindications to ET; inflammatory arthritis; on the waitlist for hip joint replacement within the next 7 months; intra-articular hip corticosteroid injection in the previous 30 days; insufficient understanding of the English language to complete questionnaires
Recruitment/selection of patients	People were referred from rheumatologists, orthopaedic surgeons, other hospital consultants, and general practitioners (from physiotherapy waitlists)
Age, gender and ethnicity	Age - Mean (SD): 62.5 (9.9). Gender (M:F): 47:84. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (2.2 (1.4)). 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 34.5 (45.5) months
Indirectness of population	No indirectness
Interventions	(n=45) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Exercise therapy - 6 to 8 individual 30 minute physiotherapy sessions over 8 weeks, which included flexibility and strengthening exercises delivered using a semi-structured protocol. Strengthening on low-load exercise, commencing in non-weight-bearing positions, and progressing to functional positions. A daily home exercise program supplemented the clinic based treatment.

	Participants were also encouraged to undertake aerobic exercise, such as walking, cycling or swimming for at least 30 minutes, 5 days a week, and were given written and verbal information on the principles of aerobic conditiong, such as pacing, gradually progressing intensity and time of exercise, and incorporating exercise into daily life.. Duration 18 weeks. Concurrent medication/care: All groups received standardised written information on hip osteoarthritis. All nonconsenting and excluded participants were treated as usual by the physiotherapy department in each trial center. Participants were asked to avoid all other interventions for the duration of the RCT, apart from routine doctor care and analgesics.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and flexibility). (n=43) Intervention 2: Other. Manual therapy with the exercise program. Duration 18 weeks. Concurrent medication/care: All groups received standardised written information on hip osteoarthritis. All nonconsenting and excluded participants were treated as usual by the physiotherapy department in each trial center. Participants were asked to avoid all other interventions for the duration of the RCT, apart from routine doctor care and analgesics.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group is not included in the analysis as it did not fulfill the inclusion criteria in the protocol (n=43) Intervention 3: No treatment. Waiting list for 8 weeks, then randomisation into the exercise or exercise and manual therapy groups at week 9. Duration 8 weeks (then re-randomised for the remaining 10 weeks). Concurrent medication/care: All groups received standardised written information on hip osteoarthritis. All nonconsenting and excluded participants were treated as usual by the physiotherapy department in each trial center. Participants were asked to avoid all other interventions for the duration of the RCT, apart from routine doctor care and analgesics.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by a Fellowship for the Therapy Professions from the Health Research Board, Ireland (grant no. CTPF-06-12))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND	

STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical component summary at 9 weeks; Group 1: mean 37.03 (SD 11.25); n=45, Group 2: mean 33.82 (SD 9.67); n=43; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 36.51 (9.87). Baseline control: 36.60 (9.32).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons; Group 2 Number missing: 0, Reason: 0 lost to follow up
- Actual outcome: SF-36 mental component summary at 9 weeks; Group 1: mean 48.92 (SD 12.5); n=45, Group 2: mean 48.52 (SD 13.75); n=43; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 52.78 (10.75). Baseline control: 52.82 (11.75).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons; Group 2 Number missing: 0, Reason: 0 lost to follow up

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 9 weeks; Group 1: mean 28.08 (SD 15.48); n=45, Group 2: mean 36.09 (SD 16.41); n=43; WOMAC physical function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 32.29 (12.21). Baseline control: 32.91 (15.22).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons; Group 2 Number missing: 0, Reason: 0 lost to follow up

Protocol outcome 3: Pain at </=3 months
- Actual outcome: Pain with activity (VAS) at 9 weeks; Group 1: mean 4.02 (SD 2.88); n=45, Group 2: mean 5.62 (SD 2.84); n=43; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.62 (2.63). Baseline control: 5.65 (2.46).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons; Group 2 Number missing: 0, Reason: 0 lost to follow up

Protocol outcome 4: Psychological distress at </=3 months
- Actual outcome: HADS anxiety at 9 weeks; Group 1: mean 6.74 (SD 4.27); n=45, Group 2: mean 6.14 (SD 3.97); n=43; HADS anxiety subscale 0-21 Top=High is poor outcome; Comments: Baseline exercise: 5.80 (3.35). Baseline control: 5.07 (3.37).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons;

Group 2 Number missing: 0, Reason: 0 lost to follow up - Actual outcome: HADS depression at 9 weeks; Group 1: mean 5.02 (SD 3.39); n=45, Group 2: mean 5.58 (SD 3.45); n=43; HADS depression 0-21 Top=High is poor outcome; Comments: Baseline exercise: 4.58 (2.95). Baseline control: 4.37 (2.92). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, body mass index, pain medications, referral source, occupation, hip affected, use of walking aids, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 decline, 1 family reasons; Group 2 Number missing: 0, Reason: 0 lost to follow up	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Gill 2009[153]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=82)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 7 weeks with 8 weeks additional follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People on the waiting list for joint replacement surgery or the hip or knee
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People on the waiting list for joint replacement surgery of the hip or knee.
Exclusion criteria	People were excluded if only tibial osteotomy was performed; if they were currently completing a physiotherapy program; if surgery was scheduled before completion of the 6-week supervised program; if they were not medically fit to complete an exercise program as assessed by their local doctor; if they had inadequate communication skills in English
Recruitment/selection of patients	People were recruited from the surgical waiting list of a tertiary health care provider in regional Victoria, Australia
Age, gender and ethnicity	Age - Mean (SD): 70.4 (9.8). Gender (M:F): 31:51. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee osteoarthritis).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Exercise - Supervised strength exercise. Land based exercises completed in groups of 4 to 6 participants conducted twice a week for 6 weeks with each session lasting 1 hour. Exercises were completed at a moderate intensity between 12 and 14 on the Borg Rating of Perceived Exertion Scale. The land-based exercise sessions were held in a physiotherapy gymnasium. Exercise included: 5 to 10 minutes of forward, sideways and backward walking; 20 minutes pedaling a stationary exercise bike; resistance exercises; calf, hamstrings and quadriceps stretches (2 sets of 30 seconds).. Duration 6 weeks (then assessed again 8 weeks

	after that). Concurrent medication/care: All participants were asked to complete 30 minutes of exercise at home, 3 times a week, including walking, riding a stationary bike, or performing exercises similar to those completed in the supervised classes. After the 6-week intervention, participants were encouraged to continue to exercise at home until the follow-up assessment. All people received a home visit and environmental assessment from an occupational therapist, similar to that described by others.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=42) Intervention 2: Exercise - Other supervised exercise (including flexibility, proprioception). Water based exercises completed in groups of 4 to 6 participants conducted twice a week for 6 weeks with each session lasting 1 hour. Exercises were completed at a moderate intensity between 12 and 14 on the Borg Rating of Perceived Exertion Scale. The water-based exercise sessions were held in a hydrotherapy center. Exercises included: walking and active range of motion exercises; calf, hamstring, and quadriceps stretches (2 sets of 30 seconds).. Duration 6 weeks (then assessed again 8 weeks after that). Concurrent medication/care: All participants were asked to complete 30 minutes of exercise at home, 3 times a week, including walking, riding a stationary bike, or performing exercises similar to those completed in the supervised classes. After the 6-week intervention, participants were encouraged to continue to exercise at home until the follow-up assessment. All people received a home visit and environmental assessment from an occupational therapist, similar to that described by others.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy
Funding	Academic or government funding (Supported by Barwon Health, Australia, and the Department of Human Services, Victoria, Australia)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: SF-36 mental component scale at 7 weeks; Group 1: mean 50.6 (SD 11.2); n=32, Group 2: mean 55.7 (SD 9.3); n=34; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline hydro: 49.5 (10.8). Baseline land: 52.1 (9.6). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous	

joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Overall rate reported only. Group 1 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons, 2 unknown reasons). Before week 9 another 9 withdrew due to surgery.

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: SF-36 mental component scale at 15 weeks; Group 1: mean 51.2 (SD 10.5); n=32, Group 2: mean 51.9 (SD 12.1); n=34; SF-36 mental component scale 0-100 Top=High is good outcome; Comments: Baseline hydro: 49.5 (10.8). Baseline land: 52.1 (9.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Group 1 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons). Before week 9 another 9 withdrew due to surgery.

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: WOMAC function at 7 weeks; Group 1: mean 32.3 (SD 10.4); n=32, Group 2: mean 29.2 (SD 12.7); n=34; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline hydro: 36.0 (10.3). Baseline land: 36.9 (12.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Group 1 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: WOMAC function at 15 weeks; Group 1: mean 32.6 (SD 10.7); n=32, Group 2: mean 32.2 (SD 12.4); n=34; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline hydro: 36.0 (10.3). Baseline land: 36.9 (12.9).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Group 1 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.

Protocol outcome 5: Pain at </=3 months
- Actual outcome: WOMAC pain at 7 weeks; Group 1: mean 10.1 (SD 2.9); n=32, Group 2: mean 9.2 (SD 3.7); n=34; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline hydro: 11 (3.7). Baseline land: 11.6 (3.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Group 1 Number missing: -, Reason: Before week 7, 7

withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery. Protocol outcome 6: Pain at > 3 months - Actual outcome: WOMAC pain at 15 weeks; Group 1: mean 10.3 (SD 3.4); n=32, Group 2: mean 10 (SD 2.3); n=34; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline hydro: 11 (3.7). Baseline land: 11.6 (3.0). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, surgery type, previous joint replacement, BMI, joint pathology, current exercise program and baseline values of outcomes; Group 1 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.; Group 2 Number missing: -, Reason: Before week 7, 7 withdrew the study overall (3 for comorbidities, 2 for family reasons. Before week 9 another 9 withdrew due to surgery.	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Gomiero 2018[157]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=64)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 16 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis of the American College of Rheumatology with radiographic confirmation
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a diagnosis of tibiofemoral osteoarthritis that fulfilled the clinical criteria for knee osteoarthritis of the American College of Rheumatology (ACR), 1986; be between 50 and 75 years of age; have not done any physical activity for at least 3 months; and have reached a minimum educational level of 4th grade of elementary education.
Exclusion criteria	Uncontrolled arterial hypertension; decompensated diabetes mellitus; decompensated thyroid diseases; cardiorespiratory diseases (ischaemia, arrhythmia, precordial pain or physical exercise-induced bronchospasm); liver abnormalities; grade 4 functional impairment (Kellgren-Lawrence radiographic scale); or other rheumatic diseases; people who needed ambulatory devices; those who were on sick leave from work approved by the government agency for national insurance; any other related factor
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 61.7 (6.6). Gender (M:F): 3:61. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren and Lawrence grade 1-4, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=32) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Sensory-motor training over 16 weeks with exercise twice a week. The program included walking in different directions following verbal commands from the therapist; crossing steps while walking; crossing steps while walking backwards;

	implementing sudden changes of direction; walking on several types of surfaces (including mattresses); maintaining posture during use of a balance board; and using a mini-trampoline to expose individuals to potentially destabilising loads. Duration 16 weeks. Concurrent medication/care: Both groups had concomitant interventions such as informative talks. They also received an educational program on knee osteoarthritis, which allowed the people to clarify their doubts and concerns about the disease.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Neuromodulatory (Sensory-motor training). (n=32) Intervention 2: Exercise - Supervised strength exercise. Exercises twice a week for 16 weeks including quadriceps and hamstring strengthening exercises using ankle weights, isometric exercises for the quadriceps muscle (hip flexion with leg extended) and stretching for the lower limbs (stretching of the quadriceps, hamstrings and triceps surae). All physical exercises were performed bilaterally and at a volume of three sets of ten maximal repetitions. Duration 16 weeks. Concurrent medication/care: Both groups had concomitant interventions such as informative talks. They also received an educational program on knee osteoarthritis, which allowed the people to clarify their doubts and concerns about the disease.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: SF-36 physical role functioning at 16 weeks; Group 1: mean 54.8 (SD 24.6); n=32, Group 2: mean 51.4 (SD 25.5); n=32; SF-36 physical role functioning 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 54.8 (46.26-63.34). Reported strength exercise: 51.4 (42.57-60.23). Baseline other exercise: 51.4 (43.22-59.58). Baseline strength exercise: 38.3 (31.70-44.90).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 physical functioning at 16 weeks; Group 1: mean 57.5 (SD 43.3); n=32, Group 2: mean 50.8 (SD 38.2); n=32; SF-36 physical functioning 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 57.5 (42.50-72.50). Reported

strength exercise: 50.8 (37.57-64.03). Baseline other exercise: 32.8 (19.57-46.03). Baseline strength exercise: 30.5 (16.62-44.38).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 bodily pain at 16 weeks; Group 1: mean 59.3 (SD 26.1); n=32, Group 2: mean 54.8 (SD 28.8); n=32; SF-36 bodily pain 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 59.3 (50.14-68.46). Reported strength exercise: 54.8 (45.75-63.85). Baseline other exercise: 50.4 (40.41-60.39). Baseline strength exercise: 48.0 (38.81-57.19).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 general health perceptions at 16 weeks; Group 1: mean 60.8 (SD 20); n=32, Group 2: mean 62 (SD 21.4); n=32; SF-36 general health perceptions 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 60.8 (53.88-67.72). Reported strength exercise: 62.0 (54.57-69.43). Baseline other exercise: 55.8 (48.23-63.37). Baseline strength exercise: 54.8 (45.75-63.85).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 vitality at 16 weeks; Group 1: mean 64.5 (SD 17.6); n=32, Group 2: mean 60.3 (SD 20.7); n=32; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 64.5 (58.41-70.59). Reported strength exercise: 60.3 (53.13-67.47). Baseline other exercise: 55.6 (47.99-63.21). Baseline strength exercise: 46.4 (38.72-54.08).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 vitality at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 social role functioning at 16 weeks; Group 1: mean 74 (SD 23.7); n=32, Group 2: mean 67.3 (SD 27.2); n=32; SF-36 social role functioning 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 74.0 (65.78-82.22). Reported strength exercise: 67.3 (57.89-76.71). Baseline other exercise: 72.8 (62.63-82.97). Baseline strength exercise: 70.8 (59.80-81.80).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0
- Actual outcome: SF-36 emotional role functioning at 16 weeks; Group 1: mean 61.1 (SD 42.9); n=32, Group 2: mean 64.6 (SD 42.3); n=32; SF-36 emotional role functioning 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 61.1 (46.25-75.96). Reported strength exercise: 64.6 (49.96-79.24). Baseline other exercise: 34.7 (19.38-50.02). Baseline strength exercise: 28.1 (14.76-41.44).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for

SF-36 emotional role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0 - Actual outcome: SF-36 mental health at 16 weeks; Group 1: mean 74.1 (SD 17); n=32, Group 2: mean 65.6 (SD 19.8); n=32; SF-36 mental health 0-100 Top=High is good outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 74.1 (68.22-79.98). Reported strength exercise: 65.6 (58.75-72.45). Baseline other exercise: 65.2 (57.63-72.77). Baseline strength exercise: 28.1 (14.76-41.44). Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 mental health at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0 Protocol outcome 2: Pain at > 3 months - Actual outcome: Visual analogue scale at 16 weeks; Group 1: mean 4.6 (SD 2.2); n=32, Group 2: mean 4.1 (SD 2.7); n=32; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Reports mean and 95% confidence intervals. Reported other exercise: 4.6 (3.84-5.36). Reported strength exercise: 4.1 (3.16-5.04). Baseline other exercise: 6.3 (5.47-7.13). Baseline strength exercise: 6.7 (5.80-7.60). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 2, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0 Protocol outcome 3: Serious adverse events at > 3 months - Actual outcome: Low back pain at 16 weeks; Group 1: 1/32, Group 2: 0/32 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Strength exercise has a much lower value for SF-36 physical role functioning at baseline (meaning it has a much larger gain than the control group); Group 1 Number missing: 1, Reason: 1 lost to follow up (moved to a different city), 1 discontinued due to low back pain; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	Gondhalekar 2013[158]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=33)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People fulfilling three of the six clinical criteria listed by the American College of Rheumatology diagnosed as knee osteoarthritis confirmed using radiological investigations
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Three out of the six of: age >50 years, morning stiffness lasting <30 min, crepitus with active motion, bony tenderness, bony enlargement, no warmth to touch. People with knee pain for more than 6 weeks.
Exclusion criteria	People with bilateral involvement; a history of any lower extremity injury or underlying pathology; a history of any inflammatory joint disease and balance problems; using an assistive device for ambulation
Recruitment/selection of patients	Outpatients referred by a physician or an orthopedic surgeon to the aforementioned departments for acute knee pain were screened for knee osteoarthritis (as people often present with an acute exacerbation of chronic problems in osteoarthritis)
Age, gender and ethnicity	Age - Mean (SD): 64.43 (6.202). Gender (M:F): 15:15. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). People underwent three sessions of Retro-walking per day (10 minutes per session) for 3 weeks on a flat surface at their maximum pace and free exercises (static and dynamic quadriceps, knee bending exercise in prone lying, hip flexion exercise in supine, hip abduction in side lying and hip extension in prone lying position). All exercises were done in sets of 10 repetitions; 1 set of all

	exercises twice a day for the 1st week and progressed to 2 sets twice a day in the second week and 3 sets twice a day in the 3rd week.. Duration 3 weeks. Concurrent medication/care: Deep heating modality (short wave diathermy) 250W for 20 minutes. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and aerobic). (n=15) Intervention 2: Exercise - Unsupervised strength exercise. Free exercises (static and dynamic quadriceps, knee bending exercise in prone lying, hip flexion exercise in supine, hip abduction in side lying and hip extension in prone lying position). All exercises were done in sets of 10 repetitions; 1 set of all exercises twice a day for the 1st week and progressed to 2 sets twice a day in the second week and 3 sets twice a day in the 3rd week.. Duration 3 weeks. Concurrent medication/care: Deep heating modality (short wave diathermy) 250W for 20 minutes. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	No funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus UNSUPERVISED STRENGTH EXERCISE Protocol outcome 1: Pain at </=3 months - Actual outcome: Visual analogue scale at 3 weeks; Group 1: mean 4.07 (SD 1.18); n=15, Group 2: mean 3.53 (SD 1.33); n=15; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline mixed modality: 7.53 (1.06). Baseline strength: 7.70 (0.99). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Only reports the baseline values for gender and outcomes; Group 1 Number missing: 0, Reason: Reports 33 people were included, but 3 were lost to follow up. No information about which groups that were in.; Group 2 Number missing: 0, Reason: Reports 33 people were included, but 3 were lost to follow up. No information about which groups that were in.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Hennig 2015[173]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in Norway; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Hand osteoarthritis diagnosed by rheumatologists or orthopaedic surgeons according to the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Females with hand osteoarthritis diagnosed by rheumatologists or orthopaedic surgeons according to the American College of Rheumatology criteria, age between 18 and 80 years, stable medication over the past 3 months, a minimum of three self-reported hand osteoarthritis-related activity limitations identified by people in the Patient-Specific Functional Scale and ability to communicate in Norwegian
Exclusion criteria	Hand surgery within the past 6 months; steroid injections within the past 2 weeks; impaired hand function due to trauma or diseases other than hand osteoarthritis; cognitive or mental impairment; people who received steroid injections in the trial period; people who underwent hand surgery
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.8 (7.0). Gender (M:F): 0:80. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: High morbidity score (32 people had comorbidities). 4. Site of osteoarthritis: Hand osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (median [range]): 10.0 (0-40) years
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Exercise - Unsupervised strength exercise. Home-based hand exercise programme aimed at maximising the stable and pain-free functional range of motion of the finger joints, increasing grip strength, maintaining joint stability and preventing or delaying development of fixed deformities. A rubber ball made of polyethylene with a diameter of 7cm was used to provide resistance in the grip

	strengthening exercise, while rubber bands were used to provide resistance to the thumb abduction/extension exercise. Participants were instructed to perform three exercise sessions a week, with each exercise to be performed with 10 repetitions during the first 2 weeks, increasing to 12 repetitions over the next 2 weeks and if possible, to 15 repetitions for the rest of the 3 month exercise period. Duration 12 weeks. Concurrent medication/care: All participants received a leaflet containing information about hand osteoarthritis and ergonomic principles. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=40) Intervention 2: No treatment. Leaflet only. Duration 12 weeks. Concurrent medication/care: All participants received a leaflet containing information about hand osteoarthritis and ergonomic principles. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Funding from Martina Hansens Hospital, Norway, the Norwegian Association for Rheumatism, the Norwegian Association of Hand Therapists and the Norwegian Association for Occupational Therapy)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: Functional Index for Hand Osteoarthritis at 12 weeks; Group 1: mean -2.2 (SD 5.8); n=40, Group 2: mean 1.7 (SD 2.6); n=40; Functional Index for Hand Osteoarthritis 0-30 Top=High is poor outcome; Comments: Reported mean change and 95% confidence intervals. Reported exercise: -2.2 (-4.0 to -0.4). Reported control: 1.7 (0.8 to 2.4). Only provides median baseline values.
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports (albeit with median values) age, education, living alone, occupation, hand dominance, duration of symptoms/disease, comorbidity, disease activity and baseline values of outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up - 1 excluded due to surgery, 1 excluded due to other disease, 1 discontinued due to sustained pain; Group 2 Number missing: 5, Reason: 5 lost to follow up - 2 excluded due to debut of other rheumatic disease, 3 drop out

Protocol outcome 2: Pain at </=3 months
- Actual outcome: Joint pain (NRS) at 12 weeks; Group 1: mean -1.1 (SD 2.6); n=40, Group 2: mean 0.3 (SD 1.6); n=40; NRS 0-10 Top=High is poor outcome; Comments: Reported mean change and 95% confidence intervals. Reported exercise: -1.1 (-1.9 to -0.3). Reported control: 0.3 (-0.2 to 0.8). Only provides median baseline values.
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports (albeit with median values) age, education, living alone, occupation, hand dominance, duration of symptoms/disease, comorbidity, disease activity and baseline values of outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up - 1 excluded due to surgery, 1 excluded due to other disease, 1 discontinued due to sustained pain; Group 2 Number missing: 5, Reason: 5 lost to follow up - 2 excluded due to debut of other rheumatic disease, 3 drop out	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Henriksen 2014[177] (Bartholdy 2016[34])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical diagnosis of tibiofemoral osteoarthritis confirmed by radigopraphy
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults age at least 40 years with a clinical diagnosis of tibiofemoral osteoarthritis confirmed by radiography assessed by an experienced radiologist, and a body mass index between 20 and 35 kg/m².
Exclusion criteria	Participation in exercise therapy within the previous 3 months; systemic inflammatory and autoimmune diseases; lower extremity joint replacement; significant cardiovascular, neurologic or psychiatric disease; cervical or lumbar nerve root compression syndromes; widespread or regional pain syndromes (e.g., fibromyalgia)
Recruitment/selection of patients	Participants were recruited from the osteoarthritis outpatient clinical of Copenhagen University Hospital at Frederiksberg, Copenhagen through advertisements in newspapers
Age, gender and ethnicity	Age - Mean (SD): 63.7 (8.2). Gender (M:F): 19:48. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=31) Intervention 1: Exercise - Supervised strength exercise. Facility based function and individualised exercise therapy supervised by a physiotherapist 3 times weekly for 12 weeks. The exercise was group based and the participants consecutively joint the group as they were included in the study. The exercise program lasted approximately 1 hour and consisted of a 10-minute warm up phase (bicycle ergometer at moderate intensity) followed by a circuit training program focusing on strength and coordination

	exercises of the trunk, hips and knees. The exercises were performed with free weights, elastic rubber bands, or body weight as resistance. Progression of resistance or coordination difficult was made on an individual basis according to a prespecified progression protocol. The level of each exercise, including external load, number of repetitions, or duration was recorded for each person at each visit in a personal training diary together with current knee pain before an exercise session on a 0-10 numeric rating scale. If a participant reported symptomatic exacerbation upon attending an exercise session (defined as current knee pain exceeding a score of 5) a "rescue" training program was applied for that session, including only warm-up, trunk, and hip exercises repeated twice. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=29) Intervention 2: No treatment. No attention control. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by the Danish Council for Independent Research, Medical Sciences (grant 10-093704), the Danish Physiotherapists Association, the Lundbeck Foundation and the Oak Foundation)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: KOOS quality of life at 12 weeks; Group 1: mean 5.8 (SD 14.2); n=25, Group 2: mean -0.3 (SD 14.3); n=23; KOOS quality of life 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 5.8 (0.3, 11.4). Reported control: -0.3 (-6.2, 5.5). Baseline exercise: 37.0 (14.4). Baseline control: 44.8 (15.4) Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, body mass index, cuff pressure algometry, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost to follow up; Group 2 Number missing: 5, Reason: 5 lost to follow up Protocol outcome 2: Physical function at </=3 months	

- Actual outcome: KOOS function in daily living at 12 weeks; Group 1: mean 4.2 (SD 11); n=25, Group 2: mean 1.4 (SD 11); n=23; KOOS function in daily living 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 4.2 (-0.0, 8.5). Reported control: 1.4 (-3.1, 5.9). Baseline exercise: 64.4 (15.0). Baseline control: 74.0 (14.2). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, body mass index, cuff pressure algometry, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost to follow up; Group 2 Number missing: 5, Reason: 5 lost to follow up Protocol outcome 3: Pain at </=3 months - Actual outcome: KOOS pain at 12 weeks; Group 1: mean 6.1 (SD 9.4); n=25, Group 2: mean -0.7 (SD 9.5); n=23; KOOS pain 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 6.1 (2.4, 9.8). Reported control: -0.7 (-4.6, 3.2). Baseline exercise: 56.4 (13.6). Baseline control: 61.2 (11.2). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, height, weight, body mass index, cuff pressure algometry, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 lost to follow up; Group 2 Number missing: 5, Reason: 5 lost to follow up	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Hermann 2016[178]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with primary hip osteoarthritis scheduled for total hip arthroplasty
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with primary hip osteoarthritis scheduled for total hip arthroplasty
Exclusion criteria	Rheumatoid arthritis and other types of arthritis not diagnosed as osteoarthritis; uraemia; cancer; treatment with systemic glucocorticoids >3 months the last 5 years with a dose of at least 5mg; present or previous hip fracture (either side); other lower extremity fracture within 1 year prior to inclusion; body weight >135kg; severe walking deficits (dependency of two crutches or walker for mobilization); not speaking Danish language
Recruitment/selection of patients	People diagnosed and scheduled for surgery by hip surgeons in the Department of Orthopaedic Surgery, Herlev University Hospital, Copenhagen, Denmark
Age, gender and ethnicity	Age - Mean (SD): 70.4 (7.6). Gender (M:F): 38:52. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity: Not stated explicitly. On the waiting list for surgery. Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Exercise - Supervised strength exercise. Supervised progressive explosive-type resistance training program twice a week for 10 weeks. Each session lasted 1 hour, ten minutes of warm up on a stationary bike was followed by a random circle sequence of four resistance training exercises performed unilaterally on training machines: hip extension performed in forward standing position and knee extension, knee flexion and leg press in a seated position. Exercises were executed in three series of 8-12 repetitions each. To apply with the principles of

	explosive-type resistance training the participants were instructed to complete the concentric phase of the movement 'as fast as possible', then pause briefly, and complete the eccentric phase of hte movement in approximately 2-3 seconds. The participants were encouraged to perform the maximum number of repetitions possible within each series. If the number was below 8 or exceeded 12, the loading was adjusted for the next series. The individual progression for each participant was supervised by experienced physiotherapists. Training groups consisted of up to 8 participants with 2 physiotherapists.. Duration 10 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=40) Intervention 2: No treatment. Standardised preoperative information only (no attention control). Duration 10 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The study was conducted with financial support by a research grant from the Danish Rheumatism Association (project no: R87-A1408))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: HOOS quality of life at 10 weeks; Group 1: mean 38.8 (SD 17.2); n=40, Group 2: mean 31.2 (SD 13.9); n=40; HOOS quality of life 0-100 Top=High is good outcome; Comments: Baseline exercise: 32.1 (14.4). Baseline control: 29.2 (15.6).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 lost to follow up. 1 discontinued due to medical illness not related to the study.; Group 2 Number missing: 2, Reason: 1 unwilling to participate in the follow-up due to test-related time consumption. 1 lost to follow up.

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: HOOS activities of daily living function at 10 weeks; Group 1: mean 59.9 (SD 17.1); n=40, Group 2: mean 48.7 (SD 13.9); n=40; HOOS activities of daily living function 0-100 Top=High is good outcome; Comments: Baseline exercise: 49.2 (12.5). Baseline control: 48.1 (13.8).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 lost to follow up. 1 discontinued due to medical illness not related to the study.; Group 2 Number missing: 2, Reason: 1 unwilling to participate in the follow-up due to test-related time consumption. 1 lost to follow up.

Protocol outcome 3: Pain at </=3 months - Actual outcome: HOOS pain at 10 weeks; Group 1: mean 55.4 (SD 16.9); n=40, Group 2: mean 45.9 (SD 14.1); n=40; HOOS pain 0-100 Top=High is good outcome; Comments: Baseline exercise: 48.0 (12.7). Baseline control: 46.3 (14.4). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: 2 lost to follow up. 1 discontinued due to medical illness not related to the study.; Group 2 Number missing: 2, Reason: 1 unwilling to participate in the follow-up due to test-related time consumption. 1 lost to follow up.	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Hernandez 2019[179]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=113)
Countries and setting	Conducted in Argentina; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks of treatment and 3 months of additional follow up (6 months in total)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Medical diagnosis of knee osteoarthritis referred by the Orthopedics Department to the Physical Therapy Department of Hospital Durand. Confirmed by an orthopedist based on radiographic and clinical findings.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People over 40 years of age; had consulted for knee pain and/or difficulty in activities of daily living - such as climbing or descending stairs, walking, getting up from a chair or kneeling - over the previous month.
Exclusion criteria	Patients with a history of intraarticular knee fracture; hip osteoarthritis; lower limb joint replacement; inflammatory arthritis; spine surgery; lower limb surgery within the prior 6 months; corticoid injection within the prior 3 months; physical limitations to

	exercise; illiterate patients and/or patients with apparent communication difficulties; people with a diagnosis other than knee osteoarthritis (such as knee sprain or Baker's cyst), even when their radiographs showed degenerative symptoms, or those with a diagnosis of knee osteoarthritis whose clinical evaluation by the physical therapist at baseline was not consistent with knee osteoarthritis based on age, history and physical examination.
Recruitment/selection of patients	Carried out between July 2011 and January 2015. Consecutive patients with medical diagnosis of knee osteoarthritis referred by the Orthopedics Department to the Physical Therapy Department of Hospital Durand (Buenos Aires City).
Age, gender and ethnicity	Age - Mean (SD): 62.3 (10.6). Gender (M:F): 14:33. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of symptoms (median [range]): Experimental group = 11.5 (1-120), control group = 8.5 (1-72) (units unclear)
Indirectness of population	No indirectness
Interventions	(n=53) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Conventional exercises plus exercises aimed at the activation of the muscles considered important for core stability according to electromyography tests. Treatments were delivered in triweekly sessions for three months.. Duration 3 months. Concurrent medication/care: All groups were offered conventional exercises including warm up and mobility as well as strengthening and stretching exercises.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Specific muscular exercises to increase core stability and strengthening exercises). (n=60) Intervention 2: Exercise - Supervised strength exercise. No additional exercises. Duration 3 months. Concurrent medication/care: All groups were offered conventional exercises including warm up and mobility as well as strengthening and stretching exercises.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function short-form at 3 months; MD; 9.08 (P value: <0.01) WOMAC function 0-68 Top=High is poor outcome, Comments: Baseline values not clear. From graph, experimental = 17.8, control = 17.47.;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related

problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: WOMAC function short-form at 6 months; MD; 8.73 (P value: <0.01) WOMAC function 0-68 Top=High is poor outcome, Comments: Baseline values not clear. From graph, experimental = 17.8, control = 17.47.;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason

Protocol outcome 3: Pain at </=3 months
- Actual outcome: Visual analogue scale at 3 months; Group 1: mean 2.42 (SD 2.35); n=25, Group 2: mean 4 (SD 2.83); n=22; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline supervised mixed modality exercise: 6.92 (2.6). Baseline supervised strength exercise: 6.11 (2.11).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason

Protocol outcome 4: Pain at > 3 months
- Actual outcome: Visual analogue scale at 6 months; Group 1: mean 3.8 (SD 2.97); n=23, Group 2: mean 3.63 (SD 2.8); n=20; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline supervised mixed modality exercise: 6.92 (2.6). Baseline supervised strength exercise: 6.11 (2.11).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason

Protocol outcome 5: Serious adverse events at </=3 months
- Actual outcome: Adverse events at 3 months; Group 1: 0/53, Group 2: 0/60
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason

Protocol outcome 6: Serious adverse events at > 3 months - Actual outcome: Adverse events at 6 months; Group 1: 0/53, Group 2: 0/60 Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, BMI, time of evolution of symptoms, knee affected, NSAID use and walking aids.; Group 1 Number missing: 28, Reason: 28 eliminated: 7 transportation problems, 6 work-related problems, 4 family-related problems, 1 not satisfied, 10 lost to follow-up/unknown reason; Group 2 Number missing: 38, Reason: 38 eliminated: 6 transportation problems, 5 work-related problems, 7 family-related problems, 3 health problems, 17 lost to follow-up/unknown reason	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months

Study	Hinman 2007[180]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=61)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis was based on American College of Rheumatology classification criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Volunteers aged 50 years and older with hip osteoarthritis or knee osteoarthritis. Participants with knee osteoarthritis were included if they had knee pain on most days of the previous month and osteophytes on radiographs. Participants with hip osteoarthritis were included if they had hip pain and osteophytes and joint space narrowing on radiographs. Other inclusion criteria for all participants were an average severity of pain of greater than 3 cm on a 10cm visual analogue scale and difficulty with stair climbing, walking, or getting in or out of a chair.
Exclusion criteria	Contraindications to aquatic physical therapy; significant back or other joint pain; recent (preceding 6 months) joint injections, surgery, physical therapy, or hydrotherapy; lower-limb joint replacement; inability to understand English; and inability to safely enter and exit the pool
Recruitment/selection of patients	Recruitment by advertisements in local clubs, libraries, general practitioner's rooms, print and radio media, and the orthopedic clinic at a metropolitan hospital
Age, gender and ethnicity	Age - Mean (SD): 62.4 (8.8). Gender (M:F): 23:48. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Knee or hip).
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 8 (10.0) years
Indirectness of population	No indirectness
Interventions	(n=36) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic physical therapy program comprised of functional weight bearing and progressive exercises provided twice weekly (45-60 minutes each) for 6

	weeks. An experienced aquatic physical therapist individually instructed participants in the hydrotherapy pool with a maximum of 6 participants per session. Quality of movement was emphasized, and the therapist palpated the lower limb musculature to ensure appropriate contraction throughout the exercises. Balance without the aid of rails to maximize postural and isometric leg stance control was achieved with all participants. A neutral spinal position also was taught; feedback was provided on posture, transversus abdominis muscle contraction, and trunk control. Individual progression to subsequent phases of the program was clinically determined by the therapist and occurred upon completion of the prior phase with either no or minimal symptom exacerbation. Upon completion of the 6 week program, participants were encouraged to continue independent aquatic physical therapy twice weekly at a local pool and were provided with details of local pools and a written description of the exercises to maximize adherence.. Duration 6 weeks for supervised component, 6 weeks of self-directed exercise. Concurrent medication/care: People continued using their usual medication. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=35) Intervention 2: No treatment. No treatment for 6 weeks, then completed the aquatic physical therapy program over the next 6 weeks. Duration 6 weeks (then received treatment for 6 weeks, will not be including data after 6 weeks due to this). Concurrent medication/care: People continued using their usual medication. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The study was supported by a National Arthritis and Musculoskeletal Conditions Improvement Grant from the Australian Government Department of Health and Aging)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Assessment of Quality of Life at 6 weeks; Group 1: mean 0.43 (SD 0.2); n=36, Group 2: mean 0.5 (SD 0.2); n=35; AQoL -0.04-1.00 Top=High is good outcome; Comments: Baseline exercise: 0.38 (0.17). Baseline control: 0.52 (0.20).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in AQoL and WOMAC function

scores at baseline; Group 1 Number missing: 1, Reason: 1 dropout due to stress; Group 2 Number missing: 4, Reason: 4 dropouts, 1 due to acute disk prolapse, 1 for lack of time, 2 for family illness Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC function at 6 weeks; Group 1: mean 598 (SD 316); n=36, Group 2: mean 656 (SD 373); n=35; WOMAC function 0-1700 Top=High is poor outcome; Comments: Baseline exercise: 757 (327). Baseline control: 630 (315). Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in AQoL and WOMAC function scores at baseline; Group 1 Number missing: 1, Reason: 1 dropout due to stress; Group 2 Number missing: 4, Reason: 4 dropouts, 1 due to acute disk prolapse, 1 for lack of time, 2 for family illness Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 143 (SD 79); n=36, Group 2: mean 198 (SD 108); n=35; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline exercise: 202 (79). Baseline control: 199 (85). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in AQoL and WOMAC function scores at baseline; Group 1 Number missing: 1, Reason: 1 dropout due to stress; Group 2 Number missing: 4, Reason: 4 dropouts, 1 due to acute disk prolapse, 1 for lack of time, 2 for family illness	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	**Holm 2020[183] (Holm 2021[182])**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=90)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks

Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic and radiographic (Kellgren and Lawrence at least 2) knee osteoarthritis deemed ineligible for knee replacement surgery by orthopedic surgeons in the orthopedic outpatient clinic at Naestved Hospital.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with symptomatic and radiographic (Kellgren and Lawrence at least 2) knee osteoarthritis deemed ineligible for knee replacement surgery by orthopedic surgeons in the orthopedic outpatient clinic at Naestved Hospital. Specifically, patients who had been assessed by an orthopedic surgeon and deemed ineligible for knee replacement surgery were approached by study staff at the orthopedic outpatient clinic and invited to take part in this study. The decision to not list patients for surgery was based on a combination of criteria, which primarily included radiographic severity, symptomatic severity and the patient's willingness to undergo surgery.
Exclusion criteria	Less than "mild" symptoms (score >75 in 0-100) on the subscale activities of daily living from the Knee Injury and Osteoarthritis Outcome Score; morphine usage for pain other than knee joint pain; previous ipsilateral knee arthroplasty; rheumatoid arthritis; inability to comply with the protocol; inadequacy in written and spoken Danish.
Recruitment/selection of patients	From July 18th 2017 to October 3rd 2018. People recruited from an orthopedic outpatient clinic at Naestved Hospital.
Age, gender and ethnicity	Age - Mean (SD): 64.7 (10.2). Gender (M:F): 38:52. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren and Lawrence at least 2 Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=45) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Strength training ontop of usual exercise care. This included leg press as the primary strength training exercise. The people in this group performed one set of low-intensity, high-repetition (30-60RM) knee extensions followed by 4 sets of high-intensity (8-12RM) leg-press in gym machines. This was done approximately 10 minutes after cessation of the neuromuscular exercise session. Performing a high-repetition set prior to high-intensity strength training is aimed at causing muscular fatigue principally in lower threshold motor units (consisting of type I muscle fibers) in order to facilitate increased recruitment of higher threshold motor units (with type II fibers) in the high-intensity training sets. The combination of a single set of low-intensity, fatiguing strength training prior to traditional high-intensity strength training has previously proven to be a potent method to enhance gains in muscle mass and strength compared with high-intensity training alone in young men. This group received this training in addition to neuromuscular exercise and education.. Duration 12 weeks. Concurrent medication/care: Education was provided in the first week (where the first two exercise sessions was completed in groups after the educational sessions). These sessions focused on disease management and self-help strategies. Both groups received neuromuscular exercises twice weekly (60

	minute sessions) for 12 weeks. The exercises consisted of warm up (for 10 minutes), circuit exercises (for 40 minutes) and cool down/stretching (for 10 minutes). The circuit exercises consisted of a total of 10 exercises, two for each domain of core stability, postural orientation and functional exercises and four for leg muscle strength. All exercises were performed in 2-3 sets of 10-15 repetitions with three levels of difficulty.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session (Initially group, but for the majority individual). 3. Type of exercise: Other (Strength and neuromodulatory). (n=45) Intervention 2: Exercise - Other supervised exercise (including flexibility, proprioception). No additional therapy. Duration 12 weeks. Concurrent medication/care: Education was provided in the first week (where the first two exercise sessions was completed in groups after the educational sessions). These sessions focused on disease management and self-help strategies. Both groups received neuromuscular exercises twice weekly (60 minute sessions) for 12 weeks. The exercises consisted of warm up (for 10 minutes), circuit exercises (for 40 minutes) and cool down/stretching (for 10 minutes). The circuit exercises consisted of a total of 10 exercises, two for each domain of core stability, postural orientation and functional exercises and four for leg muscle strength. All exercises were performed in 2-3 sets of 10-15 repetitions with three levels of difficulty.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session (Initially group, but for the majority individual). 3. Type of exercise: Neuromodulatory
Funding	Academic or government funding (Financial support provided by The Danish Rheumatism Association, The Regional Health Research Grant of Region Zealand and Naestved-Slagelse-Ringsted Hosptials Research Grant. The lead author (Holm PM) is funded by a postdoc grant from Clinical Academic Group (CAG) - Research Osteoarthritis Denmark (ROAD). Dr. Skou is currently funded by a grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 801790).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION)

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: EQ-5D 5L index at 12 weeks; Group 1: mean 0.72 (SD 0.12); n=45, Group 2: mean 0.75 (SD 0.1); n=45; EQ-5D -0.11-1 Top=High is good outcome; Comments: Reported as means and 95% confidence intervals. Converted to SDs. Reported supervised mixed modality exercise: 0.72 (0.69-0.76). Reported other supervised exercise: 0.75 (0.72-0.78). Baseline supervised mixed modality exercise: 0.6 (0.2). Baseline other supervised exercise: 0.7 (0.1). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index and baseline values of outcomes; Group 1 Number missing: 10, Reason: 3 logistical reasons, 2 underwent knee replacement surgery, 2 unable to adhere to intervention procedures, 1 unrelated health reasons, 1 family reasons, 1 unknown reasons; Group 2 Number missing: 3, Reason: 1 logistical reasons, 1 exacerbation of knee pain, 1 unrelated hospitalisation

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS activities of daily living at 12 weeks; Group 1: mean 67 (SD 13); n=45, Group 2: mean 68.1 (SD 14); n=45; KOOS activities of daily

living 0-100 Top=High is good outcome; Comments: Reported as means and 95% confidence intervals. Converted to SDs. Reported supervised mixed modality exercise: 67 (63.2-70.8). Reported other supervised exercise: 68.1 (64-72.2). Baseline supervised mixed modality exercise: 48.2 (14.5). Baseline other supervised exercise: 54.3 (11.8). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index and baseline values of outcomes; Group 1 Number missing: 10, Reason: 3 logistical reasons, 2 underwent knee replacement surgery, 2 unable to adhere to intervention procedures, 1 unrelated health reasons, 1 family reasons, 1 unknown reasons; Group 2 Number missing: 3, Reason: 1 logistical reasons, 1 exacerbation of knee pain, 1 unrelated hospitalisation Protocol outcome 3: Pain at </=3 months - Actual outcome: KOOS pain at 12 weeks; Group 1: mean 58.5 (SD 14.7); n=45, Group 2: mean 61.2 (SD 13.7); n=45; KOOS pain 0-100 Top=High is good outcome; Comments: Reported as means and 95% confidence intervals. Converted to SDs. Reported supervised mixed modality exercise: 58.5 (54.2-62.8). Reported other supervised exercise: 61.2 (57.2-65.2). Baseline supervised mixed modality exercise: 43.4 (16.3). Baseline other supervised exercise: 49.1 (12.8). Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index and baseline values of outcomes; Group 1 Number missing: 10, Reason: 3 logistical reasons, 2 underwent knee replacement surgery, 2 unable to adhere to intervention procedures, 1 unrelated health reasons, 1 family reasons, 1 unknown reasons; Group 2 Number missing: 3, Reason: 1 logistical reasons, 1 exacerbation of knee pain, 1 unrelated hospitalisation Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Serious adverse events at 12 weeks; Group 1: 3/45, Group 2: 5/45; Comments: Supervised mixed modality exercise: 2 renal system, 1 other. Other supervised exercise: 1 renal system, 1 deep vein thrombosis, 3 other. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, body mass index and baseline values of outcomes; Group 1 Number missing: 10, Reason: 3 logistical reasons, 2 underwent knee replacement surgery, 2 unable to adhere to intervention procedures, 1 unrelated health reasons, 1 family reasons, 1 unknown reasons; Group 2 Number missing: 3, Reason: 1 logistical reasons, 1 exacerbation of knee pain, 1 unrelated hospitalisation
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Holsgaard-larsen 2017[185] (Clausen 2014[90], Holsgaard-larsen 2018[184])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=93)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 year
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical diagnosis of knee osteoarthritis in accordance with the American College of Rheumatology criteria, with or without radiographic changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women with a clinical diagnosis of knee osteoarthritis aged 40-70 years. People were accepted with or without radiographic changes, had no contraindication for exercise, non-steroidal anti-inflammatory drugs or X-ray and had not had any leg surgery/trauma within the last 6 months
Exclusion criteria	People demonstrating radiographic signs of lateral compartment osteoarthritis (greater joint space narrowing in the lateral compared to medial compartment assessed qualitatively) and/or at clinical examination (area of pain and bony tenderness) were excluded
Recruitment/selection of patients	People were recruited via general practitioners in the communities of Odense and Middelfat, Denmark, and from advertisements in local clubs, libraries, print media, and Facebook
Age, gender and ethnicity	Age - Mean (SD): 58.1 (8.0). Gender (M:F): 39:54. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear (Mixed). 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 0-3, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=47) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Exercise in five parts: warming up (10 minutes of aerobic activity at 'rather strenuous levels'), functional, proprioceptive, endurance strengthening, and cooling down. The functional part comprised five exercises with

	focus on: core stability/postural function, postural orientation, and lower-extremity muscle strength. The proprioceptive party comprised three exercises, with focus on balance and functional stability. The endurance strengthening part comprised three exercise circuits, with focus on postural and functional stability of the trunk and knee. No restrictions on home exercises or participation in additional exercise programs besides the supervised NEMEX were provided. Duration 8 weeks (follow up for up to 1 year). Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Proprioceptive, functional, strengthening). (n=46) Intervention 2: Pharmacological treatment - NSAIDs. People received best information on how to use paracetamol and oral NSAIDs, in doses consistent with the Danish guidelines. If pain relief from over-the-counter paracetamol was not sufficient the pamphlet informed participants to contact their GPs to prescribe additional NSAIDs. Good compliance was defined as taking at least 2000mg/daily of paracetamol or equivalent dose of NSAID for at least 28 days. Duration 8 weeks (follow up for 1 year). Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Oral treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This project was funded by: the region of Southern Demark PhD Fund; The Danish Rheumatism Association; The Danish Rheumatism Association Ryholts grant; the University of Southern Denmark Scholarship; The Association of Danish Physiotherapists; Odense University Hospital free research funds; and Family Hede Nielsens Fund. The Parker Institute is supported by unrestricted grants from the Oak Foundation.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NSAIDS Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: KOOS QOL at 8 weeks; Group 1: mean 3.14 (SD 12.63); n=47, Group 2: mean 4.5 (SD 13.05); n=46; KOOS Quality of Life 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported exercise: 3.14 (-0.51, 6.79). Reported pharma: 4.50 (0.77, 8.23). Baseline exercise: 45.2 (16.6). Baseline pharma: 45.6 (16.6).	

Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 7, Reason: 7 did not respond to the patient-reported outcomes (reasons not given)

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: KOOS QOL at 52 weeks; Group 1: mean 10 (SD 15.1); n=47, Group 2: mean 8.7 (SD 15.4); n=46; KOOS Quality of Life 0-100 Top=High is good outcome; Comments: Reports change scores and standard error. Reported exercise: 10.0 (2.2). Reported pharma: 8.7 (2.3). Baseline exercise: 45.2 (16.6). Baseline pharma: 45.6 (16.6).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 6, Reason: 6 did not respond to the patient-reported outcomes (reasons not given)

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: KOOS ADL at 8 weeks; Group 1: mean 6.96 (SD 11.19); n=47, Group 2: mean 7.46 (SD 11.06); n=46; KOOS activities of daily living 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported exercise: 6.96 (3.76, 10.16). Reported pharma: 7.46 (4.26, 10.65). Baseline exercise: 68.2 (15.5). Baseline pharma: 68.4 (17.1).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 7, Reason: 7 did not respond to the patient-reported outcomes (reasons not given)

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: KOOS ADL at 52 weeks; Group 1: mean 11.4 (SD 13.7); n=47, Group 2: mean 7.9 (SD 13.4); n=46; KOOS activities of daily living 0-100 Top=High is good outcome; Comments: Reports change scores and standard error. Reported exercise: 11.4 (2.0). Reported pharma: 7.9 (2.0). Baseline exercise: 68.2 (15.5). Baseline pharma: 68.4 (17.1).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 6, Reason: 6 did not respond to the patient-reported outcomes (reasons not given)

Protocol outcome 5: Pain at </=3 months
- Actual outcome: KOOS pain at 8 weeks; Group 1: mean 7.23 (SD 10.77); n=47, Group 2: mean 5.15 (SD 10.68); n=46; KOOS pain 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported exercise: 7.23 (4.14, 10.3). Reported pharma: 5.15 (2.06, 8.23). Baseline exercise: 61.6 (13.7). Baseline pharma: 60.1 (15.0).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic

severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 7, Reason: 7 did not respond to the patient-reported outcomes (reasons not given) Protocol outcome 6: Pain at > 3 months - Actual outcome: KOOS pain at 52 weeks; Group 1: mean 13.6 (SD 13.7); n=47, Group 2: mean 9.4 (SD 14.1); n=46; KOOS pain 0-100 Top=High is good outcome; Comments: Reports change scores and standard error. Reported exercise: 13.6 (2.0). Reported pharma: 9.4 (2.1). Baseline exercise: 61.6 (13.7). Baseline pharma: 60.1 (15.0). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, weight, BMI, study knee, radiographic severity, social economic status, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 did not respond to the patient-reported outcomes (reasons not given); Group 2 Number missing: 6, Reason: 6 did not respond to the patient-reported outcomes (reasons not given)	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Huang 2003[192]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=132)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 week with additional follow up 52 weeks afterwards
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Moderate bilateral knee osteoarthritis (Altman grade II)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with moderate bilateral knee osteoarthritis (Altman grade 2)
Exclusion criteria	People with respiratory or cardiac dysfunction, or combined ankle or hip pain
Recruitment/selection of patients	People were selected by clinicians from outpatients attending the department of rehabilitation
Age, gender and ethnicity	Age - Mean (SD): 62 (4.5). Gender (M:F): 39:93. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Altman grade II Duration of symptoms (range): 4 months - 9 years
Indirectness of population	No indirectness
Interventions	(n=99) Intervention 1: Exercise - Supervised strength exercise. Three groups: One had isokinetic exercise (speed constant), one had isotonic exercise (speed variable) and one had isometric (speed constant but isometric hold angles were used in the range of motion, the speed of passive forward or backward motion was set at 30 degrees/second). Exercise 3 times weekly for 8 weeks.. Duration 8 weeks. Concurrent medication/care: The people in all groups also received 20 minutes of hot packs and passive range motion exercise by an electric stationary bike (20 cycles per minute) for 5 minutes to both knees before exercise (unclear as to whether this applied to the control group).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable Comments: The three types of exercise were pooled together for class effect as

	agreed in the protocol (n=33) Intervention 2: No treatment. Controls (no treatment). Duration 8 weeks. Concurrent medication/care: The people in all groups also received 20 minutes of hot packs and passive range motion exercise by an electric stationary bike (20 cycles per minute) for 5 minutes to both knees before exercise (unclear as to whether this applied to the control group).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by a project grant from the National Science Council of Taiwan)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analogue scale score of knee pain at 8 weeks; Group 1: mean 3.1 (SD 1); n=99, Group 2: mean 4.4 (SD 0.4); n=33; VAS 0-100 Top=High is poor outcome; Comments: Reports isokinetic: 3.1 (1.2). Reported isotonic: 2.6 (0.7). Reported isometric: 3.6 (0.6). Reported control: 4.4 (0.4). Baseline isokinetic: 4.8 (1.4). Baseline isotonic: 4.6 (.7). Baseline isometric: 4.7 (1.4). Baseline control: 4.6 (1.3).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at > 3 months
- Actual outcome: Visual analogue scale score of knee pain at 52 weeks; Group 1: mean 2.6 (SD 1.7); n=99, Group 2: mean 6.1 (SD 1.3); n=33; VAS 0-10 Top=High is poor outcome; Comments: Reports isokinetic: 2.5 (1.8). Reported isotonic: 2.0 (1.4). Reported isometric: 3.2 (1.6). Reported control: 6.1 (1.3). Baseline isokinetic: 4.8 (1.4). Baseline isotonic: 4.6 (.7). Baseline isometric: 4.7 (1.4). Baseline control: 4.6 (1.3).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Huang 2005[191]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (with additional follow up 1 year later)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Bilateral moderate knee osteoarthritis (Altman grade 2) with periarticular soft tissue pain, as identified by painful sensations during palpation or passive stretching of the arthritis knee under orthopedic examination. Confirmed by radiography.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with bilateral moderate knee osteoarthritis with periarticular soft tissue pain, as identified by painful sensations during palpation or passive stretching of the arthritic knee under orthopedic examination.
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.0 (8.4). Gender (M:F): 1:4.2 (as reported by the study). Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Altman grade 2 Duration of symptoms (range): 6 months - 11 years
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise - Supervised strength exercise. Isokinetic exercise - 5 minute warm up exercise on a stational bicycle set without resistance. Then exerise for both knees with 60% of the average peak torque, with increase being increased from 1 set to 5 sets during the first through fifth sessions, and remained at 6 sets for the remaining 6th through 24th sessions. Each set consisted of 5 repetitions of concentric contraction in angular velocities of 30 degrees/s and 120 degrees/s for extensors, and 5 repetitions of eccentric and concentric contractions in angular velocities of 30 degrees/s and 120 degrees/s for flexors. The start and stop angles for

	extension exercises were 40 degrees and 70 degrees, and the start and stop angles for flexion exercises were 70 degrees and 40 degrees. People were allowed 5 seconds of rest between sets, 10 seconds of rest between different modes of training, and 10 minutes of rest between right and left knee training. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=30) Intervention 2: No treatment. Control group (no additional information). Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=60) Intervention 3: Other. Two additional groups: Exercise and ultrasound therapy, and ultrasound therapy alone. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These groups were not included in the analysis as they did not fulfill the inclusion criteria in the protocol
Funding	Academic or government funding (Supported by National Science Council of Taiwan (grant no. NSC-92-2314-B-037-067))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Pain at </=3 months - Actual outcome: Visual analogue scale score for knee pain at 8 weeks; Group 1: mean 1.2 (SD 1.4); n=30, Group 2: mean 0.4 (SD 1.6); n=30; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.9 (1.5). Baseline control: 4.8 (1.8). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Pain at > 3 months - Actual outcome: Visual analogue scale score for knee pain at 52 weeks; Group 1: mean 3.5 (SD 1.7); n=30, Group 2: mean 6 (SD 1.3); n=30; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.9 (1.5). Baseline control: 4.8 (1.8). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,	

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Huang 2005[193]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=140)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 week (additional follow up 1 year later)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Bilateral moderate knee osteoarthritis (Altman grade 2)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with bilateral moderate knee osteoarthritis
Exclusion criteria	No additional information
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 65.0 (6.4). Gender (M:F): 27:113. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Altman grade 2 Duration of symptoms (range): 5 months - 12 years
Indirectness of population	No indirectness
Interventions	(n=35) Intervention 1: Exercise - Supervised strength exercise. Isokinetic muscular strengthening exercises. People received a warmup exercise with 20 minutes of hot packs and underwent passive range or motion exercises on an electric stationary bike (20 cycles per minute) for 5 minutes to both knees before undergoing the exercise. The exercise is a mode of speed-constant exercise. The isokinetic exercise program began with 60% of the mean peak torque preset in the Kin-Com, and the person reached the present intensity by visual biofeedback. An increasing dose program was used in the first 5 sessions (1 set to 5 sets), and a dose of 6 sets was applied from the sixth to twennty-fourth sessions, with the density rising from 60% to 80% of the mean peak torque as the person was able. Each set consisted of 5 repetitions of concentric contraction in angular velocities of 30 degrees/second and 120 degrees/second for extensors and 5 repetitions of eccentric and concentric contractions in angular velocities of 30 degrees/second and 120 degrees/second for flexors. The start and

	stop angles for extension exercise were 40 degrees and 70 degrees, and the start and stop angles for flexion exercise were 70 degrees and 40 degrees. People were allowed 5 seconds of rest between sets, 10 seconds of rest between extensors and flexors strengthening modes, and 10 minutes of rest between right and left knee training. After completing treatment, people in the treated groups received a home exercise program with 15 minutes of stationary bicycling exercise, using an exercise bike or a common bicycle with a device attached to elevate the posterior wheel to executre/perform the bicycling exercise for people who did not have an exercise bike at home.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=35) Intervention 2: No treatment. Control group (no treatment). Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=70) Intervention 3: Other. Two additional groups, one receiving exercise and ultrasound treatment, one receiving exercise, ultrasound and hyaluronic acid injection. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These two groups were not included in the analysis as they did not fulfill the inclusion criteria in the protocol
Funding	Academic or government funding (Supported by a project grant from the National Science Council of Taiwan)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Pain at </=3 months - Actual outcome: Visual analogue scale score for knee pain at 8 weeks; Group 1: mean 1.2 (SD 1.6); n=35, Group 2: mean 0.5 (SD 1.7); n=35; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.3 (1.5). Baseline control: 5.4 (1.7). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	

Protocol outcome 2: Pain at > 3 months - Actual outcome: Visual analogue scale score for knee pain at 52 weeks; Group 1: mean 3.9 (SD 1.4); n=35, Group 2: mean 6.6 (SD 1.5); n=35; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.3 (1.5). Baseline control: 5.4 (1.7). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Hunt 2018[196]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=79)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 5 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Definitive medial tibiofemoral osteophytes on x-ray with joint space narrowing greater in the medial tibiofemoral compartment compared to the lateral compartment
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Definitive medial tibiofemoral osteophytes on X-ray; joint space narrowing greater in the medial tibiofemoral compartment compared to the lateral compartment; history of knee pain longer than 6 months; average knee pain of at least 3 out of 10 over the 1 month period prior to initial screening
Exclusion criteria	Knee surgery or intra-articular pain relief injection within 6 months; current or past (within 6 months) oral corticosteroid use; history of knee joint replacement or tibial osteotomy; any other condition affecting lower limb function; participation in a new structured exercise program within the past 3 months, or planning to commence exercise or other treatment for knee osteoarthritis in the next 4 months; an inability to travel to the university to attend testing and training sessions
Recruitment/selection of patients	Community dwelling individuals recruited via an existing laboratory database as well as advertisements in print media
Age, gender and ethnicity	Age - Mean (SD): 65.0 (8.7). Gender (M:F): 24:55. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Median radiographic severity - Moderate Duration of symptoms: at least 6 months
Indirectness of population	No indirectness
Interventions	(n=39) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Toe-out gait modification program. Trained to perform walking with 15 degrees more toe-out than the self-selected amount

	measured at the baseline testing session. Toe-out modification during the training sessions was facilitated with mirror-guided biofeedback or performance. People placed their study foot on a protractor device at the target toe out angle for that session, and verbally instructed the therapist int he placement of a piece of green tape on the mirror to best cover the reflection of the foot in this target position. The tape remained on the mirror during the training session to guide foot placement during treadmill walking. To promote motor learning, a faded feedback paradigm was used with removal of real-time biofeedback commencing at session 4.. Duration 4 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Proprioceptive(? - gait adjustment) and aerobic). (n=40) Intervention 2: Exercise - Supervised aerobic exercise . People underwent all training procedure as those in the other group, with the exception of receiving no training or instruction related to toe-out walking. This included walking on the treadmill in front of a mirror during training sessions, but without foot placement guide tape as per the toe-out gait modification training protocol. Duration 4 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (Funding for this study was received from The Arthritis Society (SOG-13-024) (Canada). Salary support from provided by the Michael Smith Foundation for Health Research, the Canadian Institutes of Health Research,and the Natural Sciences and Engineering Research Council of Canada)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED AEROBIC EXERCISE

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: WOMAC physical function at 5 months; Group 1: mean -9.4 (SD 9.7); n=39, Group 2: mean -6.6 (SD 10.3); n=40; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported mixed: -9.4 (-12.4, -6.3). Reported aerobic: -6.6 (-9.8, -3.4). Baseline mixed: 28.1 (1.9). Baseline aerobic: 21.4 (1.5).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, heigh, body mass, BMI, radiographic severity, and baseline values for outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up - 2 unable to commit further, 1 unable to contact; Group 2 Number missing: 4, Reason: 4 lost to follow up - 2 unable to commit further, 1 unrelated health issues, 1 family emergency

Protocol outcome 2: Pain at > 3 months - Actual outcome: WOMAC pain at 5 months; Group 1: mean -2.5 (SD 3); n=39, Group 2: mean -1.5 (SD 3.2); n=40; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported mixed: -2.5 (-3.5, -1.6). Reported aerobic: -1.5 (-2.5, -0.5). Baseline mixed: 7.6 (0.5). Baseline aerobic: 6.4 (0.4). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, heigh, body mass, BMI, radiographic severity, and baseline values for outcomes; Group 1 Number missing: 3, Reason: 3 lost to follow up - 2 unable to commit further, 1 unable to contact; Group 2 Number missing: 4, Reason: 4 lost to follow up - 2 unable to commit further, 1 unrelated health issues, 1 family emergency	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Imoto 2012[203]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age between 50 and 75 years, diagnosis of knee osteoarthritis according to the criteria of the American College of Rheumatology based on history, physical examination and radiographic findings (pain in the knee and one of the following items - over 50 years of age, less than 30 minutes of morning stiffness and crepitation in active movement and osteophytes), knee x-ray in the last 12 months and grade 2 or above in the Kellgren and Lawrence radiographic classification.
Exclusion criteria	People with a diagnosis of fibromyalgia, unstable heart condition, physical activity more often than twice a week, inability to pedal a stationary bicycle and previous knee arthroplasty, the occurrence of adverse events
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.1 (8.5). Gender (M:F): 8:92. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=50) Intervention 1: Exercise - Supervised strength exercise. Muscle strengthening group activities based on the 10 maximum repetitions test. After estimating 100% of the load, 50-60% of this load was established for use in the strengthening of people from the study. This was completed through group sessions lasting from 30 to 40 minutes, with a weekly frequency of twice a week. The exercise protocol used by us consisted of 10 minutes of warm-up on a stationary bicycle, ischiotibial stretching

	exercises and three series of 15 repetitions of knee extension exercises, aiming to strengthen the quadriceps muscle. The interval between series was from 30-45 seconds. the load used in the exercise was increased according to tolerance. The person's positioning for the exercise was: seated in a chair, with 90 degrees of knee and hip flexion.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=50) Intervention 2: No treatment. An explanation about a manual after initial evaluation. The orientation manual consisted of a description of knee osteoarthritis, as well as the possible signs and symptoms presented by the patients, and pointed them in the direction of a better way of dealing with the functional difficulties.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 functional capacity at 8 weeks; Group 1: mean 20.28 (SD 27.42); n=50, Group 2: mean 6.96 (SD 26.79); n=50; SF-36 functional capacity 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 20.28 (12.68, 27.88). Reported control: 6.96 (-0.46, 14.39). Baseline exercise: 31 (19.59). Baseline control: 34.53 (24.76).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 physical role at 8 weeks; Group 1: mean 27.85 (SD 67.5); n=50, Group 2: mean 13.39 (SD 64.9); n=50; SF-36 physical role 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 27.85 (9.13, 46.57). Reported control: 13.39 (-4.59, 31.38). Baseline exercise: 27.16 (38.74). Baseline control: 25.6 (38.39).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 pain at 8 weeks; Group 1: mean 16.4 (SD 36.85); n=50, Group 2: mean 6.14 (SD 41.16); n=50; SF-36 pain 0-100 Top=High is good

outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 16.40 (6.18, 26.61). Reported control: 6.14 (-5.27, 17.55). Baseline exercise: 34.47 (18.27). Baseline control: 34.51 (24.3).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 general health at 8 weeks; Group 1: mean 8.05 (SD 24.68); n=50, Group 2: mean 5.89 (SD 26.1); n=50; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 8.05 (1.21, 14.89). Reported control: 5.89 (-1.34, 13.13). Baseline exercise: 52.24 (27.72). Baseline control: 50.77 (21.43).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 vitality at 8 weeks; Group 1: mean 10 (SD 37.9); n=50, Group 2: mean 3.17 (SD 40.62); n=50; SF-36 vitality 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 10 (2.26, 17.73). Reported control: 3.17 (-8.08, 14.44). Baseline exercise: 53.11 (23.04). Baseline control: 52.53 (22.08).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 social aspect at 8 weeks; Group 1: mean 9.57 (SD 37.27); n=50, Group 2: mean 0.35 (SD 51.7); n=50; SF-36 social aspect 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 9.57 (-0.76, 19.9). Reported control: 0.35 (-13.97, 14.69). Baseline exercise: 71.24 (26.01). Baseline control: 63.56 (29.27).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 mental health at 8 weeks; Group 1: mean 3.77 (SD 29.15); n=50, Group 2: mean 1.28 (SD 32.99); n=50; SF-36 mental health 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 3.77 (-4.31, 11.85). Reported control: 1.28 (-7.86, 10.43). Baseline exercise: 59.27 (24.86). Baseline control: 55.88 (24.15).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return
- Actual outcome: SF-36 emotional role at 8 weeks; Group 1: mean 16.22 (SD 68.22); n=50, Group 2: mean 13.21 (SD 86.46); n=50; SF-36 emotional role 0-100 Top=High is good outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: 16.22 (-2.68, 35.14). Reported control: 13.21 (-10.75, 37.18). Baseline exercise: 47.22 (46.7). Baseline control: 35.49 (42.07).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return Protocol outcome 2: Pain at </=3 months - Actual outcome: Pain (NRS) at 8 weeks; Group 1: mean -3.17 (SD 3.84); n=50, Group 2: mean -0.88 (SD 3.73); n=50; NRS 0-10 Top=High is poor outcome; Comments: Reported mean difference and 95% confidence intervals. Reported exercise: -3.17 (-4.23, -2.10). Reported control: -0.88 (-1.92, 0.15). Baseline exercise: 7.43 (2.01). Baseline control: 6.92 (2.60). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, side treated, age, BMI, Kellgren Lawrence grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 non-adherence, 2 significant knee inflammation, 1 death in the family, 2 found new job, 1 treatment closer to home; Group 2 Number missing: 12, Reason: 12 losses, 1 ankle fracture, 11 did not return	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Jan 2008[206]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=98)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Bilateral knee pain that fulfilled the American College of Rheumatology criteria for knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with bilateral knee pain that fulfilled the American College of Rheumatology criteria for knee osteoarthritis (knee pain with osteophytes confirmed by radiography and the following 3: experiencing stiffness for less than 30 minutes in the morning, having crepitus and being older than 50 years of age). Additionally: an osteoarthritis grade of 3 or lower on the Kellgren/Lawrence classification based on plain radiographs, as assessed by the same orthopedic surgeon, who had more than 30 years of clinical experience; a history of knee pain longer than 6 months (chronic knee osteoarthritis)
Exclusion criteria	If they had received knee physical therapy during the preceding 3 months or had other musculoskeletal problems associated with the knee joint (such as tendon or ligament tears), central or peripheral neuropathy, or other unstable medical conditions
Recruitment/selection of patients	Recruited from the Department of Orthopedics, National Taiwan University Hospital
Age, gender and ethnicity	Age - Mean (SD): 62.6 (6.7). Gender (M:F): 19:79. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 3.2 (2.7) years
Indirectness of population	No indirectness
Interventions	(n=68) Intervention 1: Exercise - Supervised strength exercise. High resistance or low resistance exercise using the EN-Dynamic Track leg press machine. People performed knee resistance training in a sitting position, with one foot placed on the center of the pedal of the EN-Dynamic Track machine. Subjects were asked to fully

	extend and flex their knee joint from 90 degrees or knee flexion. Each action was completed rhythmically, with the first second spent extending the knee and the following second spent flexing the knee. People in both groups underwent 3 training sessions per week for 8 weeks. The program was delivered individually. The high intensity exercise was performed at 60% of 1 RM, while the low intensity exercise was performed at 10% of 1 RM. Cool packs were applied to the subjects' knees for 10 minutes after exercise completion.. Duration 8 weeks. Concurrent medication/care: People were not allowed to take non-steroidal anti-inflammatory medication during the study. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable Comments: The two groups were combined for analysis due to class effect (n=30) Intervention 2: No treatment. No treatment control. Duration 8 weeks. Concurrent medication/care: People were not allowed to take non-steroidal anti-inflammatory medication during the study. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 8 weeks; Group 1: mean 14.8 (SD 8.9); n=68, Group 2: mean 22.5 (SD 10.9); n=30; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reported high resistance: 14.7 (8.5), reported low resistance: 14.8 (9.2). Baseline high resistance: 26.4 (9.0). Baseline low resistance: 26.1 (8.1). Baseline control: 25.4 (11.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, heigh, weight, osteoarthritis duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: High resistance: 3 discontinued due to knee pain. Low resistance: No one discontinued; Group 2 Number missing: 4, Reason: 4 lost to follow up due to personal reasons other than knee pain

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain subscale at 8 weeks; Group 1: mean 4.8 (SD 3.1); n=68, Group 2: mean 7.1 (SD 3.4); n=30; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reported high resistance: 4.8 (3.5), reported low resistance: 4.8 (2.7). Baseline high resistance: 8.5 (3.8). Baseline low resistance: 7.8 (3.3). Baseline control: 8.3 (4.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, heigh, weight, osteoarthritis

duration, radiographic grade and baseline values of outcomes; Group 1 Number missing: 3, Reason: High resistance: 3 discontinued due to knee pain. Low resistance: No one discontinued; Group 2 Number missing: 4, Reason: 4 lost to follow up due to personal reasons other than knee pain	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Jorge 2015[214]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Unilateral or bilateral osteoarthritis of the knee, based on the classification criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee based on the criteria of the American College of Rheumatology; age between 40 and 70 years; and pain at rest between 3 and 8 out of 10 on the visual analogue scale for one or both knees.
Exclusion criteria	Inflammatory conditions or any medical condition that prevented physical activity; joint injection in the previous 3 months; regular physical activity at the time; or travel plans for the subsequent 12 weeks
Recruitment/selection of patients	People were selected by telephone using a database of people with osteoarthritis from the Universidade Federal de Sao
Age, gender and ethnicity	Age - Mean (SD): 60.8 (7.0). Gender (M:F): 0:60. Ethnicity: 69-71% were Caucasian, no other information given
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Radiographic grade 1-2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=29) Intervention 1: Exercise - Supervised strength exercise. Progressive resistance exercise program that included four different exercises: knee extension/flexion and hip abduction/adduction using two gym machines (knee flexion -extension and abduction-adduction) with free weights. The exercises were preceded by five-minute warm-up on an exercise bicycle. The initial load was based on the 1RM. The porgram was structured as follows: 2 sets of 8 repetitions, the first set employing 50% of 1RM and the second set employing 75% of 1RM. A 1 minute rest interval was given between

	sets. The exercise program was performed twice a week over a 12 week period.. Duration 12 weeks. Concurrent medication/care: When pain exceeded a 7 on the visual analog scale, the subject could take 50mg of diclofenac every 8 hours.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=31) Intervention 2: No treatment. No exercise control. Duration 12 weeks. Concurrent medication/care: When pain exceeded a 7 on the visual analog scale, the subject could take 50mg of diclofenac every 8 hours.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The present study was supported by grants from Brazilian fostering agencies - Fundacao de Amparo a Pesquisa do Estado de Sao Paulo (FAPESP) and Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical function at 12 weeks; Group 1: mean 49.8 (SD 21.9); n=29, Group 2: mean 30.8 (SD 16.8); n=31; SF-36 physical function 0-100 Top=High is good outcome; Comments: Baseline exercise: 39.3 (16.3). Baseline control: 32.4 (16.0).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 physical role limitation at 12 weeks; Group 1: mean 48.3 (SD 41.7); n=29, Group 2: mean 16.9 (SD 23.6); n=31; SF-36 physical role 0-100 Top=High is good outcome; Comments: Baseline exercise: 25.9 (36.3). Baseline control: 22.6 (26.1).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 pain at 12 weeks; Group 1: mean 58.6 (SD 25); n=29, Group 2: mean 41.7 (SD 20.6); n=31; SF-36 pain 0-100 Top=High is good outcome; Comments: Baseline exercise: 44.9 (21.9). Baseline control: 39.0 (15.7).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or

health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 general health at 12 weeks; Group 1: mean 66.1 (SD 21.8); n=29, Group 2: mean 52.6 (SD 21.8); n=31; SF-36 general health 0-100 Top=High is good outcome; Comments: Baseline exercise: 65.4 (22.3). Baseline control: 53.1 (23.1).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 vitality at 12 weeks; Group 1: mean 64 (SD 25.2); n=29, Group 2: mean 52.4 (SD 21.3); n=31; SF-36 vitality 0-100 Top=High is good outcome; Comments: Baseline exercise: 55.9 (23.1). Baseline control: 50.0 (24.2).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 social aspects at 12 weeks; Group 1: mean 77.2 (SD 28.9); n=29, Group 2: mean 57.7 (SD 27.9); n=31; SF-36 social aspects 0-100 Top=High is good outcome; Comments: Baseline exercise: 62.1 (28.8). Baseline control: 55.2 (30.8).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 emotional role at 12 weeks; Group 1: mean 72.4 (SD 39.9); n=29, Group 2: mean 49.5 (SD 39.3); n=31; SF-36 emotional role 0-100 Top=High is good outcome; Comments: Baseline exercise: 56.3 (44.6). Baseline control: 44.1 (38.9).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days
- Actual outcome: SF-36 mental health at 12 weeks; Group 1: mean 76.4 (SD 18.7); n=29, Group 2: mean 59.5 (SD 21.2); n=31; SF-36 mental health 0-100 Top=High is good outcome; Comments: Baseline exercise: 67.9 (19.9). Baseline control: 61.5 (21.3).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC function subscale at 12 weeks; Group 1: mean 17.3 (SD 12.4); n=29, Group 2: mean 26.7 (SD 10.2); n=31; WOMAC function subscale 0-68 Top=High is poor outcome; Comments: Baseline exercise: 27.7 (9.3). Baseline control: 28.4 (10.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or

health problems, 2 more abandoned experiment at 90 days Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean 4.9 (SD 4.2); n=29, Group 2: mean 9.5 (SD 3.2); n=31; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 9.0 (2.9). Baseline control: 9.3 (3.3). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Increased pain at 12 weeks; Group 1: 3/29, Group 2: 0/31 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different for all of the subscales of SF-36; Group 1 Number missing: 2, Reason: 2 abandoned due to personal or health problems; Group 2 Number missing: 3, Reason: 2 abandoned due to personal or health problems, 2 more abandoned experiment at 90 days	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	**Joshi 2019[215]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed by an orthopedician

Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with age more than 40 years diagnosed with knee osteoarthritis by an orthopedician who were referred to or attended the O.P.D, Department of Physiotherapy, Guru Jambheshwar University of Science and Technology, Hisar.
Exclusion criteria	People having a history of any inflammatory, infection, traumatic condition or the knee joint; with any previous surgery or any invasive procedure of knee joint; history of cardiac disease; lower limb injury or pathology; fixed deformity of the knee; any skin problems around the knee joint; lacking independent ambulation or requiring use of any walking aid; neurological disorders; patients with severe knee osteoarthritis (grade 4 or those referred for knee replacement surgery); those unable to comply with study protocol.
Recruitment/selection of patients	People who were referred to or attended the O.P.D, Department of Physiotherapy, Guru Jambheshwar University of Science and Technology, Hisar.
Age, gender and ethnicity	Age - Mean (SD): 52.5 (9.5). Gender (M:F): 20:22. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=21) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Retrowalking group. Retrowalking protocol on a treadmill in addition to the conventional exercise program. The intervention began with a session of 3-5 minute forward walking on the treadmill for warm up followed by 10 minutes of retrowalking on the treadmill at a comfortable speed of 0 degrees inclination under the supervision of the therapist. The treadmill was placed within a specially designed metal framework with handrails support for safety purposes. Both interventions were given for three session in a week for a total duration of six weeks.. Duration 6 weeks. Concurrent medication/care: Both groups received conventional exercise program which consisted of hot packs for 10 minutes followed by exercises. These consisted of range of motion exercises, muscle strengthening exercise in the form of isometric and isotonic exercises, muscle stretching exercises and flexibility exercises.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Aerobic and strengthening). (n=21) Intervention 2: Exercise - Supervised strength exercise. Conventional exercise only. Three sessions per week for six weeks.. Duration 6 weeks. Concurrent medication/care: Both groups received conventional exercise program which consisted of hot packs for 10 minutes followed by exercises. These consisted of range of motion exercises, muscle strengthening exercise in the form of isometric and isotonic exercises, muscle stretching exercises and flexibility exercises.. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: Physical function (WOMAC) at 6 weeks; Group 1: mean 20.79 (SD 4.2); n=21, Group 2: mean 31.79 (SD 6.1); n=21; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed modality: 46.84 (6.44). Baseline strength: 51.69 (7.32).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, sex and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: Pain (VAS) at 6 weeks; Group 1: mean 2.85 (SD 0.88); n=21, Group 2: mean 5.2 (SD 1.17); n=21; VAS 0-10 Top=High is poor outcome; Comments: Baseline mixed modality: 7.82 (1.08). Baseline strength: 7.92 (0.98).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, sex and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Juhakoski 2011[216]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Finland; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 24 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with unilateral or bilateral hip osteoarthritis fulfilling the clinical and radiological criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People of age 55-80 years, willingness to take part in a study lasting for two years, and unilateral or bilateral hip osteoarthritis with Kellgren Lawrence grade at least 1 (X-ray less than 3 years old) and pain experienced in the hip region (groin and lateral hip region) within the preceding month as indicated in the clinical criteria of the American College of Rheumatology.
Exclusion criteria	Total hip replacement; rheumatoid arthritis; cognitive impairment; a major surgical operation within the preceding six months in the lower limb or lower back area; acute or subacute lower back pain; cardiovascular or pulmonary disease or some other chronic disease that would prevent full participation in the training programme.
Recruitment/selection of patients	Recruitment from newspaper advertisements and a small number being selected from specialists' clinics (2) or by general practitioners (5)
Age, gender and ethnicity	Age - Mean (SD): 66.6 (6.5). Gender (M:F): 35:83. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (No chronic disease = 49, 1 chronic disease = 53, 2 or more chronic diseases = 16). 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity: Radiological grade 1-4, median grade 2 Duration of follow up (mean [SD]): 5.5 (5.5) years
Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: Exercise - Unsupervised strength exercise. Home exercise program of hip strengthening exercises, conducted over around 30-35 minutes where exercises were made with the maximal effort in order to achieve the highest possible

	movement velocity with 1-2 minutes rest between each exercise. 2-3 repetitions with each legs on average. This was taught over 12 supervised sessions 9once per week for 45 minutes) with four additional booster sessions one year later.. Duration 24 months. Concurrent medication/care: All people received an hour long instruction session regarding the basic principles of non-operative treatment for hip osteoarthritis. All people received GP standard care.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=58) Intervention 2: No treatment. No exercise treatment. Duration 24 months. Concurrent medication/care: All people received an hour long instruction session regarding the basic principles of non-operative treatment for hip osteoarthritis. All people received GP standard care.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by an EVO-grant from Mikkeli Central Hospital)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 3 months; Group 1: mean 27.4 (SD 13.9); n=60, Group 2: mean 25.9 (SD 14.5); n=58; WOMAC function 0-100 Top=High is poor outcome; Comments: Reports mean and SE. Reported exercise: 27.4 (1.8). Reported no treatment: 25.9 (1.9). Baseline exercise: 24.7 (16.7). Baseline no treatment: 28.9 (22.4).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, duration of symptoms, working status, radiographic grades, existence of knee osteoarthritis, comorbidities, body mass index and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: WOMAC function at 24 months; Group 1: mean 24.4 (SD 20.9); n=60, Group 2: mean 30 (SD 21.3); n=58; WOMAC function 0-100 Top=High is poor outcome; Comments: Reports mean and SE. Reported exercise: 24.4 (2.7). Reported no treatment: 30.0 (2.8). Baseline exercise: 24.7 (16.7). Baseline no treatment: 28.9 (22.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, duration of symptoms, working status, radiographic grades, existence of knee osteoarthritis, comorbidities, body mass index and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 dropped out in total - Reasons unclear (given for different time periods). Overall - 1 lost due to neck pain, 1 lost due to other disease, 3

lost due to total hip replacement; Group 2 Number missing: 5, Reason: 5 dropped out in total - reasons unclear (given for different time periods). Overall - 1 deceased, 3 total hip replacements Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 3 months; Group 1: mean 27.6 (SD 16.3); n=60, Group 2: mean 24.3 (SD 16.8); n=58; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reports mean and SE. Reported exercise: 27.6 (2.1). Reported no treatment: 24.3 (2.2). Baseline exercise: 21.5 (14.8. Baseline no treatment: 29.1 (20.2). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, duration of symptoms, working status, radiographic grades, existence of knee osteoarthritis, comorbidities, body mass index and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Pain at > 3 months - Actual outcome: WOMAC pain at 24 months; Group 1: mean 24.1 (SD 22.5); n=60, Group 2: mean 27.9 (SD 22.8); n=58; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reports mean and SE. Reported exercise: 24.1 (2.9). Reported no treatment: 27.9 (3.0). Baseline exercise: 21.5 (14.8. Baseline no treatment: 29.1 (20.2). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, duration of symptoms, working status, radiographic grades, existence of knee osteoarthritis, comorbidities, body mass index and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 dropped out in total - Reasons unclear (given for different time periods). Overall - 1 lost due to neck pain, 1 lost due to other disease, 3 lost due to total hip replacement; Group 2 Number missing: 5, Reason: 5 dropped out in total - reasons unclear (given for different time periods). Overall - 1 deceased, 3 total hip replacements	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kang 2019[221]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=29)
Countries and setting	Conducted in South Korea; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Hand osteoarthritis as suggested by the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Male career workers recruited from an automobile assembly line with hand osteoarthritis according to the inclusion criteria suggested by the American College of Rheumatology: hand pain or stiffness and hard tissue enlargement in at least 2 out of 10 selected joints; hard tissue enlargement of at least 2 DIP joints; 3 or fewer swollen metacarpophalangeal joints; deformity in at least 1 out of 10 selected joints; at least 5 points in the functional index for hand osteoarthritis
Exclusion criteria	Presence of cognitive disorder; history of recent serious trauma; history of recent surgery for osteoarthritis or other major operations; having received a corticosteroid injection in a hand joint in the prior 2 months
Recruitment/selection of patients	Male career workers recruited from an automobile assembly line
Age, gender and ethnicity	Age - Mean (SD): 47.3 (4.4). Gender (M:F): 29:0. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hand osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 3.5 (1.1) years.
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Supervised strength exercise. Finger exercise program. Exercises 1 and 2 maintain or increase the flexibility of the MCP, PIP and DIP joints. Exercise 3 increases opponens pollicis strength and grip strength. Exercise 4 strengthens the extensor and abductor pollicis muscles. The purpose was to maintain the thumb web space, increase thumb stability and counteract the strong pull from the adductor pollicis muscle, combined with the increasing weakness of the

	opposing thenar instrinsic musculature, which can be seen in individuals with carpometacarpal joint osteoarthritis, thereby leading to thumb adduction deformity. Exercises 5 and 6 increase grip strength. The exercises were performed for 30 minutes per day, 5 times a week for 8 weeks. All exercises were performed with 10 repetitions for the initial 2 weeks and 15 repetitions for weeks 3 to 8.. . Duration 8 weeks. Concurrent medication/care: Both groups received dip-wrap paraffin bath therapy sessions. The temperature of the paraffin bath was 50 degrees centigrade. Subjects dipped the affect hand in, removed the hand, and waited for the paraffin to harden and become opaque. They then re-dipped the hand up to 10 times. When the last layer hardened, the hand was covered with a towel for 20 minutes.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=14) Intervention 2: No treatment. No additional treatment. Duration 8 weeks. Concurrent medication/care: Both groups received dip-wrap paraffin bath therapy sessions. The temperature of the paraffin bath was 50 degrees centigrade. Subjects dipped the affect hand in, removed the hand, and waited for the paraffin to harden and become opaque. They then re-dipped the hand up to 10 times. When the last layer hardened, the hand was covered with a towel for 20 minutes.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: AUSCAN physical function score at 8 weeks; Group 1: mean 50.93 (SD 7.01); n=15, Group 2: mean 56.64 (SD 5.26); n=14; AUSCAN physical function score 0-100 Top=High is poor outcome; Comments: Baseline exercise: 67.73 (9.42). Baseline no treatment: 68.07 (6.72). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, career length, height, mass, BMI, symptom duration, number of painful hand joints, number of stiff hand joints, number of bony knobs, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Pain at </=3 months - Actual outcome: AUSCAN pain score at 8 weeks; Group 1: mean 42.07 (SD 5.26); n=15, Group 2: mean 56.5 (SD 6.19); n=14; AUSCAN pain 0-100	

Top=High is poor outcome; Comments: Baseline exercise: 63.67 (9.42). Baseline control: 64.36 (9.36). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, career length, height, mass, BMI, symptom duration, number of painful hand joints, number of stiff hand joints, number of bony knobs, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Karadag 2019[222]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=62)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with a diagnosis of knee osteoarthritis according to the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were diagnosed with bilateral knee osteoarthritis for at least 6 months
Exclusion criteria	People who did not have any communication and psychiatric problem; whose VAS-P scores were 4 and above according to the pain scale; who did not have acute trauma, inflammation or oedema on their legs; did not have malignity; did not have circulatory disorder and peripheral vascular disease; did not receive intra-articular steroid treatment and physical therapies in the last 6 months
Recruitment/selection of patients	People with a diagnosis of knee osteoarthritis who applied to the Physical Therapy and Rehabilitation polyclinic of a university hospital of a city in Turkey between January 2014 and February 2015
Age, gender and ethnicity	Age - Mean (SD): 57.9 (10.8). Gender (M:F): 10:52. Ethnicity: Not stated/unclear

Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: High morbidity score (People with any other chronic disease: 39. People without chronic disease: 23.). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Stage 2-4, median stage 3 Duration of symptoms (SD): 32.4 (6.6) years
Indirectness of population	No indirectness
Interventions	(n=36) Intervention 1: Exercise - Unsupervised strength exercise. Combination of exercise after heat application and exercise only group. Practiced twice a day for 5 days a week for 4 weeks. Seven movements specified by the consulting physiotherapist to strengthen their muscles (in standing, sitting, lying positions). They were delivered brochures and were asked to do these exercises at home for 10 minutes twice a day, 5 days a week. The exercise and heat pack group received two hot-packs to be applied to both knees and were recommended to use them for 20 minutes, twice a day for 5 days a week. They were informed to apply hot packs in a sitting position with legs stretched out and by putting them in their cases after keeping hot-packs in boiling water for 5 minutes.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable Comments: The two groups containing exercises were pooled for the analysis as the other two comparators in the study (hot packs only and a no treatment control) could both be classified as no treatment when compared to the exercise and hot pack and exercise only arms respectively. (n=36) Intervention 2: No treatment. Combination of the heat pack and control group. The heat pack group received two hot-packs to be applied to both knees and were recommended to use them for 20 minutes, twice a day for 5 days a week. They were informed to apply hot packs in a sitting position with legs stretched out and by putting them in their cases after keeping hot-packs in boiling water for 5 minutes. The control group received no additional information apart from any usual care provided by physicians.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable Comments: The two groups containing exercises were pooled for the analysis as the other two comparators in the study (hot packs only and a no treatment control) could both be classified as no treatment when compared to the exercise and hot pack and exercise only arms respectively.
Funding	Academic or government funding (Scientific Research Projects Unit of Erciyes University, Grant/Award Number: TSA-2013-4788.)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT	

Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC disability at 4 weeks; Group 1: mean -16.97 (SD 4.69); n=30, Group 2: mean -7.72 (SD 9.03); n=32; WOMAC disability 0-68 Top=High is poor outcome; Comments: The two exercise groups and the two non-exercise groups were pooled together for the analysis. Change scores. Reported exercise and hot pack: -13.53 (1.98). Reported exercise alone: -20.40 (4.05). Reported heat: -17.06 (1.77). Reported control: 0.52 (2.40). Baseline exercise and hot pack: 39.70 (3.59). Baseline exercise alone: 42.13 (3.36). Baseline heat: 38.00 (4.22). Baseline control: 40.70 (3.72). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 6, Reason: Exercise and heat: 2 withdrew, 1 patient not reached. Exercise: 1 withdrew, 1 not reached, 1 received steroid injection.; Group 2 Number missing: 4, Reason: Heat: 1 diagnosed with cancer, 1 'patient was administered', 1 received steroid injection. Control: 1 not reached. Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 4 weeks; Group 1: mean -6.2 (SD 1.12); n=30, Group 2: mean -3.28 (SD 2.71); n=32; WOMAC pain 0-20 Top=High is poor outcome; Comments: The two exercise groups and the two non-exercise groups were pooled together for the analysis. Reported exercise and hot pack: -5.66 (0.70). Reported exercise alone: -6.73 (1.20). Reported heat: -5.86 (0.72). Reported control: -0.70 (0.95). Baseline exercise and hot pack: 12.60 (1.01). Baseline exercise alone: 13.33 (0.95). Baseline heat: 12.86 (1.09). Baseline control: 12.23 (0.78). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 6, Reason: Exercise and heat: 2 withdrew, 1 patient not reached. Exercise: 1 withdrew, 1 not reached, 1 received steroid injection.; Group 2 Number missing: 4, Reason: Heat: 1 diagnosed with cancer, 1 'patient was administered', 1 received steroid injection. Control: 1 not reached.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Karatosun 2006[223]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=105)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the knee as defined by the American College of Rheumatology criteria. All people had Kellgren Lawrence grade 3 osteoarthritis with narrowing of joint space and sclerosis of the subchondral bone
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with primary osteoarthritis of the knee as defined by the American College of Rheumatology criteria. All people had Kellgren Lawrence grade 3 osteoarthritis with narrowing of joint space and sclerosis of the subchondral bone
Exclusion criteria	People with radiographic appearance of pseudocysts (defined as Kellgren Lawrence grade 4 osteoarthritis); previous fracture around the knee; people receiving NSAIDs 15 days prior to the study; inflammatory arthritis; previous intra-articular injections or any other invasive procedure in the knee; significant comorbidity (renal, hepatic or heart disease) and chicken or egg allergy
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 56.5 (12.9). Gender (M:F): 15:90. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=53) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Progressive exercise program. Week 1 included isometric exercises (quadriceps femoris muscle), terminal knee extension exercises (quadriceps femoris muscle), stretching exercises (hamstrings and hip flexor muscles), active knee range of motion exercises, advices for daily living activities.

	Week 2/3 added strengthening exercises for hip muscles, progressive resistive exercises (Quadriceps femoris and hamstring muscles), week 6 added proprioceptive exercises, closed kinetic chain exercises. The exercises were taught on the weeks stated, and the participants otherwise performed the exercises at home.. Duration 18 months. Concurrent medication/care: No treatment with non-steroidal anti-inflammatory drugs. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Proprioceptive, strengthening, flexibility, range of motion). (n=52) Intervention 2: Pharmacological treatment - Intra-articular hyaluronic acid. Three injections of hyaluronic acid (Synvisc, Hylan G-F 20) separated by one week intervals. In bilateral cases, both knees were injected. Duration Injection over 3 weeks, follow up for 18 months. Concurrent medication/care: No treatment with non-steroidal anti-inflammatory drugs. Indirectness: No indirectness Further details: 1. Class of medicine: Intrarticular treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus INTRA-ARTICULAR HYALURONIC ACID Protocol outcome 1: Pain at > 3 months - Actual outcome: Pain during activity (VAS) at 18 months; Group 1: mean 12.1 (SD 3.1); n=53, Group 2: mean 12.9 (SD 3.4); n=52; Hospital for Special Surgery pain during activity score Range unclear Top=High is good outcome; Comments: Baseline exercise: 4.5 (4.7). Baseline hyaluronate: 4.2 (4.4). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, BMI, and baseline values for outcomes; Group 1 Number missing: 0; Group 2 Number missing: 21, Reason: 21 excluded as they had sought after additional therapy (possibly due to only receiving one intervention early on, while the exercise group got continuing care)	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Karatosun 2008[224]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Secondary ankle osteoarthritis defined by the clinical and radiographic findings (However, they ultimately ended up including people with primary osteoarthritis, 17 primary: 13 secondary)
Stratum	Overall:
Subgroup analysis within study	Not applicable
Inclusion criteria	People with secondary osteoarthritis of the ankle with a definite history of severe trauma of Kellgren Lawrence grade 3 radiographic severity.
Exclusion criteria	Appearance of definite deformity of the bony contour (Kellgren Lawrence grade 4); inflammatory arthritis; previous intra-articular injections or any other invasive procedures in the ankle; significant comorbidity (renal, hepatic or heart disease) and chicken or egg allergy
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 55.1 (12.1). Gender (M:F): 9:21. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Ankle osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). An exercise program taught over 6 weeks of progressive, simple, isometric, isotonic range of motion, resistance, closed kinetic chain and proprioceptive exercises, taught over 4 visits. People then repeated these exercises at home.. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other

	(n=15) Intervention 2: Pharmacological treatment - Intra-articular hyaluronic acid. The hyaluronic acid group received three injections of hyaluronic acid at 1 week intervals by the same physician. The dose of hyaluronic acid was 2.5mg in each injection. The injection was performed with the person in half lying position with the knee flexed and the foot flat on the plinth. Then the anterior ankle joint line was palpated and the needle was inserted slightly upward in order to run upper surface of the talus which is slightly convex. When it was felt that the capsule was passed, then the joint fluid was aspirated if present, and then hyaluronic acid was injected. People were advised not to take part in strenuous activity for a few days. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Intrarticular treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus INTRA-ARTICULAR HYALURONIC ACID Protocol outcome 1: Pain at > 3 months - Actual outcome: Pain during activity (VAS) at 12 months; Group 1: mean 2.4 (SD 3.1); n=15, Group 2: mean 1.4 (SD 1.9); n=15; Visual analogue score 0-10 Top=High is poor outcome; Comments: Baseline exercise: 2.1 (2.4). Baseline hyaluronic acid: 2.4 (3.0). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, gender, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kars Fertelli 2018[225]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee or hip osteoarthritis as diagnosed by the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis who were in the 25 years and above age group, who were able to communicate and walk, whose pain level score was at least 5 according to the WOMAC scale, who had a medical report indicating their eligibility for aquatic exercise, who lived within the municipal boundaries of Sivas, and who volunteered to participate in study and individuals receiving pharmacological treatment
Exclusion criteria	People who had previously undergone hip or knee joint surgery; who had rheumatoid arthritis, hypertension or myocardial infarction; who had undergone intra-articular corticosteroid therapy in the last month
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 55.6 (7.8). Gender (M:F): 10:110. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee).
Extra comments	Severity: Median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). 40 minute long aquatic exercise program including: 10 minutes of warming up exercises, 20 minutes of basic exercises, and 10 minutes of cooling down exercises. During the study, not to put the body under much strain, the intensity and repetition of the exercises were increased gradually (8-15 repetitions, one to three sets). During the exercise, swim foam boards and balls were used to help the participants move their joints more easily. Completed 3 days a week for 8 weeks..

	Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=60) Intervention 2: Exercise - Unsupervised strength exercise. People in the control group were informed about how to do exercises that should be done by people with osteoarthritis and told to do exercises at home. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus UNSUPERVISED STRENGTH EXERCISE Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 26.1 (SD 15.59); n=60, Group 2: mean 46.9 (SD 17.22); n=60; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline hydro: 50.32 (16.04). Baseline home: 50.97 (11.70). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, job, educational status, affected joints, classification of diseases, exercising daily life situations, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 7 (SD 4.44); n=60, Group 2: mean 14.43 (SD 6.4); n=60; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline hydro: 14.43 (3.82). Baseline home: 15.35 (4.41). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, job, educational status, affected joints, classification of diseases, exercising daily life situations, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3

	months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kawasaki 2009[227]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=102)
Countries and setting	Conducted in Japan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary osteoarthritis of the medial femorotibial compartment of the knee according to the clinical and radiographic criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Postmenopausal females (>50 years of age) with primary osteoarthritis of the medial femorotibial compartment of the knee and no other inflammatory diseases according to the clinical and radiographic criteria of the American College of Rheumatology as well as standard exclusion criteria
Exclusion criteria	'Standard exclusion criteria' - no additional information
Recruitment/selection of patients	People who visited five hospitals (Juntendo University Hospital, Juntendo University Urayasu Hospital, Juntendo University Nerima Hospital, Koto Hospital, Tokyo Rinkai Hospital)
Age, gender and ethnicity	Age - Mean (SD): 70.4 (7.8). Gender (M:F): 0:102. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Isometric muscle exercises of the bilateral lower limbs: one set (each exercise done 20 times) of straight leg raising training and hip abduction and adduction exercises performed twice a day. Range or motion exercises: maximum flexion and maximum extension performed twice a day in the morning and evening after the knee was warmed. They also recommended to walk as much as they could without pain during their daily living.. Duration 8 weeks. Concurrent

	medication/care: All people were supplied with 100mg sodium loxoprofen tablets for pain rescue analgesia (300mg/day maximum allowed use) in the treated knee only. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and range of motion). (n=60) Intervention 2: Pharmacological treatment - Intra-articular hyaluronic acid. Intraarticular injections of 25mg/2.5mL hyaluronate sodium (Artz) in the affected knee once a week for the first 5 weeks. This was followed by a once-a-month injection to maintain effects until the 24th week. All treatments were performed under aseptic conditions and after aspirating any existing effusion as completely as possible. The frequency of the injection was determined by the precautions given by the pharmaceutical firm and with reference to past reports.. Duration 8 weeks. Concurrent medication/care: All people were supplied with 100mg sodium loxoprofen tablets for pain rescue analgesia (300mg/day maximum allowed use) in the treated knee only. Indirectness: No indirectness Further details: 1. Class of medicine: Intrarticular treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus INTRA-ARTICULAR HYALURONIC ACID

Protocol outcome 1: Pain at > 3 months
- Actual outcome: Visual analogue scale at 24 weeks; Group 1: mean -21.29 (SD 27.6); n=60, Group 2: mean -20.46 (SD 36.04); n=60; VAS 0-100 Top=High is poor outcome; Comments: Baseline exercise: 55.2 (22.6). Baseline hyaluronic acid: 59.8 (22.7)
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, joint space width, femorotibial angle and baseline values of outcomes; Group 1 Number missing: 10, Reason: 10 withdrew: 1 other treatments, 4 poor execution, 5 lost to follow up; Group 2 Number missing: 18, Reason: 18 withdrew: 8 other treatments, 10 lost to follow up

Protocol outcome 2: Serious adverse events at > 3 months
- Actual outcome: Severe adverse events, such as worsening pain, effusion, synovitis, haemarthrosis, or septic arthritis at 24 weeks; Group 1: 0/60, Group 2: 0/60
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, joint space width, femorotibial angle and

baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	Keefe 2004[228]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=72)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Persistent knee pain due to osteoarthritis and were diagnosed as having osteoarthritis of the knees
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Married people with persistent knee pain due to osteoarthritis and who were diagnosed as having osteoarthritis of the knees and their respective spouses
Exclusion criteria	Comorbid medical conditions that could affect their health status over the course of the trial (e.g. a recent myocardial infarction), a n abnormal cardiac response to exercise (e.g. exercise-induced ventricular tachycardia, abnormal blood pressure response); or other known organic disease that would contraindicate safe participation in the study (e.g. chronic obstructive pulmonary disease, congestive heart failure, or cancer).
Recruitment/selection of patients	Recruited from rheumatology clinics and advertisements placed in newspapers
Age, gender and ethnicity	Age - Mean (SD): 59.50 (11.36). Gender (M:F): 33:39. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=16) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Three supervised group exercise sessions per week for 12 consecutive week. Their spouses did not attend the exercise sessions. Included: cardiopulmonary endurance training; strength training and flexibility/range of motion training. People participated in 30 minutes of aerobic training three days a week at an intensity of 50-70% of heart rate reserve, gradually increased to 70-85% over 12 weeks. These sessions involved a warm up, low intensity biking or walking, 30

	minutes of continuous aerobic activity (walking, biking or water aerobics) and a cool down period. People also participated in 30 minutes of strength training two days per week.. Duration 12 weeks. Concurrent medication/care: People were allowed to continue to receive their routine care. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength and aerobic). (n=18) Intervention 2: No treatment. No exercise care. Duration 12 weeks. Concurrent medication/care: People were allowed to continue to receive their routine care. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=38) Intervention 3: Other. Spouse assisted coping skills training, with of without exercise. Duration 12 weeks. Concurrent medication/care: People were allowed to continue to receive their routine care. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These two groups were not included in the analysis as they did not fulfill the inclusion criteria
Funding	Academic or government funding (This research was supported by National Institute of Arthritis and Musculoskeletal Diseases Grant No. AR-35270)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Pain at </=3 months
- Actual outcome: AIMS pain at 12 weeks; Group 1: mean 3.19 (SD 1.85); n=16, Group 2: mean 4.03 (SD 2.08); n=18; Arthritis Impact Measurement Scale pain 0-10 Top=High is poor outcome; Comments: Baseline exercise: 3.91 (1.64). Baseline no treatment: 3.91 (1.73).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: AIMS psychological disability difference at baseline; Group 1 Number missing: 0; Group 2 Number missing: 2, Reason: No additional information

Protocol outcome 2: Psychological distress at </=3 months
- Actual outcome: AIMS psychological disability at 12 weeks; Group 1: mean 1.88 (SD 0.87); n=16, Group 2: mean 1.8 (SD 1.04); n=18; Arthritis Impact Measurement Scale psychological disability 0-10 Top=High is poor outcome; Comments: Baseline exercise: 2.36 (1.22). Baseline no exercise: 1.85 (0.33).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: AIMS psychological disability difference at baseline; Group 1 Number missing: 0; Group 2 Number missing: 2, Reason: No additional information	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Khruakhorn 2021[233]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=34)
Countries and setting	Conducted in Thailand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (end of intervention) and 6 months after end of intervention
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee diagnosed with radiography
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Define
Exclusion criteria	Define
Recruitment/selection of patients	No additional information

Age, gender and ethnicity	Age - Mean (SD): 61.4 (8.4). Gender (M:F): Define. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren-Lawrence grade 2-3 Duration of symptoms: Not stated/unclear. Thai Clinical Trials Registry identification number: TCTR20170527001.
Indirectness of population	No indirectness
Interventions	(n=17) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Hydrotherapy with progressive strengthening exercises. Both groups attended the exercise classes for 45-60 minutes, three times per week for 6 weeks. For the strengthening exercises, the participants had to hold the position for 10 seconds in 10 sets. For the strengthening exercises, 15 repetitions of three sets and cycling for 10 and 15 minutes in the second and third phases were conducted. There was a minute of rest between exercises. The phase was changed every six sessions of exercise, including the number of exercises, extra resistance and time duration. Floatation noodles were used to strengthen the knee extensor. Cycling in water with floating noodles were used to enhance total leg muscle. Jogging in water was used for the calf muscle. The exercises were performed at a hydrotherapy pool (32-33 degrees centigrade). Noodles and water floatation were used for extra water resistance in strengthening exercises and deep water cycling. Both exercise groups performed the exercises under the supervision of a physiotherapist.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=17) Intervention 2: Exercise - Supervised strength exercise. Land based therapy with progressive strengthening exercises. Both groups attended the exercise classes for 45-60 minutes, three times per week for 6 weeks. For the strengthening exercises, the participants had to hold the position for 10 seconds in 10 sets. For the strengthening exercises, 15 repetitions of three sets and cycling for 10 and 15 minutes in the second and third phases were conducted. There was a minute of rest between exercises. The phase was changed every six sessions of exercise, including the number of exercises, extra resistance and time duration. Land-based exercises were performed on an exercise mat for 45-60 minutes per session, 3 sessions per week for 6 weeks. Elastic bands were used to strengthen the knee extensor. A stationary bike was used to enhance total leg muscle. Tiptoe was used for the calf muscle. Both exercise groups performed the exercises under the supervision of a physiotherapist.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (The Thammasat University research funding supported this study.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months

- Actual outcome: WHO Quality of Life Total Score at 6 weeks; Group 1: mean 93.06 (SD 5.8); n=17, Group 2: mean 92.88 (SD 12.23); n=17; WHO quality of life total 0-100 Top=High is good outcome; Comments: Baseline hydrotherapy: 83.88 (10.99). Baseline land-based: 86.76 (11.91). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Health related quality of life at > 3 months - Actual outcome: WHO Quality of Life Total Score at 6 months; Group 1: mean 98.18 (SD 3.43); n=17, Group 2: mean 96.24 (SD 8.05); n=17; WHO Quality of Life Total Score 0-100 Top=High is good outcome; Comments: Baseline hydrotherapy: 83.88 (10.99). Baseline land-based: 86.76 (11.91). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Physical function at </=3 months - Actual outcome: WOMAC function at 6 weeks; Group 1: mean 20.24 (SD 18.81); n=17, Group 2: mean 24.35 (SD 28.61); n=17; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline hydrotherapy: 51.47 (36.43). Baseline land-based: 64.29 (32.41). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Pain at </=3 months - Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 7.47 (SD 6.85); n=17, Group 2: mean 7.94 (SD 9.22); n=17; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline hydrotherapy: 17.53 (12.44). Baseline land-based: 18.82 (10.89). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kigozi 2018[234]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=514)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Current knee pain and/or stiffness in one or both knees who met the criteria recommended by the National Institute for Health and Care Excellence guidelines for a clinical diagnosis of knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults aged at least 35 years with current knee pain and/or stiffness in one or both knees who met the criteria recommended by the National Institute for Health and Care Excellence guidelines for a clinical diagnosis of knee osteoarthritis were invited to take part.
Exclusion criteria	Those with potentially serious pathology (such as inflammatory arthritis, malignancy); those who have had a total hip or knee replacement on the affected side; those who are on a waiting list for a total knee or hip replacement; those for whom their knee problem was caused by a recent trauma (sports injury, fall or accident),; those for whom exercise interventions are contra-indicated (such as those with unstable cardiovascular disorders, severe hypertension, unstable angina or congestive heart failure); those who have received an exercise programme for a physiotherapist or a knee joint injection in the last three months; those residing in nursing home accomodation; those who are so severely physically restricted that they cannot get to the physiotherapy treatment centres and those who have a close family member already participating in the BEEP trial
Recruitment/selection of patients	People were recruited from up to 100 general practices and their local physiotherapy services in the West Midlands and North West regions of the UK
Age, gender and ethnicity	Age - Other: Mean: 63 years. Gender (M:F): 51% were female. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not applicable 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated

Indirectness of population	No indirectness
Interventions	(n=176) Intervention 1: Exercise - Supervised strength exercise. A supervised, individually tailored and progressed lower limb exercise programme provided in six to eight one-to-one treatment sessions over 12 weeks. Participants received a print-out of a specific exercise prescription individualised for them based on their progress on the programme. . Duration 12 weeks (18 months follow up). Concurrent medication/care: All participants received an information booklet providing information about benefits of exercise and physical activity and a home exercise program. Usual physical therapy care included advice and lower-limb exercise provided in up to four individual one-to-one treatment sessions over 12 weeks, in line with usual physical therapy practice in the National Health Service. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=175) Intervention 2: No treatment. Usual care only. Duration 18 months. Concurrent medication/care: All participants received an information booklet providing information about benefits of exercise and physical activity and a home exercise program. Usual physical therapy care included advice and lower-limb exercise provided in up to four individual one-to-one treatment sessions over 12 weeks, in line with usual physical therapy practice in the National Health Service. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=163) Intervention 3: Other. Targeted exercise adherence with face to face and telephone contact to help improve general physical activity adherence over 6 months. Duration 6 months (18 months follow up). Concurrent medication/care: All participants received an information booklet providing information about benefits of exercise and physical activity and a home exercise program. Usual physical therapy care included advice and lower-limb exercise provided in up to four individual one-to-one treatment sessions over 12 weeks, in line with usual physical therapy practice in the National Health Service. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for

	Applied Research (grant number: RP-PG-0407-10386) and the Arthritis Research UK Centre in Primary Care grant (grant number: 18139).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: EQ-5D at 3 months; Group 1: mean 0.708 (SD 0.188); n=176, Group 2: mean 0.686 (SD 0.201); n=175; EQ-5D -0.11-1 Top=High is good outcome; Comments: Baseline exercise: 0.644 (0.229). Baseline no treatment: 0.636 (0.230).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Only reports the baseline value for the outcome; Group 1 Number missing: 36, Reason: Reports that only 80% completed the trial and 78% were available at 18 months; Group 2 Number missing: 35, Reason: Reports that only 80% completed the trial and 78% were available at 18 months

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: EQ-5D at 18 months; Group 1: mean 0.7 (SD 0.206); n=176, Group 2: mean 0.7 (SD 0.219); n=175; EQ-5D -0.11-1 Top=High is good outcome; Comments: Baseline exercise: 0.644 (0.229). Baseline no treatment: 0.636 (0.230).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Only reports the baseline value for the outcome; Group 1 Number missing: 39, Reason: Reports that only 80% completed the trial and 78% were available at 18 months; Group 2 Number missing: 39, Reason: Reports that only 80% completed the trial and 78% were available at 18 months

Protocol outcomes not reported by the study	Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Knoop 2013[237] (Knoop 2015[238], Knoop 2014[239])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=159)
Countries and setting	Conducted in Netherlands; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks of treatment, 6 months in all
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of knee osteoarthritis according to clinical American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of knee osteoarthritis according to clinical American College of Rheumatology criteria; age between 40 and 75 years; presence of self-reported and/or biomechanically assessed knee instability. Self reported knee instability was defined as at least one episode of buckling, shifting or giving way of the knee in the past 3 months, reported by the person. Biomechanically assessed knee instability was defined as the presence of muscle weakness in combination with presence of: impaired proprioceptive accuracy and/or high passive varus-valgus laxity.
Exclusion criteria	Other forms of arthritis than osteoarthritis (e.g. crystal arthropathy, septic arthritis, spondyloarthropathy) identified by radiography and/or blood and urine samples; presence of comorbidity resulting in severe activity limitations; total knee arthroplasty or total knee arthroplasty in the near future; severe knee pain (NRS >8); insufficient comprehension of Dutch language; inability to be scheduled for therapy; unwillingness to give informed consent
Recruitment/selection of patients	People were recruited through advertisements in local and regional newspapers and from regular referral from rheumatologists or rehabilitation physicians from our rehabilitation center
Age, gender and ethnicity	Age - Mean (SD): 62.0 (7.1). Gender (M:F): 62:97. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 0-4, median grade 2 Duration of symptoms (mean [SD]): 10.8 (9.3) years
Indirectness of population	No indirectness

Interventions	(n=80) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). An exercise program of 12 weeks with 2 sessions of 60 minutes weekly and a home exercise program for 5 days weekly with gradual increase in training intensity, knee load and exercise difficulty. The exercise consisted of three phases: first phase (week 1-4) targeting knee joint stabilisation, second phase (week 5-8) targeting muscle strength in addition to knee joint stabilisation, third phase (week 9-12) targeting performance of daily activities in addition to the previous components. People were encouraged to remain active after the intervention. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Proprioceptive and strength). (n=79) Intervention 2: Exercise - Supervised strength exercise. Exercise therapy sessions of 50 minutes twice weekly focusing on muscle strength (hydrotherapy in first week, land based therapy from week 2), including home exercises for 5 days a week, similar to the experimental group. Two phases: first phase (week 1-8) targeting muscle strength, second phase (week 9-12) targeting performance of daily activities in addition to muscle strength. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by the Dutch Arthritis Association)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 12 weeks; Group 1: mean 17.4 (SD 11.6); n=80, Group 2: mean 19.3 (SD 11.4); n=79; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 26.2 (11.8). Baseline strength: 27.1 (12.7).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, BMI, unilateral/bilateral, radiographic severity, education level, comorbidity score, use of analgesia, use of walking device, alignment, biomechanical outcomes, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 protocol violators - 2 discontinued treatment due to health condition, 4 missed more than 8 sessions; Group 2 Number missing: 7, Reason: 4 lost to follow up - 3 withdrawals due to health condition, 1 lack of time. 3 protocol violators, 3 missed more

than 8 sessions. Protocol outcome 2: Physical function at > 3 months - Actual outcome: WOMAC physical function at 38 weeks; Group 1: mean 18.9 (SD 13.3); n=80, Group 2: mean 19.2 (SD 13.2); n=79; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 25.2 (11.8). Baseline control: 27.1 (12.7). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, BMI, unilateral/bilateral, radiographic severity, education level, comorbidity score, use of analgesia, use of walking device, alignment, biomechanical outcomes, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 protocol violators, 2 discontinued treatment due to health condition, 4 missed more than 8 sessions, 1 total knee arthroplasty in the follow up period, 1 knee arthroscopy in the follow up period; Group 2 Number missing: 9, Reason: 4 lost to follow up - 3 withdrawals due to health condition, 1 lack of time. 5 protocol violators, 3 missed >8 sessions, 2 total knee arthroplasties in the follow up period Protocol outcome 3: Pain at </=3 months - Actual outcome: NRS (knee pain severity) at 12 weeks; Group 1: mean 2.8 (SD 2.1); n=80, Group 2: mean 3.3 (SD 2.1); n=79; NRS 0-10 Top=High is poor outcome; Comments: Baseline mixed: 4.8 (2.2). Baseline strength: 5.2 (2.0). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, BMI, unilateral/bilateral, radiographic severity, education level, comorbidity score, use of analgesia, use of walking device, alignment, biomechanical outcomes, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 protocol violators - 2 discontinued treatment due to health condition, 4 missed more than 8 sessions; Group 2 Number missing: 7, Reason: 4 lost to follow up - 3 withdrawals due to health condition, 1 lack of time. 3 protocol violators, 3 missed more than 8 sessions. Protocol outcome 4: Pain at > 3 months - Actual outcome: NRS (knee pain severity) at 38 weeks; Group 1: mean 3.1 (SD 2.5); n=80, Group 2: mean 3.7 (SD 2.4); n=79; NRS 0-10 Top=High is poor outcome; Comments: Baseline mixed: 4.8 (2.2). Baseline strength: 5.2 (2.0). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, BMI, unilateral/bilateral, radiographic severity, education level, comorbidity score, use of analgesia, use of walking device, alignment, biomechanical outcomes, and baseline values of outcomes; Group 1 Number missing: 8, Reason: 8 protocol violators, 2 discontinued treatment due to health condition, 4 missed more than 8 sessions, 1 total knee arthroplasty in the follow up period, 1 knee arthroscopy in the follow up period; Group 2 Number missing: 9, Reason: 4 lost to follow up - 3 withdrawals due to health condition, 1 lack of time. 5 protocol violators, 3 missed >8 sessions, 2 total knee arthroplasties in the follow up period	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Kraus 2014[242] (Krauss 2011[244])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=218)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of one or both hip joints according to the clinical criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age between 18 and 85 years; osteoarthritis of one or both hip joint(s) (clinical criteria of the American College of Rheumatology); the subject gives voluntary consent to study participation after receiving oral and written information about study content and objectives; the subject has the time available to undertake the exercises and attend the measurings; the subject is physically fit for the intervention measure (as ascertained during the examination conducted by the principal investigator). "Fitness" in this setting relates to the physical as well as the psychological condition of the subject (subjects will not be excluded if they have one hip endoprosthesis, as long as the contralateral hip is affected by osteoarthritis according to the listed criteria); the subject has capacity to consent
Exclusion criteria	Unstable anchoring of endoprosthetic hip joint; hip dislocation after endoprosthetic joint replacement; further disorders affecting the lower extremities or lower back that require treatment by a physician/therapist and which are not connected to the osteoarthritis and are currently being treated; the presence of osteoarthritis in several joints (for example, hip and knee) is NOT and exclusion criteria; medication or alcohol misuse; participation in a clinical study in the preceding 4 weeks; lack of compliance; acute illness; use of walking aids; previous trauma in the hip and pelvis area with accompanying development of secondary osteoarthritis; known endocrinological causes of hip osteoarthritis; confirmed metabolic causes of hip osteoarthritis; state after aseptic bone necrosis (Perthes' disease); cardiocirculatory disorders or other comorbidities that result in severely restricted everyday physical capacity and that are contraindications to physical exertion (for example, heart failure NYHA III-IV, terminal renal failure stage IV); medical exercise therapy, physiotherapy on resistance machines in the preceding 3 months; with a total treatment frequency of more than 6 units; systematic group or individual therapy to treat the osteoarthritis (systematic in

	the sense of a minimum of 1/week for 30 minutes or more) in the preceding 3 months; physical therapy to treat the osteoarthritis (systematic in the sense of regular, prescribed application at least 1x/week) in the preceding 3 months; newly initiated exercise/movement therapy in the receding 3 months (sports and movement therapy defined as taking place a minimum of 1x/week, getting out of breath, minimum duration 30 minutes); corticosteroid injection into the hip joint in the preceding 12 months
Recruitment/selection of patients	People were recruited via sports orthopedics outpatient clinics and via the press
Age, gender and ethnicity	Age - Mean (SD): 59 (10). Gender (M:F): 130:88. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=71) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Tübinger exercise therapy approach - entailed a once-weekly group intervention (60-90 minutes) in addition to a twice-weekly home exercise program (30-40 minutes each). The therapeutic program entailed education and social interaction as well as exercises to strengthen the muscles and to improve proprioception, balance and flexibility.. Duration 12 weeks. Concurrent medication/care: No additional information Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength, proprioception, flexibility). (n=69) Intervention 2: No treatment. No exercise control. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=78) Intervention 3: Other. Ultrasound therapy or placebo ultrasound therapy. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These groups were not included in the analysis as they did not fulfill the

	inclusion criteria
Funding	Equipment / drugs provided by industry (The study was supported with training aterials by the companies Theraband and Ludwig Artzt)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 bodily pain subscale at 12 weeks; Group 1: mean 5.2 (SD 17.6); n=70, Group 2: mean -0.1 (SD 17.3); n=68; SF-36 bodily pain subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 57.9 (18.4). Baseline control: 56.6 (17.5).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 physical functioning subscale at 12 weeks; Group 1: mean 2 (SD 14); n=71, Group 2: mean 2 (SD 18); n=69; SF-36 physical functioning subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 66 (20). Baseline control: 65 (20).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 role-physical subscale at 12 weeks; Group 1: mean 2 (SD 35); n=71, Group 2: mean 3 (SD 33); n=69; SF-36 role-physical subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 79 (34). Baseline control: 74 (33).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 general health subscale at 12 weeks; Group 1: mean 3 (SD 14); n=71, Group 2: mean 0 (SD 16); n=69; SF-36 general health subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 65 (16). Baseline control: 68 (14).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy,

supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 vitality subscale at 12 weeks; Group 1: mean -1 (SD 15); n=71, Group 2: mean 0 (SD 12); n=69; SF-36 vitality subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 64 (17). Baseline control: 63 (17).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 social functioning subscale at 12 weeks; Group 1: mean -2 (SD 13); n=71, Group 2: mean -2 (SD 15); n=69; SF-36 social functioning subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 91 (24). Baseline control: 87 (18).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 role-emotional subscale at 12 weeks; Group 1: mean 1 (SD 27); n=71, Group 2: mean 2 (SD 14); n=69; SF-36 role-emotional subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 91 (24). Baseline control: 93 (20).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)
- Actual outcome: SF-36 mental health subscale at 12 weeks; Group 1: mean -1 (SD 11); n=71, Group 2: mean -2 (SD 10); n=69; SF-36 mental health subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 80 (13). Baseline control: 80 (13).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4)

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function subscale at 12 weeks; Group 1: mean -8.4 (SD 13.4); n=71, Group 2: mean -2.1 (SD 12.9); n=68; WOMAC

physical function subscale 0-100 Top=High is poor outcome; Comments: Baseline exercise: 26.5 (17.5). Baseline control: 26.7 (15.5). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4) Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 12 weeks; Group 1: mean -8.5 (SD 13.9); n=71, Group 2: mean -1.3 (SD 15.3); n=69; WOMAC pain subscale 0-100 Top=High is poor outcome; Comments: Baseline exercise: 27.5 (16.7). Baseline control: 28.3 (16.9). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 6, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4); Group 2 Number missing: 2, Reason: Reports overall reasons for withdrawal: Lack of compliance with therapy, supplementary therapy (8), orthopedic/internal medical exclusion criteria (2), surgery (1), illness on follow up (4) Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Leaving the study due to adverse events at 12 weeks; Group 1: 0/71, Group 2: 0/69 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reported BMI, age, SF-36 subscales and WOMAC subscale; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Kumar 2013[248]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=44)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People referred with knee osteoarthritis (diagnosed radiologically or clinically)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age between 40 and 65 years involving either one or both knee joints
Exclusion criteria	Grade IV knee osteoarthritis on the Kellgren-Lawrence Scale; any systemic infection; neurological or vestibular disorder; deformity of the back, hip, knee and ankle; history of either knee trauma during the last 3 months or knee surgery; uncontrolled cardiac insufficiency; clinically significant anteroposterior or mediolateral instability of knee; taken steroid injection within 6 months in knee joint; uncooperative person
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 53.3 (6.2). Gender (M:F): 19:25. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear (Mixed). 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=22) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Resistive exercise (beginning with knee flexor and extensor strengthening on quads table and after that on a plinth for hip flexors, hip extensors, hip abductors and hip external rotators. Exercises include 3 sets of 10 repetitions of open chain exercises for each exercise. Exercises were performed in a medium slow rate. The rest period between repetition and sets were 30s and 60s respectively, and 5 minutes between exercises or legs. Resistance was gradually increased every week by 10%) with proprioceptive training (given on alternate days of resistive exercise, gradually increased in difficulty, no additional information). Duration

	4 weeks. Concurrent medication/care: Ultrasound therapy (frequency 1mHz, spatial and temporal peak intensity of 2.5W/cm² and pulsed at a duty cycle of 25% for 5 minutes).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and proprioception). (n=22) Intervention 2: Exercise - Supervised strength exercise. Resistive exercise only (beginning with knee flexor and extensor strengthening on quads table and after that on a plinth for hip flexors, hip extensors, hip abductors and hip external rotators. Exercises include 3 sets of 10 repetitions of open chain exercises for each exercise. Exercises were performed in a medium slow rate. The rest period between repetition and sets were 30s and 60s respectively, and 5 minutes between exercises or legs. Resistance was gradually increased every week by 10%). Duration 4 weeks. Concurrent medication/care: Ultrasound therapy (frequency 1mHz, spatial and temporal peak intensity of 2.5W/cm² and pulsed at a duty cycle of 25% for 5 minutes).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 4 weeks; Group 1: mean 6.95 (SD 2.34); n=22, Group 2: mean 10 (SD 3.49); n=22; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 23.32 (1.67). Baseline strength: 23.59 (2.56).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Uses the 'chit box' method of randomisation - essentially pulling cases out of a box until everyone is allocated to a group; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, and baseline values for outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: Pain intensity (NRS) at 4 weeks; Group 1: mean 2.18 (SD 0.66); n=22, Group 2: mean 2.91 (SD 0.81); n=22; NRS 0-10 Top=High is poor outcome; Comments: Baseline mixed: 4.86 (0.99). Baseline strength: 5.14 (0.77).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - Uses the 'chit box' method of randomisation - essentially pulling cases out of a box until everyone is allocated to a group; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, and baseline values for

outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kuptniratsaikul 2002[250]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=392)
Countries and setting	Conducted in Thailand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 month
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the knee, Kellgren Lawrence grade 2-3
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee with grade 2 or 3 osteoarthritis, as judged by the criteria of Kellgren and Lawrence, based on weight-bearing radiographs.
Exclusion criteria	No additional information
Recruitment/selection of patients	Volunteers from the urban community of the Bangkok Metropolitan area around Siriraj Hospital
Age, gender and ethnicity	Age - Mean (SD): 67.8 (5.9). Gender (M:F): 86:306. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Median Mild Duration of knee pain (mean [SD]): 45.3 (40.2) months
Indirectness of population	No indirectness
Interventions	(n=193) Intervention 1: Exercise - Supervised strength exercise. An exercise class, emphasising quadriceps muscle strengthening, for two sessions per week lasting 1 hour. This was continued for 8 weeks.. Duration 8 weeks, followed up for 12 months. Concurrent medication/care: People were allowed to continue their usual medical treatments. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Not applicable (n=199) Intervention 2: No treatment. No exercise. Duration 12 months. Concurrent medication/care: People were allowed to continue their usual medical treatments. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by the National Research Council of Thailand)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: Modified Bandi's criteria of functional incapacity scale score at 2 months; Group 1: mean 6.08 (SD 3.14); n=193, Group 2: mean 6.38 (SD 3.58); n=199; Modified Bandi's criteria of functional incapacity scale score 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.74 (3.15). Baseline control: 6.90 (3.75).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, duration of symptoms, income, education, medication use, BMI, severity of knee osteoarthritis and baseline values of outcomes; Group 1 Number missing: 20, Reason: 173 completed the 8 week study; Group 2 Number missing: 6, Reason: 193 completed the 8 week study period

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: Modified Bandi's criteria of functional incapacity scale score at 12 months; Group 1: mean 5.28 (SD 3.46); n=193, Group 2: mean 6.32 (SD 3.63); n=199; Modified Bandi's criteria of functional incapacity scale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.74 (3.15). Baseline control: 6.90 (3.75).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, duration of symptoms, income, education, medication use, BMI, severity of knee osteoarthritis and baseline values of outcomes; Group 1 Number missing: 20, Reason: 173 completed the 8 week study. Doesn't state for after this period.; Group 2 Number missing: 6, Reason: 193 completed the 8 week study period. Doesn't state for after this period.

Protocol outcome 3: Pain at </=3 months
- Actual outcome: Pain (VAS) at 2 months; Group 1: mean 4.14 (SD 2.28); n=193, Group 2: mean 5.15 (SD 2.26); n=199; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.35 (2.01). Baseline control: 5.71 (2.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, duration of symptoms, income, education, medication use, BMI, severity of knee osteoarthritis and baseline values of outcomes; Group 1 Number missing: 20, Reason: 173 completed the 8 week study; Group 2 Number missing: 6, Reason: 193 completed the 8 week study period

Protocol outcome 4: Pain at > 3 months
- Actual outcome: Pain (VAS) at 12 months; Group 1: mean 4.25 (SD 2.7); n=193, Group 2: mean 4.57 (SD 2.69); n=199; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.35 (2.01). Baseline control: 5.71 (2.0).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports sex, age, duration of symptoms, income, education, medication use, BMI, severity of knee osteoarthritis and baseline values of outcomes; Group 1 Number missing: 20, Reason: 173 completed the 8 week study. Doesn't state for after this period.; Group 2 Number missing: 6, Reason: 193 completed the 8 week study period. Doesn't state for after this period.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Kuptniratsaikul 2019[249]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in Thailand; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis with mild to moderate knee pain
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of primary knee osteoarthritis, age 50-80 years, mild to moderate knee pain with numeric rating scale or at least 5, and body mass index of at least 25kg/m²
Exclusion criteria	Bowel and/or bladder incontinence; skin ulcer; inability to walk due to a serious medical condition (e.g. cardiopulmonary system, spinal sternosis, or severe back, hip or ankle joint pain)
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 61.9 (6.7). Gender (M:F): 5:75. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (median [range}): Home exercise: 4.0 (0.2, 20.0), UTM: 3.0 (0.1, 30.0) years
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Underwater treadmill exercise with moderate intensity (NRS 5-6/10) for 30 minutes, including warm up and cool down three times per week for 4 weeks and training from a certified physical therapist on how to perform quadriceps exercise, being asked to repeat this 10-20 repetitions/set with a 1-2 minute rest, routine daily while at home.. Duration 4 weeks. Concurrent medication/care: All participants received a leaflet advising them on how to use their knee joints in daily practice (i.e. warm compression for pain relief, regular isometric quadriceps exercise, and avoid bending the knee more than 90 degrees).. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=40) Intervention 2: Exercise - Unsupervised strength exercise. Training from a certified physical therapist on how to perform quadriceps exercise, being asked to repeat this 10-20 repetitions/set with a 1-2 minute rest, routine daily while at home.. Duration 4 weeks. Concurrent medication/care: All participants received a leaflet advising them on how to use their knee joints in daily practice (i.e. warm compression for pain relief, regular isometric quadriceps exercise, and avoid bending the knee more than 90 degrees).. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This research project was supported by a grant from the Faculty ofmedicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (grant no. (IO) R015733003).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Pain score (NRS) at 4 weeks; Group 1: mean -2.3 (SD 1.9); n=40, Group 2: mean -1.8 (SD 1.7); n=40; NRS 0-10 Top=High is poor outcome; Comments: Baseline hydro: 6.5 (1.3). Baseline home exercise: 6.4 (1.1).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, disease duration, sides affected, use of gait aids, use of knee support, use of pain medication, knee exercise and baseline values for outcomes; Group 1 Number missing: 7, Reason: 7 lost to follow up: 1 severe knee pain, 1 unable to contact, 5 inconvenience; Group 2 Number missing: 3, Reason: 3 lost to follow up: 2 unable to contact, 1 inconvenience

Protocol outcome 2: Serious adverse events at </=3 months
- Actual outcome: Joint pain, muscle pain and others at 4 weeks; Group 1: 8/40, Group 2: 14/40; Comments: Hydrotherapy: 3 joint pain, 4 muscle pain, 1 other. Home exercise: 4 joint pain, 7 muscle pain, 5 others.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, disease duration, sides affected, use of gait aids, use of knee support, use of pain medication, knee exercise and baseline values for outcomes; Group 1 Number missing: 7, Reason: 7 lost to follow up: 1 severe knee pain, 1 unable to contact, 5 inconvenience; Group 2 Number missing: 3, Reason: 3 lost to follow up: 2 unable to

contact, 1 inconvenience	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Kuru colak 2017[251]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=78)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Kellgren Lawrence grade 2-3 knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 45 years or older; Kellgren-Lawrence grade 2-3 osteoarthritis determined clinically and radiographically
Exclusion criteria	History of surgery or malunion of fractures in the lower extremity; infection or malignancy; vestibular problems; uncontrolled hypertension; chronic disease or disability that would inhibit completion of the program; history of heart disease or cerebrovascular attack; contraindication to the ability to undertake the exercise training; injections or other invasive joint therapies
Recruitment/selection of patients	People who presented at the Department of Orthopedics and traumatology, Dr. Lutfi Kirdar Kartal Training and Research Hospital
Age, gender and ethnicity	Age - Mean (range): 60 (49-84). Gender (M:F): 17:39. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=39) Intervention 1: Exercise - Supervised strength exercise. Low-intensity therapeutic isometric and isotonic exercises for major muscle groups in both lower extremities and simple balance exercises. Rest periods of 30-60s were given between exercise sets.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable

	(n=39) Intervention 2: Exercise - Unsupervised strength exercise. Exercises taught under the supervision and guidance of a physiotherapist in an exercise session, and these people were asked to perform the same exercise protocol at home at least three times a week. These people received structured telephone follow-up (once a week), to check compliance and answer any questions/adjust the intensity. People who wanted to do a supervised exercise program was offered this at the end of the study period.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by the Scientific Research Project Committee of Marmara University (SAG-A-100713-0300))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus UNSUPERVISED STRENGTH EXERCISE Protocol outcome 1: Pain at </=3 months - Actual outcome: VAS pain score (After activity) at 6 weeks; Group 1: mean 39.58 (SD 27.9); n=39, Group 2: mean 50.09 (SD 44.4); n=39; VAS 0-100 Top=High is poor outcome; Comments: Reports means and 95% confidence intervals. Reported supervised: 39.58 (30.8-48.3). Reported unsupervised: 50.09 (36.1-64). Baseline supervised: 67.61 (58.3-76.9). Baseline unsupervised: 62.61 (50.8-74.3). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, heigh, weight, BMI and baseline values of outcomes; Group 1 Number missing: 6, Reason: 5 discontinued as unable to contact, moving out of city, person decision. 1 did not attend the post-treatment assessments; Group 2 Number missing: 16, Reason: 12 discontinued intervention (unable to contact, person decision). 4 did not attend the post-treatment assessments.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Lee 2009[256]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=34)
Countries and setting	Conducted in South Korea; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic osteoarthritis with radiologic alterations in the knee joint of grade 2 or higher (Kellgren-Lawrence Scale) at least 6 months prior to study entry
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Symptomatic osteoarthritis with radiologic alterations in the knee joint of grade 2 or higher (Kellgren-Lawrence Scale) at least 6 months prior to study entry; no current participation in an exercise program; 50-80 years of age
Exclusion criteria	Had received a corticosteroid injection in the symptomatic knee within 6 months of study entry; had received medication fo rosteoarthritis within 6 months; had a history of knee surgery or a prior diagnosis of inflammatory arthritis
Recruitment/selection of patients	People were recruited from the Hawseong City Health Center, Republic of Korea
Age, gender and ethnicity	Age - Mean (SD): 69.1 (5.5). Gender (M:F): 3:41. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: median Kellgren-Lawrence grade 2-3 Duration of symptoms: At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=29) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Tai Chi Qigong performed for 1 hour, repeated twice a week for 8 weeks. Tai Chi Qigong consists of 18 movements, in which traditional warm-up exercises include weight shifting, arm swinging, visualisation techniques and gentle stretches of the neck, shoulder, spine, arms and legs. These exercises focus on releasing tension in the physical body, incorporating mindfulness and imagery into movement, increasing awareness of breathing and promoting overall relaxation of the body and mind.. Duration 8 weeks. Concurrent medication/care: No additional

	information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Tai Chi Qigong). (n=15) Intervention 2: No treatment. No treatment. People in this group were offered to complete the treatment programme after the study finished.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This work was supported by the Korea Science and Engineering Foundation grant funded by the Korean government (R11-2005-014))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical health at 8 weeks; Group 1: mean 17.1 (SD 14.9); n=29, Group 2: mean 5.6 (SD 12.9); n=15; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 38.7 (14.1). Baseline no treatment: 40.0 (11.9).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, Kellgren-Lawrence grade and baseline value of outcomes; Group 1 Number missing: 1, Reason: 1 person did not complete the study 0 withdrew due to professional activities not related to her clinical condition; Group 2 Number missing: 2, Reason: Unclear
- Actual outcome: SF-36 mental health at 8 weeks; Group 1: mean 19.2 (SD 15.9); n=29, Group 2: mean 9.1 (SD 10.3); n=15; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline exercise: 48.0 (17.1). Baseline no treatment: 43.4 (11.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, Kellgren-Lawrence grade and baseline value of outcomes; Group 1 Number missing: 1, Reason: 1 person did not complete the study 0 withdrew due to professional activities not related to her clinical condition; Group 2 Number missing: 2, Reason: Unclear

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean -9.4 (SD 14.4); n=29, Group 2: mean -2.7 (SD 10.8); n=15; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline exercise: 24.2 (14.7). Baseline no treatment: 23.5 (13.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, Kellgren-Lawrence grade and baseline value of outcomes; Group 1 Number missing: 1, Reason: 1 person did not complete the study 0 withdrew due to professional activities not related to her clinical condition; Group 2 Number missing: 2, Reason: Unclear

Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean -2.2 (SD 4.1); n=29, Group 2: mean -0.2 (SD 1.8); n=15; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.8 (4.2). Baseline no exercise: 6.1 (3.4). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, Kellgren-Lawrence grade and baseline value of outcomes; Group 1 Number missing: 1, Reason: 1 person did not complete the study 0 withdrew due to professional activities not related to her clinical condition; Group 2 Number missing: 2, Reason: Unclear	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Lim 2008[264]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=107)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Tibiofemoral joint osteoarthritis in at least 1 knee fulfilling the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	All people had tibiofemoral joint osteoarthritis in at least 1 knee fulfilling the American Colege of Rheumatology classification criteria. This included: medial knee pain, medial compartment osteophytes, and medial joint space narrowing greater than lateral joint space narrowing
Exclusion criteria	A history of lower limb joint replacement,; knee surgery within the previous 6 months; intraarticular steroid or hylan G-F20 injection within the previous 6 months; a systemic arthritic condition; more than 5 degrees of valgus malalignment on radiograph; intention to start or current participation in physiotherapy for knee osteoarthritis or a lower limb strengthening program; and/or presence of a severe medical condition that precluded safe participation in an exercise program
Recruitment/selection of patients	People were recruited form the community in Melbourne, Australia through advertisements in newspapers and at local community clubs
Age, gender and ethnicity	Age - Mean (SD): 64.6 (8.6). Gender (M:F): 48:59. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: median Kellgren Lawrence grade 3 Duration of symptoms (mean [SD]): 6.7 (6.5) years
Indirectness of population	No indirectness
Interventions	(n=53) Intervention 1: Exercise - Unsupervised strength exercise. Quadriceps muscle strength exercise including long arc knee extension, inner range knee extension, straight leg raise, isometric knee extension and isometric knee extension at a different ankle. This was supported by using ankle weights and a Thera-band. This was taught

	by physiotherapists to be performed on the study leg 5 days a week for 12 weeks. The physiotherapist visited 7 times to check progress.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=54) Intervention 2: No treatment. No treatment. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Study funded by industry (Supported in part by united Pacific Industries through a grant from the Physiotherapy Research Foundation, Australia)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean -5.7 (SD 4.1); n=53, Group 2: mean -2.9 (SD 2.2); n=54; WOMAC function 0-100 Top=High is poor outcome; Comments: Reports values subgrouped by knee malalignment. Combined for this analysis. Reported exercise more malaligned: -2.1 (2.1). Reported exercise more neutrally aligned: -9.2 (2.1). Reported control more malaligned: -2.0 (2.1). Reported control more neutrally aligned: -3.7 (2.0). Baseline exercise more malaligned: 31.4 (17.4). Baseline exercise more neutrally aligned: 33.1 (15.4). Baseline control more malaligned: 38.5 (14.5). Baseline control more neutrally aligned: 36.1 (15.7).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, body mass, BMI, physical activity level, symptom duration, bilateral symptoms, disease severity, varus malalignment, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 withdrew - 1 neck pain, 2 lack of time, 1 unrelated surgery; Group 2 Number missing: 6, Reason: 6 withdrew - 3 loss of motivation, 2 illness, 1 unrelated surgery

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean -8.9 (SD 4.8); n=53, Group 2: mean -1.9 (SD 2.9); n=54; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reports values subgrouped by knee malalignment. Combined for this analysis. Reported exercise more malaligned: -4.6 (2.5). Reported exercise more neutrally aligned: -13.0 (2.3). Reported control more malaligned: -3.1 (2.68). Reported control more neutrally aligned: -0.7 (2.5). Baseline exercise more malaligned: 33.1 (15.4). Baseline exercise more neutrally aligned: 35.7 (14.6). Baseline control more malaligned: 39.2 (14.0). Baseline control more neutrally aligned: 34.6 (16.2).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, body mass, BMI, physical activity level, symptom duration, bilateral symptoms, disease severity, varus malalignment, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 withdrew - 1 neck pain, 2 lack of time, 1 unrelated surgery; Group 2 Number missing: 6, Reason: 6 withdrew - 3 loss of motivation, 2 illness, 1

unrelated surgery	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Lim 2010[265]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=75)
Countries and setting	Conducted in South Korea; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis with Kellgren-Lawrence grade 2 or higher changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	50 years or older with a BMI more than 25kg/m² and abdominal circumferences of more than 90cm for men and 85cm for women. They showed Kellgren-Lawrence grade 2 or higher in radiologic assessments. All people were able to walk independently without walking devices.
Exclusion criteria	People were excluded if they were in progressive inflammatory or ankylosing states, or had coexisting central nervous system lesions or inadequate cardiac functions, people with infectious or skin diseases
Recruitment/selection of patients	People recruited from the patients who registered at the rehabilitation, arthritis and geriatric clinics at Seoul National University Bundang Hospital
Age, gender and ethnicity	Age - Mean (SD): 65.6 (7.7). Gender (M:F): 10:65. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2 or higher Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=26) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic exercise program performed in a water gym. 40 minute duration per session, 3 times per week for 8 weeks. Exercise intensity was maintained at a level of more than 65% of maximal heart rate by checking subject heart rates intermittently during exercise. Each training session consisted of main activities in the aquatic gym for 30 minutes. This was gradually progressed and included strength and endurance training.. Duration 8 weeks. Concurrent medication/care: No additional

	information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=25) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Land based exercise program with generalised conditioning and knee-specific exercises. Each exercise therapy was applied for 40 minutes at each session. The intensity began from 40% of the 1-repetition maximum for the beginning, and 60% of 1 repetition maximum for the advanced classes. The exercises consisted of joint mobilisation and strengthening exercises. Range of motion and stretching exercises of the hamstring, rectus femoris, tensor fascia latae, and calf muscles were included. Bicycling was also included for aerobic conditioning and fitness. A quadriceps isometric strengthening exercise was done, along with other strengthening exercises, such as leg presses and leg extensions.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strengthening, aerobic, stretching/range of motion). (n=24) Intervention 3: Exercise - Unsupervised strength exercise. Provided only the classes for home-based exercise, including the Q-sets exercise for strengthening of quadriceps muscles and a partial squatting along with behavioral correction of their daily activities and lifestyles. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by a grant from Health Promotion Fund 2005, ministry of Health and Welfare, Republic of Korea)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: SF-36 PCS at 8 weeks; Group 1: mean 38.8 (SD 7.7); n=24, Group 2: mean 40.4 (SD 7.9); n=22; SF-36 PCS 0-100 Top=High is good outcome; Comments: Baseline hydro: 34.4 (7.4). Baseline land: 33.6 (12.6). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass,	

body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).

- Actual outcome: SF-36 MCS at 8 weeks; Group 1: mean 54.8 (SD 8.8); n=24, Group 2: mean 52.9 (SD 8.3); n=22; SF-36 MCS 0-100 Top=High is good outcome; Comments: Baseline hydro: 47.3 (12.1). Baseline land: 50.6 (8.9).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).

Protocol outcome 2: Pain at </=3 months
- Actual outcome: BPI mean pain at 8 weeks; Group 1: mean 3.27 (SD 1.67); n=24, Group 2: mean 3.46 (SD 1.3); n=22; NRS 0-10 Top=High is poor outcome; Comments: Baseline hydro: 4.41 (1.43). Baseline land: 4.02 (1.45).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 PCS at 8 weeks; Group 1: mean 38.8 (SD 7.7); n=24, Group 2: mean 36.9 (SD 9.6); n=20; SF-36 PCS 0-100 Top=High is good outcome; Comments: Baseline hydro: 34.4 (7.4). Baseline home: 35.7 (9.4).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

- Actual outcome: SF-36 MCS at 8 weeks; Group 1: mean 54.8 (SD 8.8); n=24, Group 2: mean 48.4 (SD 14.3); n=20; SF-36 MCS 0-100 Top=High is good outcome; Comments: Baseline hydro: 47.3 (12.1). Baseline home: 47.4 (12.2).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

Protocol outcome 2: Pain at </=3 months
- Actual outcome: BPI mean pain at 8 weeks; Group 1: mean 3.27 (SD 1.67); n=22, Group 2: mean 4.55 (SD 1.88); n=20; NRS 0-10 Top=High is poor outcome; Comments: Baseline hydro: 4.41 (1.43). Baseline home: 4.12 (2.08).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Hydro: 2 drop outs (1 heart problems, 1 time constraint).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 PCS at 8 weeks; Group 1: mean 40.4 (SD 7.9); n=22, Group 2: mean 36.9 (SD 9.6); n=20; SF-36 PCS 0-100 Top=High is good outcome; Comments: Baseline land: 33.6 (12.6). Baseline home: 35.7 (9.4).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

- Actual outcome: SF-36 MCS at 8 weeks; Group 1: mean 52.9 (SD 8.3); n=22, Group 2: mean 48.4 (SD 14.3); n=20; SF-36 MCS 0-100 Top=High is good outcome; Comments: Baseline land: 50.6 (8.9). Baseline home: 47.4 (12.2).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

Protocol outcome 2: Pain at </=3 months
- Actual outcome: BPI mean pain at 8 weeks; Group 1: mean 3.46 (SD 1.3); n=22, Group 2: mean 4.55 (SD 1.88); n=20; NRS 0-10 Top=High is poor outcome; Comments: Baseline land: 4.02 (1.45). Baseline home: 4.12 (2.08).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI, body mass, body fat mass, body fat proportion, waist hip ratio, and baseline values of outcomes; Group 1 Number missing: 3, Reason: Land: 3 drop outs (3 pain and discomfort).; Group 2 Number missing: 4, Reason: Control: 4 dropouts (personal reasons).

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Lin 2004[268]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=106)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People treated for osteoarthritis of the knee/hip from their general practitioner, rheumatologist or orthopaedic surgeon
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged >60 years, with current symptoms of pain and joint stiffness in the knee and/or hip. People who answered yes to the question 'have you ever had pain in or around the knee on most days for at least a month?'. People being treated for osteoarthritis of the knee/hip from their general practitioner, rheumatologist or orthopaedic surgeon.
Exclusion criteria	People currently receiving hydrotherapy, physiotherapy or regularly participating in an exercise class (defined as more than once a week for 20 minutes or longer); or had a medical condition that precluded water based exercise (acute intermittent illness, unstable cardiac failure, myocardial infarction in the last three months, urinary infection or incontinence, open wounds or skin disease, advanced chronic obstructive pulmonary disease, paralysis, or dementia).
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 69.2 (6.00. Gender (M:F): 13:93. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Knee or hip osteoarthritis).
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 12.2 (11.1) years
Indirectness of population	No indirectness
Interventions	(n=66) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Water exercise programme consisted of 1 hour sessions twice a week over a period of 12 months. Accounting for holidays, the programme was run for a total of 46 weeks. The exercise facilitators were qualified exercise instructors who

	used a standard exercise protocol specifically designed for this groups based on a progressive five-phase plan. Each session lasted approximately 1 hour with a standard warm-up period and exercises included: joint range-of-motion, muscle strengthening, balance and co-ordination and cardiovascular fitness. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=40) Intervention 2: No treatment. Health education leaflets from the Arthritis and Rheumatism Council (UK) and the Arthritis Foundation (USA) were posted to each participant monthly. Quarterly telephone calls were made to maximize continued participation and to record any change(s), which could confound the results of the study (i.e. hospitalisation, starting a new exercise class, etc.). Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by the Department of Health under the NHS Health Technology Assessment R&D programme (Project number: 96/32/99) as a pilot study to a larger trial.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: WOMAC physical function at 12 months; Group 1: mean 30.16 (SD 14.03); n=56, Group 2: mean 34.96 (SD 9.87); n=38; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline exercise: 34.19 (9.88). Baseline control: 34.54 (10.32).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, BMI, affected joint sites, medication use and baseline values of outcomes; Group 1 Number missing: 10, Reason: No reasons given; Group 2 Number missing: 2, Reason: No reasons given

Protocol outcome 2: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean 8.62 (SD 4.34); n=56, Group 2: mean 9.32 (SD 2.84); n=38; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline exercise: 9.94 (3.14). Baseline control: 9.48 (3.76).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, BMI, affected joint sites,

medication use and baseline values of outcomes; Group 1 Number missing: 10, Reason: No reasons given; Group 2 Number missing: 2, Reason: No reasons given	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Lin 2009[266]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=108)
Countries and setting	Conducted in Taiwan; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis diagnosed by an orthopedic surgeon based on the clinical history, radiographic imaging and physical assessment
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	50 or more years of age, an osteoarthritis grade of 3 or lower as determined using the Kellgren and Lawrence plain radiograph classification, and a history of knee pain for longer than 6 months (chronic knee osteoarthritis)
Exclusion criteria	Received physical therapy treatment for their knee during the preceding 3 months; had other musculoskeletal conditions involving the knee joint (eg, tendon/ligament tears); had central or peripheral neurological disorders or hypertension.
Recruitment/selection of patients	People were recruited from the Department of Orthopedics, National Taiwan University Hospital
Age, gender and ethnicity	Age - Mean (SD): 62.5 (7.5). Gender (M:F): 33:75. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Radiographic grade 2-3, median grade 3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=36) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). A previously designed computer game foot-stepping exercise that predominantly involves knee movement in a sitting position. People were asked to repetitively step on target pedals in multiple directions. People were also required to perform fast and accurate range-of-motion exercises involving the knee.An approximately 150- to 250-N force was applied to the foot of the participating subject. Training was performed for 20 minutes for each lower extremity, with a 10-minute break between sides to prevent fatigue. All people underwent 3 training sessions per

	week for 8 weeks.. Duration 8 weeks. Concurrent medication/care: All people were asked to cease any exercise activity outside of the exercise training.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Proprioception (n=36) Intervention 2: Exercise - Supervised strength exercise. People sat comfortably in a chair with their back against a back support and their knees at 90 degrees of flexion. Both hands were used to grasp the sides of the seat. A pad attached to the dynamometer cable was placed on the distal, anterior portion of the leg. During training people were asked to fully extend their knee using a concentric quadriceps action, then to lower the leg using an eccentric quadriceps action. The baseline resistance was set at 50% of 1-RM, with a progressive increment of 5% of the original 1-RM every 2 weeks, as long as the increased resistance did not elicit knee pain. All people trained 3 sessions weekly for 8 weeks. Each session consisted of 4 sets, with 6 repetitions per set. There was a 1-minute rest between sets. Both lower extremities were trained, with a 5-minute interval between the training of each side.. Duration 8 weeks. Concurrent medication/care: All people were asked to cease any exercise activity outside of the exercise training.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=36) Intervention 3: No treatment. No exercise. Duration 8 weeks. Concurrent medication/care: All people were asked to cease any exercise activity outside of the exercise training.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Financial support from the National Science Council Grant in Taiwan (NSC 91-2218-E-002-033))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC function at 8 weeks; Group 1: mean 14.6 (SD 9.6); n=36, Group 2: mean 10.1 (SD 8.3); n=36; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline proprioception: 23.1 (9.4). Baseline strength: 27.3 (9.5). Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,	

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in WOMAC physical function subscale; Group 1 Number missing: 0, Reason: Proprioception: 0; Group 2 Number missing: 2, Reason: Strength: 2 knee pain

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 4.3 (SD 2.3); n=36, Group 2: mean 4.2 (SD 3); n=36; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline proprioception: 8.0 (3.7). Baseline strength: 8.8 (3.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in WOMAC physical function subscale; Group 1 Number missing: 0, Reason: Proprioception: 0; Group 2 Number missing: 2, Reason: Strength: 2 knee pain

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 14.6 (SD 9.6); n=36, Group 2: mean 24.9 (SD 11.8); n=36; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline proprioception: 23.1 (9.4). Baseline control: 24.8 (10.7).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, mass, bilateral involvement, radiographic osteoarthritis grade, and baseline values of outcomes; Group 1 Number missing: 0, Reason: Proprioception: 0; Group 2 Number missing: 3, Reason: Control: 3 personal reasons

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 4.3 (SD 2.3); n=36, Group 2: mean 7.3 (SD 3.4); n=36; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline proprioception: 8.0 (3.7). Baseline control: 8.5 (4.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, mass, bilateral involvement, radiographic osteoarthritis grade, and baseline values of outcomes; Group 1 Number missing: 0, Reason: Proprioception: 0; Group 2 Number missing: 3, Reason: Control: 3 personal reasons

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 10.1 (SD 8.3); n=36, Group 2: mean 24.9 (SD 11.8); n=36; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline strength: 27.3 (9.5). Baseline control: 24.8 (10.7).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, mass, bilateral involvement, radiographic osteoarthritis grade, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Strength: 2 knee pain; Group 2 Number missing: 3, Reason: Control: 3 personal reasons

Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 4.2 (SD 3); n=36, Group 2: mean 7.3 (SD 3.4); n=36; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline strength: 8.8 (3.6). Baseline control: 8.5 (4.6). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, mass, bilateral involvement, radiographic osteoarthritis grade, and baseline values of outcomes; Group 1 Number missing: 2, Reason: Strength: 2 knee pain; Group 2 Number missing: 3, Reason: Control: 3 personal reasons	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Mccaffrey 2019[287]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=18)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Reported pain associated with lower extremity osteoarthritis (hip, knee or other lower extremities) verified by a nurse practitioner
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 62 years or older; living independently; reported pain associated with lower extremity osteoarthritis verified by a nurse practitioner; ability to ambulate independently with minimal assistance (eg, cane or walker); chronic pain at least 15 days of the month for 3 months or longer; self-reported inability to participate in regular standing yoga or standing exercise (eg, aerobics) due to pain, physical disability, fear of falling, or balance problems; not currently participating in chair yoga or any other exercise program; ability to read and understand English.
Exclusion criteria	No additional information.

Recruitment/selection of patients	No additional information.
Age, gender and ethnicity	Age - Mean (SD): 78.5 (2.4). Gender (M:F): 2:10. Ethnicity: Not stated/unclear
Further population details	1. Age: Over 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Knee, hip and other lower extremity).
Extra comments	Severity: Not stated/unclear Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=9) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Chair yoga program based on traditional Hatha yoga postures, practiced in a chair or standing using a chair as support. The chair yoga intervention consisted of twice-weekly 50-minute sessions for 8 weeks, a total of 16 sessions, led by a certified yoga instructor. The yoga instructor held certification from the National Yoga Alliance and had been trained in the chair yoga program. The yoga intervention had 3 components - breathing (10 minutes), physical postures (25 minutes), and relaxation (10 minutes) - all completed while sitting in a chair. The yoga program consisted of a 50-minute yoga session held twice weekly for 8 weeks. Each of the 16 yoga sessions followed the same format.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not stated / Unclear 2. Group or individual : Individual session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Yoga). (n=9) Intervention 2: Exercise - Supervised strength exercise. Chair exercise for older adults. An exercise program adapted from the standing Go4Life program designed for older adults to increase muscle strength, range of motion and activities of daily living. The program uses progressive resistive exercises incorporating body weight and/or external resistance using cuff weights, resistance bands and balls. Each participant sites in a chair for seated exercises but may stand with the support of the chair. The program is suitable for beginners. Participants in this intervention group completed the 50-minute program twice weekly for 8 weeks, for a total of 16 sessions.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported by Mercer University.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean -17.4 (SD 14.4); n=9, Group 2: mean -14.9 (SD 13.6); n=9; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline values unclear (provides pre values, but they appear to be change scores and so do not appear

appropriate for understanding the true baseline values). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, regular exercise, pain location and medications used.; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean -4.4 (SD 2.1); n=9, Group 2: mean -4.4 (SD 2.4); n=9; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline values unclear (provides pre values, but they appear to be change scores and so do not appear appropriate for understanding the true baseline values). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, regular exercise, pain location and medications used.; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Serious adverse events at </=3 months - Actual outcome: Adverse events at 8 weeks; Group 1: 0/9, Group 2: 0/9 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, regular exercise, pain location and medications used.; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Mcilroy 2017[289]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=14)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 week
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Adults with persistent knee pain of at >3 months duration
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults at least 50 years old with self-reported persistent knee pain for >3 months; knee pan over the past 7 days of >3/10 on a Numerical Rating Scale and able and willing to provide informed consent
Exclusion criteria	Self-reported early morning stiffness for ≥30 minutes; contraindications to aquatic therapy; upcoming knee surgery in the next 3 months; physiotherapy, aquatic therapy or surgery for their persistent knee pain in the previous 6 months; the presence of another condition primarily limiting their mobility.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 63.3 (7.8). Gender (M:F): 0:14. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: >3 months
Indirectness of population	No indirectness
Interventions	(n=7) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). People who were randomised to receive aquatic therapy completed 6x30 minutes weekly group sessions delivered by one of two Senior Physiotherapists who had undertaken postgraduate aquatic therapy training. All participants completed a circuit of exercises aimed to increase function. This included strengthening exercises and cycling. People continued to receive usual medical care as directed by their referring Physician.. Duration 6 weeks. Concurrent medication/care: All participants attended one 30 minute individual, self-management education session

	with a physiotherapist. This comprised: information on the causes of persistent knee pain, physical activity/aerobic exercise and knee exercise (e.g. quadriceps strengthening exercises); footwear advice and the use of shock absorbing insoles; activity pacing; pain (e.g. thermotherapy) and weight management. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=7) Intervention 2: No treatment. Usual care (available to both groups) - medication and adjunctive therapies as directed by their referring Physician. Duration 6 weeks. Concurrent medication/care: All participants attended one 30 minute individual, self-management education session with a physiotherapist. This comprised: information on the causes of persistent knee pain, physical activity/aerobic exercise and knee exercise (e.g. quadriceps strengthening exercises); footwear advice and the use of shock absorbing insoles; activity pacing; pain (e.g. thermotherapy) and weight management. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Study funded by industry (Funding for an MSc in Musculoskeletal Physiotherapy was provided by the University College London Hospital's NHS trust and the Chartered Society of Physiotherapy Charitable Trust)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-12 physical subscale at 6 weeks; Group 1: mean 4.3 (SD 12.8); n=7, Group 2: mean 0.01 (SD 4.1); n=6; SF-12 physical 0-100 Top=High is good outcome; Comments: Baseline exercise: 33.4 (8.7). Baseline no treatment: 34.6 (10.5).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC function subscale was different at baseline; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 withdrew due to group allocation
- Actual outcome: SF-12 mental subscale at 6 weeks; Group 1: mean 8.2 (SD 9.2); n=7, Group 2: mean 1.2 (SD 5.2); n=6; SF-12 mental subscale 0-100 Top=High is good outcome; Comments: Baseline exercise: 42.9 (16.8). Baseline no treatment: 43.4 (13.1).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC function subscale was different at baseline; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 withdrew due to group allocation

Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC function subscale at 6 weeks; Group 1: mean -10.7 (SD 8.9); n=7, Group 2: mean 2 (SD 9.6); n=6; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline exercise: 58.1 (14.7). Baseline no treatment: 44.1 (16.7). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC function subscale was different at baseline; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 withdrew due to group allocation Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain subscale at 6 weeks; Group 1: mean -5.1 (SD 4.2); n=7, Group 2: mean 0.3 (SD 1.9); n=6; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline exercise: 16.7 (4.3). Baseline no treatment: 14.3 (4.2). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC function subscale was different at baseline; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 withdrew due to group allocation Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Serious adverse events at 6 weeks; Group 1: 0/7, Group 2: 0/7 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC function subscale was different at baseline; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Munukka 2016[310]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=87)
Countries and setting	Conducted in Finland; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 16 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Mild knee osteoarthritis demonstrated through radiography grade 1-2 changes according to the Kellgren-Lawrence classification experiencing knee pain on most days
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Postmenopausal women aged 60-68 years with mild knee osteoarthritis, experiencing knee pain on most days, participating in intesnive exercise no more than twice a week, radiographic changes, no previous cancer or chemotherapy, no medical contraindications or other limitations to full participation in an intensive aquatic training program and complete transverse relaxation time (T2) data.
Exclusion criteria	T-score <-2.5 (indicating osteoporosis) measured from the femoral neck using dual-energy X-ray absorptiometry; resting knee pain visual analogue scale >50/100; surgery of the knee due to trauma or knee instability; meniscectomy within the last 12 months; inflammatory joint disease; intra-articular steroid injections in the knee during the previous 12 months; contraindications to MRI and allergies to contrast agents or renal insufficiency; due to confounding factors related to obesity, a body mass index >34kg/m² was an exclusion criteria.
Recruitment/selection of patients	A multistage recruitment process was implemented. Initially, postmenopausal women from the Jyväskylä region in Central Finland were voluntarily recruited through advertisements in local newspapers.
Age, gender and ethnicity	Age - Mean (SD): 64 (2). Gender (M:F): 0:87. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-2 Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=43) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). 1 hour of supervised lower limb aquatic resistance training three times a week for 16 weeks, for a total of 48 training sessions. Resistance of exercises was progressed with three different levels: barefoot, small fins and large resistance boots and the training leg performed all the movements without contact with the pool walls or bottom i.e. non-weight bearing. The intervention was completed in groups of 6-8 subjects with 2 instructors. Intensity was set at @as hard and as fast as possible@ to ensure maximal muscle contraction. Training intensity was adjusted during the process.. Duration 16 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=44) Intervention 2: No treatment. Usual care only. Asked to continue their usual leisure time activities. They were offered the possibility of participating in two sessions consisting of 1 hours of light stretching and relaxation during the period.. Duration 16 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Financial supports were: Academy of Finland (ref: 253198), Social Insurance Institution of Finland (KELA) (ref: 24/26/2011), Finnish Cultural Foundation and Yrjö Jahnsson Foundation.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: KOOS QOL at 16 weeks; Group 1: mean 7 (SD 13); n=42, Group 2: mean 3 (SD 15); n=42; KOOS QOL 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported training: 7 (3 to 11). Reported control: 3 (-1 to 8). Baseline training: 65 (17). Baseline control: 71 (20).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body mass, BMi, time from menopause, pain killers for knee pain, glucosamine use, radiographic severity and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 withdrew; Group 2 Number missing: 2, Reason: 1 withdrew, 1 corrupted T2 data

Protocol outcome 2: Physical function at > 3 months - Actual outcome: KOOS ADL at 16 weeks; Group 1: mean 4 (SD 10); n=42, Group 2: mean 0 (SD 8); n=42; KOOS ADL 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported training: 4 (1 to 7). Reported control: 0 (-2 to 3). Baseline training: 84 (10). Baseline control: 85 (11). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body mass, BMi, time from menopause, pain killers for knee pain, glucosamine use, radiographic severity and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 withdrew; Group 2 Number missing: 2, Reason: 1 withdrew, 1 corrupted T2 data Protocol outcome 3: Pain at > 3 months - Actual outcome: KOOS pain at 16 weeks; Group 1: mean 4 (SD 10); n=42, Group 2: mean 1 (SD 10); n=42; KOOS pain 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported training: 4 (1 to 7). Reported control: 1 (-2 to 4). Baseline training: 80 (10). Baseline control: 82 (12). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body mass, BMi, time from menopause, pain killers for knee pain, glucosamine use, radiographic severity and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 withdrew; Group 2 Number missing: 2, Reason: 1 withdrew, 1 corrupted T2 data Protocol outcome 4: Serious adverse events at > 3 months - Actual outcome: Adverse events at 16 weeks; Group 1: 2/43, Group 2: 1/44; Comments: Two medical consultations (bilateral knee pain and dyspnoea) as a result of the aquatic training. One subject in the control group required a medical consultation for knee pain after the baseline physical performance measures. Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body mass, BMi, time from menopause, pain killers for knee pain, glucosamine use, radiographic severity and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 withdrew; Group 2 Number missing: 2, Reason: 1 withdrew, 1 corrupted T2 data	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	Nahayatbin 2018[314]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=48)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 months (1 month with treatment, 1 month post treatment
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis with grade 2-3 changes based on the Kellgren Lawrence classification
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Grades 2 to 3 of knee osteoarthritis based on the Kellgren Lawrence classification, and at least grade 3 of muscle strength for lower limb muscles based on the Oxford scale.
Exclusion criteria	Malignancy; infection; hypermobility; history of knee injury; other musculoskeletal disorders; corticosteroid injections and use of non-steroid anti-inflammatory drugs during the last month prior to the study; surgery of the lower limb; under physiotherapy or having attended any fitness classes during the last six months prior to the study
Recruitment/selection of patients	People admitted to a private clinic in district thirteen in Tehran, Iran
Age, gender and ethnicity	Age - Mean (SD): 55.79 (5.97). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=16) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Tai Chi - five minutes of Tai Chi warm-up, ten minutes of exercises according to form six of Yang style in Tai Chi as basic exercises, and five minutes of specific tai chi cool-down exercises. Duration 1 month. Concurrent medication/care: All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Tai Chi).

	(n=16) Intervention 2: Exercise - Supervised strength exercise. Closed chain kinetic exercises - 20 minutes in each second, including five minutes of static stretching warm up, ten minutes of exercises, and five minutes of cool-down exercises. The exercises consisted of standing terminal extension with ten seconds hold and ten seconds rest, mini qsquat with an angle of fifteen degrees with ten seconds hold and ten seconds rest, front and side step-up exercises each for ten times and lung exercise for ten times with ten seconds hold and ten seconds rest. The exercises were carried out with two-minute rest intervals.. Duration 1 month. Concurrent medication/care: All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=16) Intervention 3: No treatment. No exercise treatment. Duration 1 month. Concurrent medication/care: All people had routine physiotherapy - including: 15 minutes of infrared and 5 minutes of pulsed ultrasound.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS Quality of Life at 8 weeks; Group 1: mean 57.13 (SD 16.41); n=16, Group 2: mean 57.31 (SD 19.39); n=16; KOOS Quality of Life 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS Activity of Daily Living at 8 weeks; Group 1: mean 74.69 (SD 12.54); n=16, Group 2: mean 59 (SD 10.25); n=16; KOOS Activities of Daily Living 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome:

No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS Pain at 8 weeks; Group 1: mean 70.13 (SD 11.8); n=16, Group 2: mean 58.44 (SD 9.51); n=16; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS Quality of Life at 8 weeks; Group 1: mean 57.13 (SD 16.41); n=16, Group 2: mean 40 (SD 15.24); n=16; KOOS Quality of Life 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS Activity of Daily Living at 8 weeks; Group 1: mean 74.69 (SD 12.54); n=16, Group 2: mean 61.31 (SD 10.39); n=16; KOOS activities of daily living 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS Pain at 8 weeks; Group 1: mean 70.13 (SD 11.8); n=16, Group 2: mean 50.31 (SD 10.77); n=16; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: KOOS Quality of Life at 8 weeks; Group 1: mean 57.31 (SD 19.39); n=16, Group 2: mean 40 (SD 15.24); n=16; KOOS Quality of Life 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at </=3 months - Actual outcome: KOOS Activity of Daily Living at 8 weeks; Group 1: mean 59 (SD 10.25); n=16, Group 2: mean 61.31 (SD 10.39); n=16; KOOS Activities of Daily Living 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Pain at </=3 months - Actual outcome: KOOS Pain at 8 weeks; Group 1: mean 58.44 (SD 9.51); n=16, Group 2: mean 50.31 (SD 10.77); n=16; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline values not reported (first reports the amounts after 1 session). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - References a nonrandom sampling method (convenience sampling); Indirectness of outcome: No indirectness ; Baseline details: Reports age, height, weight and BMI. Selectively reports the first value for outcomes as the values after the first exercise session; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Nambi 2020[315]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Saudi Arabia; Setting: Outpatient follow up

Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks (post-intervention), 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Chronic osteoarthritis after ACL injury (secondary osteoarthritis)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	University football players; age group 18-25 years; male players; chronic (at least 3 months) osteoarthritis after anterior cruciate ligament injury; 4-8 pain intensity on the visual analogue scale (VAS).
Exclusion criteria	Severe musculoskeletal, neural, somatic or psychiatric conditions; waiting for surgery; alcohol or drug abuse; involvement in other weight and balance training programmes; people with other soft-tissue injuries, fracture of the lower limbs and pelvic bone, or deformities.
Recruitment/selection of patients	University football players with post-traumatic osteoarthritis after anterior cruciate ligament injury.
Age, gender and ethnicity	Age - Mean (SD): 22.5 (1.5). Gender (M:F): 60:0. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of injury (SD): 5.4 (0.4) months
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Exercise - Supervised strength exercise. People were instructed to perform a 5-minute warm-up, followed by slow stretching of the hamstring and quadriceps muscles. They were instructed to sit in an isokinetic dynamometer with their hips flexed at 90 degrees. Velcro fixation straps were tied around the chest, hip and the distal thigh of the training limb to prevent unnecessary movements. The training knee was kept at a 90 degrees flexed position, and the dynamometer axis was aligned with the centre of the lateral femoral condyle. The lever arm was customized according to the subject's leg length, and resistance was applied anterior to the ankle joint. The knee was tested from 0 degrees to 120 degrees of flexion, where 0 degrees was considered full extension. Subjects were familiarised with the exercise by showing them video clips of a model, then allowing them to practice attempts. Once they had mastered the exercise they were instructed to perform it at an angular speed of 60, 90 and 120 degrees/s with 15 repetitions in 3 sets. Rest periods of 30 s between each set, and 60 s between each speed, were given. Training was performed on 5 days a week for 4 weeks. Subjects were monitored and instructed by a supervisor throughout the training.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not stated / Unclear

	(n=20) Intervention 2: Exercise - Other supervised exercise (including flexibility, proprioception). Sensory motor training. The training was given in 3 stages; static, dynamic and functional. All exercises were performed 5 times in 1 set, for 3 sets, with a sufficient with 5 minutes rest period between sets. The exercise protocol was designed so that the level of difficulty increased. The subjects were not progressed to the next level of difficulty until they had completed the previous level. In the static phase subjects were instructed to stand in an erect position for 30 s on a firm surface and 30 s on a soft surface. They were then instructed to stand on one leg (the affected limb) with eyes closed for 10 s on a firm surface and 10 s on a soft surface, followed by a half knee-bend position for 10 s. In the dynamic phase they were instructed to perform a forward stepping thrust for 30 s and T-band kick exercise for 30 s. The functional phase began with toe skipping for 20 m (straight, inward and outward rotation) and heel skipping for 20 m (straight, inward and outward rotation). Subjects were then instructed to perform regular squats (10 times) bilaterally and unilaterally with the support of a wall and away from the wall. Once trained in the above exercises they were instructed to perform the balance exercises on a wobble board.. Duration 4 weeks. Concurrent medication/care: No additional information.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Neuromodulatory (n=20) Intervention 3: Exercise - Unsupervised strength exercise. People were given instructions about the type of exercise and the procedure for stretching and strengthening the specific muscles to be performed at home. Initially the exercises were demonstrated by the therapist and clarifications were given. These home-based exercises were printed in a hand manual with easy-to-follow images and language. The first part of the manual contained the do's and don'ts for the study period. The next part of the manual contained different stretching and strengthening exercises for quadriceps, hamstrings, glutei and calf muscles. Subjects performed these exercises 10-15 repetitions/day, 5 days a week for 4 weeks. Stretching was focused on each muscle group for 3 repetitions of 15 s per muscle group.. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by the Deanship of scientific research, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analogue scale at 8 weeks; Group 1: mean 1.8 (SD 0.4); n=19, Group 2: mean 3.8 (SD 0.4); n=18; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline supervised strength: 7.5 (0.4). Baseline unsupervised strength: 7.5 (0.5).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 1, Reason: Reasons not provided; Group 2 Number missing: 2, Reason:

Reasons not provided

Protocol outcome 2: Pain at > 3 months
- Actual outcome: Visual analogue scale at 6 months; Group 1: mean 0.8 (SD 0.3); n=18, Group 2: mean 3.1 (SD 0.2); n=18; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline supervised strength: 7.5 (0.4). Baseline unsupervised strength: 7.5 (0.5).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 2, Reason: Reasons not provided; Group 2 Number missing: 2, Reason: Reasons not provided

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analogue scale at 8 weeks; Group 1: mean 3.6 (SD 0.3); n=18, Group 2: mean 1.8 (SD 0.4); n=19; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline other supervised: 7.7 (0.6). Baseline supervised strength: 7.5 (0.4).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 2, Reason: Reasons not provided; Group 2 Number missing: 1, Reason: Reasons not provided

Protocol outcome 2: Pain at > 3 months
- Actual outcome: Visual analogue scale at 6 months; Group 1: mean 2.9 (SD 0.2); n=18, Group 2: mean 0.8 (SD 0.3); n=18; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline other supervised: 7.7 (0.6). Baseline supervised strength: 7.5 (0.4).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 2, Reason: Reasons not provided; Group 2 Number missing: 2, Reason: Reasons not provided

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus UNSUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analogue scale at 8 weeks; Group 1: mean 3.6 (SD 0.3); n=18, Group 2: mean 3.8 (SD 0.4); n=18; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline other supervised: 7.7 (0.6). Baseline unsupervised strength: 7.5 (0.5).
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 2, Reason: Reasons not provided; Group 2 Number missing: 2, Reason:

Reasons not provided Protocol outcome 2: Pain at > 3 months - Actual outcome: Visual analogue scale at 6 months; Group 1: mean 2.9 (SD 0.2); n=18, Group 2: mean 3.1 (SD 0.2); n=18; Visual analogue scale 0-10 Top=High is poor outcome; Comments: Baseline other supervised: 7.7 (0.6). Baseline unsupervised strength: 7.5 (0.5). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight BMI, VO2, HR, years of playing, duration of injury and baseline values of outcomes; Group 1 Number missing: 2, Reason: Reasons not provided; Group 2 Number missing: 2, Reason: Reasons not provided	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Nejati 2015[320]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=56)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months (states 12 months follow up in total, unclear but likely 3 months of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology criteria with radiographic Kellgren Lawrence grade 2-4 changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Pain in the knee for more than 3 months in most days of the week; having grade 2-4 radiographic knee osteoarthritis according to the criteria of Kellgren-Lawrence; having a BMI in the 18-30kg/m² range.
Exclusion criteria	Having limitations for performing strength exercise, icnluding: uncontrolled hypertension, uncontrolled ventricular arrhythmias, uncontrolled heart failure, and severe valvular problems.
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 61.3 (9.2). Gender (M:F): 20:30. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=28) Intervention 1: Exercise - Supervised strength exercise. Strengthening and stretching exercises for muscles around the knee (hamstrings, quadriceps and calf muscles). People were asked to perform the stretching exercises daily and keep doing each exercise for a maximum of 15 seconds in stretching form and repeat them 4 times. Strengthening exercise were performed daily and each time every exercise was repeated 10 times in three sets. There was 1-3 minute rest between sets. The weight of cuff weights tied to the person's ankle were selected according to the tolerance of

	people and their basic status. The weight of cuff was added 250 grams each 2 weeks until it met 2kg.. Duration 3 months. Concurrent medication/care: In both groups people received acupuncture during 10 sessions, twice per week, physical modalities during 1- sessions, three times a week (including TENS, ultrasound and infrared) and could receive diclofenac 100mg once daily for pain. All people were recommended to use 1500mg glucosamine and 800mg chondroitin. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=28) Intervention 2: No treatment. No exercise treatment. Duration 3 months. Concurrent medication/care: In both groups people received acupuncture during 10 sessions, twice per week, physical modalities during 1- sessions, three times a week (including TENS, ultrasound and infrared) and could receive diclofenac 100mg once daily for pain. All people were recommended to use 1500mg glucosamine and 800mg chondroitin. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS QOL at 3 months; Group 1: mean 39.4 (SD 3.26); n=28, Group 2: mean 35.74 (SD 3.26); n=28; KOOS QOL 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 39.40. Reported control: 35.74. P-value = 0.000 (using 0.0001 for calculations). Reports baseline means only. Baseline exercise: 28.52. Baseline control: 23.06.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 8, Reason: 8 lost by the third month. 12 lost by the twelfth month.; Group 2 Number missing: 4, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: KOOS QOL at 12 months; Group 1: mean 30.26 (SD 18.7); n=28, Group 2: mean 38.21 (SD 18.7); n=28; KOOS QOL 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 30.26. Reported control: 38.21. P-value = 0.118. Reports baseline means only. Baseline exercise: 28.52. Baseline control: 23.06.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 12, Reason: 8 lost by the third

month. 12 lost by the twelfth month.; Group 2 Number missing: 11, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: KOOS ADL at 3 months; Group 1: mean 64.99 (SD 3.37); n=28, Group 2: mean 50.81 (SD 3.37); n=28; KOOS ADL 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 64.99. Reported control: 50.81. P-value = <0.0001. Reports baseline means only. Baseline exercise: 49.96. Baseline control: 41.24.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 8, Reason: 8 lost by the third month. 12 lost by the twelfth month.; Group 2 Number missing: 4, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: KOOS ADL at 12 months; Group 1: mean 46.98 (SD 20.3); n=28, Group 2: mean 58.88 (SD 20.3); n=28; KOOS ADL 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 46.98. Reported control: 58.88. P-value = 0.033. Reports baseline means only. Baseline exercise: 49.96. Baseline control: 41.24.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 12, Reason: 8 lost by the third month. 12 lost by the twelfth month.; Group 2 Number missing: 11, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcome 5: Pain at </=3 months
- Actual outcome: KOOS pain at 3 months; Group 1: mean 63.39 (SD 19.3); n=28, Group 2: mean 46.65 (SD 19.3); n=28; KOOS pain 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 63.39. Reported control: 46.65. P-value = 0.002. Reports baseline means only. Baseline exercise: 46.96. Baseline control: 36.92.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 8, Reason: 8 lost by the third month. 12 lost by the twelfth month.; Group 2 Number missing: 4, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcome 6: Pain at > 3 months
- Actual outcome: KOOS pain at 12 months; Group 1: mean 48.07 (SD 1.73); n=28, Group 2: mean 49.03 (SD 1.73); n=28; KOOS pain 0-100 Top=High is good outcome; Comments: Reports means and p-values. Reported exercise: 48.07. Reported control: 49.03. P-value = 0.043. Reports baseline means only. Baseline exercise: 46.96. Baseline control: 36.92.
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes. Baseline values of outcomes were different between groups (but no comparison of SD).; Group 1 Number missing: 12, Reason: 8 lost by the third month. 12 lost by the twelfth month.; Group 2 Number missing: 11, Reason: 4 lost by the third month. 11 lost by the twelfth month.

Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Nery 2021[322]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Hand osteoarthritis under the American College of Rheumatology classification criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People older than 50 years of age; with hand osteoarthritis for at least one year under the American College of Rheumatology classification criteria; a history of pain in the interphalangeal joints between three and eight in the Numerical pain Scale and with a stable dosage of disease-modifying antirheumatic drugs for at least three months before the study.
Exclusion criteria	People with a history of a systemic illness associated with the upper limbs; inflammatory rheumatic diseases; prior hand surgery; deformities which prevented them performing the exercises; rehabilitation and joint injection in their upper limbs in the last three months.
Recruitment/selection of patients	People who had sought care for hand osteoarthritis at the University Rheumatology Division and were recruited from primary health care by a physiotherapist, an expert in rheumatology. A rheumatology-expert doctor conducted the selection of these individuals. Carried out at the Rheumatology Rehabilitation Section of the Federal University of Sao Paulo (UNIFESP) in the city of Sao Paulo, Brazil between October 2013 and February 2015.
Age, gender and ethnicity	Age - Mean (SD): 66.8 (9.1). Gender (M:F): 1:59. Ethnicity: White = 37. Non-white = 23.

Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hand osteoarthritis
Extra comments	Severity: Kellgren Lawrence grades I-IV, median grade III. Duration of symptoms (SD): 7.1 (5.1) years
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise - Supervised strength exercise. The exercise group besides the education session were engaged in a program of progressive resistance exercises for hands. This was done in groups of up to five patients supervised by a physiotherapist, experienced in rheumatology. The exercises proposed in the program covered the following muscle groups: flexor digitorum superficialis and flexor digitorum profundus, from the second to the fifth finger, extensor digitorum communis and dorsal, palmar and lumbrical interosseous. The exercises selection was based on the main muscle groups related to hand function. The training was followed according to the recommendations of the American College of Sports and Medicine. Patients took part in two sessions per week over twelve weeks. Each session lasted at least 35 minutes and three sets of 10 repetitions for each muscle group were performed. The exercises were done by alternating the hands for each set. Warming and stretching exercises were not done before or after hand exercise. To perform the exercises, two devices were used: the Digi-Extend for the extensor digitorum communis and Power-Web for the other muscle groups. The resistance progression was performed according to the load available and to the manufacturer's instruction. These were identified by different colours to indicate types of rubber with different levels of resistance. The colours used were yellow, red, green and blue, indicating light, medium light, medium and strong resistance, respectively. The same colour scheme was used for both devices. The resistance was increased every three weeks, beginning with the yellow colour and ending with the blue colour.. Duration 12 weeks. Concurrent medication/care: Both groups had a single education session to receive information about the illness before the randomisation. This briefing included information about the disease and the impairment it caused, treatment and guidelines for joint protection and energy conservation. The participants did not receive any extra material. Both groups were instructed to continue medication without change and orthoses were not allowed during the study.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=30) Intervention 2: No treatment. No additional treatment.. Duration 12 weeks. Concurrent medication/care: Both groups had a single education session to receive information about the illness before the randomisation. This briefing included information about the disease and the impairment it caused, treatment and guidelines for joint protection and energy conservation. The participants did not receive any extra material. Both groups were instructed to continue medication without change and orthoses were not allowed during the study.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months - Actual outcome: AUSCAN function at 12 weeks; Group 1: mean 8.77 (SD 7.4); n=30, Group 2: mean 13.8 (SD 7.42); n=30; AUSCAN function 0-36 Top=High is poor outcome; Comments: Baseline exercise: 12.67 (7.99). Baseline control: 14.07 (7.05). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, disease duration, education, race, dominant hand, medication, intraarticular injection, Kelgren and Lawrence classification, erosive hand osteoarthritis, dominant or non-dominant hand; Group 1 Number missing: 3, Reason: 3 discontinued due to personal problems.; Group 2 Number missing: 1, Reason: 1 discontinued due to personal problems. Protocol outcome 2: Pain at </=3 months - Actual outcome: AUSCAN pain at 12 weeks; Group 1: mean 4.97 (SD 4.07); n=30, Group 2: mean 8.23 (SD 4.42); n=30; AUSCAN pain 0-20 Top=High is poor outcome; Comments: Baseline exercise: 7.90 (4.49). Baseline control: 8.47 (3.01). Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, disease duration, education, race, dominant hand, medication, intraarticular injection, Kelgren and Lawrence classification, erosive hand osteoarthritis, dominant or non-dominant hand; Group 1 Number missing: 3, Reason: 3 discontinued due to personal problems.; Group 2 Number missing: 1, Reason: 1 discontinued due to personal problems.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Ojoawo 2016[330]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=50)
Countries and setting	Conducted in Nigeria; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis of knee osteoarthritis (in people with symptomatic and radiologic evidence) with symptoms of pain, stiffness and functional difficulty of no less than 6 weeks duration
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Males and females with diagnosis of knee osteoarthritis with symptoms of pain, stiffness and functional difficulty of no less than 6 weeks duration
Exclusion criteria	People with osteoporosis; acute inflammation; those with a history of traumatic injury and surgery of the knee joint; people who had previously been on physiotherapy treatment
Recruitment/selection of patients	People referred with treatment at the Department of Physiotherapy, Obafemi Awolowo University Teaching Hospital Complex
Age, gender and ethnicity	Age - Mean (SD): 68.89 (10.28). Gender (M:F): Not stated. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: at least 6 weeks
Indirectness of population	No indirectness
Interventions	(n=25) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Proprioceptive exercise. One leg balance - each subject stood on the affected leg with relaxed upright posture and the other leg was flexed at the knee, hip and ankle joint off from the ground. This position was held for 1 minute followed by rest for 10-20 seconds and was repeated twice. They rested for 2-3 minutes after which the procedure was repeated twice for the unaffected leg. Blind advanced one leg balance: the same as the previous exercise, except the person was asked to completely close their eyes.. Duration 6 weeks. Concurrent medication/care: Infrared

	radiation therapy was applied with a methyl salicylate ointment for 20 minutes twice a week for 6 weeks. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Proprioception (n=25) Intervention 2: Exercise - Supervised strength exercise. Isometric quadriceps strengthening exercise completed in high sitting with ankles in dorsi-flexion and standard weights hung with a bag on the ankle joint. The amount of weight depended on the size of the weight the person could carry. The person was asked to sustain the knee joint in extension with the ankle in dorsi-flexion, the subjects were asked to release the leg after the count of ten (equivalent to 10 seconds). The person was then allowed to rest for 6 seconds and repeated the same procedure for 10 repetitions.. Duration 6 weeks. Concurrent medication/care: Infrared radiation therapy was applied with a methyl salicylate ointment for 20 minutes twice a week for 6 weeks. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 6 weeks; Group 1: mean 10.14 (SD 11.48); n=23, Group 2: mean 17.67 (SD 8.66); n=22; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline proprioception: 23.71 (10.37). Baseline strength: 23.67 (8.33).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 for undisclosed reasons; Group 2 Number missing: 3, Reason: 3 engagement at office

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 3.71 (SD 3.4); n=23, Group 2: mean 6.5 (SD 3.83); n=22; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline proprioception: 10.71 (3.04). Baseline strength: 9.00 (3.46).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 for undisclosed reasons; Group 2 Number missing: 3, Reason: 3 engagement at office

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Oliveira 2012[332]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed according to the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age between 50 and 75 years; osteoarthritis classified as grade 2 and over based on the Kellgren and Lawrence radiological classification; knee osteoarthritis diagnosed according to the American College of Rheumatology criteria
Exclusion criteria	Pacemaker use; unstable heart conditions; participation in another physical activity program; inability to pedal a stationary bike; inability to walk; previous knee or hip arthroplasty; diagnosis of fibromyalgia; epilepsy; presence of a tumour or cutaneous lesion that could interfere with the procedure
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.1 (8.5). Gender (M:F): 6:94. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=50) Intervention 1: Exercise - Supervised strength exercise. Strength exercise performed twice a week for a period of 8 weeks including: a warm up for 10 minutes with a stationary bike; stretching of the hamstring muscle with the aid of an elastic band (three sets of 30 seconds); and three sets of 15 repetitions of knee extension exercises, with 30-45 second intervals between the sets. The exercise was performed in a sitting position, with the hip and knees flexed at 90 degrees. The load used was defined based on the ten repetition maximum test rather than the one-repetition maximum test. Fifty to sixty percent of the estimated maximum load was used..

	Duration 8 weeks. Concurrent medication/care: Both groups received a manual with instructions to prevent knee overload during daily activities and instructions about the use of knee ice packs for pain with inflammation, and warm dressing for pain with no inflammatory signs. In addition, people in both groups were already prescribed medication.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=50) Intervention 2: No treatment. No exercise intervention. Encouragement to follow the instruction manual on weeks 2 and 6.. Duration 8 weeks. Concurrent medication/care: Both groups received a manual with instructions to prevent knee overload during daily activities and instructions about the use of knee ice packs for pain with inflammation, and warm dressing for pain with no inflammatory signs. In addition, people in both groups were already prescribed medication.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Study funded by Fundação de Apoio à Pesquisa do Estado deSão Paulo)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean -10.95 (SD 14.05); n=50, Group 2: mean -1.97 (SD 16.58); n=50; WOMAC function 0-68 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported exercise: -10.95 (-14.84, -7.05). Reported no treatment: -1.97 (-6.56, 2.63). Baseline exercise: 35.15 (11.88). Baseline no treatment: 33.40 (12.58).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, body side treated, BMI, KL grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 no treatment compliance, 6 intervention interrupted, 2 knee pain, 1 death in the family, 2 found a new job; Group 2 Number missing: 12, Reason: 12 losses (did not return for final assessment)

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean -3.87 (SD 4.15); n=50, Group 2: mean -1.05 (SD 4.65); n=50; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported exercise: -3.87 (-5.02, -2.72). Reported no treatment: -1.05 (-2.35, 0.23). Baseline exercise: 10.32 (3.54). Baseline no treatment: 8.90 (4.38).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, body side treated, BMI,

KL grade, and baseline values of outcomes; Group 1 Number missing: 7, Reason: 7 losses - 1 no treatment compliance, 6 intervention interrupted, 2 knee pain, 1 death in the family, 2 found a new job; Group 2 Number missing: 12, Reason: 12 losses (did not return for final assessment) Protocol outcome 3: Serious adverse events at </=3 months - Actual outcome: Pain and inflammation increase at 8 weeks; Group 1: 2/50, Group 2: 0/50 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, body side treated, BMI, KL grade, and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	O'reilly 1999[328]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=191)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Partially adequate method of assessment/diagnosis: People with knee pain who responded affirmatively to both parts of the following questions "Have you ever had pain in or around the knee on most days for at least a month? If so, have you experienced any pain during the last year" who were then further assessed (presumedly to confirm osteoarthritis)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with knee pain who responded affirmatively to both parts of the following questions "Have you ever had pain in or around the knee on most days for at least a month? If so, have you experienced any pain during the last year" who were then further assessed (presumedly to confirm osteoarthritis)
Exclusion criteria	Already performing quadriceps exercises; clinical inflammatory arthropathy; pain referred from back or hip; serious injury within 6 months; previous knee replacement; unable to complete study because of imminent move or hospitalisation; no pain on WOMAC pain score; medical condition preventing exercise
Recruitment/selection of patients	People were registered at two general practices in Nottingham who responded to a postal survey
Age, gender and ethnicity	Age - Mean (SD): 62.05 (9.87). Gender (M:F): 67:124. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: At least 1 year
Indirectness of population	No indirectness
Interventions	(n=113) Intervention 1: Exercise - Unsupervised strength exercise. Graded exercise program including: isometric quadriceps contraction in full extension held for five seconds; isotonic quadriceps contraction held in mid flexion for five seconds; isotonic

	hamstring contraction; isotonic quadriceps contraction with resistance band held for five seconds; dynamic stepping exercise. Exercises were started in the above order and increased to a maximum of 20 repetitions on each leg. Exercises were performed at home on a daily basis, having been taught by a nurse metrologist.. Duration 6 months. Concurrent medication/care: General verbal advice concerning knee pain and knee osteoarthritis with advice on the importance of losing weight or not becoming overweight, wearing training shoes/air filled soles and maintaining fitness by walking or swimming was given to all participants. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=78) Intervention 2: No treatment. Did not receive any specific interventions. Duration 6 months. Concurrent medication/care: General verbal advice concerning knee pain and knee osteoarthritis with advice on the importance of losing weight or not becoming overweight, wearing training shoes/air filled soles and maintaining fitness by walking or swimming was given to all participants. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The authors are grateful to the Arthritis and Rheumatism Council for Research, UK for providing financial support)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: SF-36 physical function at 6 months; Group 1: mean 2.68 (SD 16.57); n=113, Group 2: mean -1.63 (SD 16.2); n=78; SF-36 physical function 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 2.68 (-0.38, 5.73). Reported control: -1.63 (-5.23, 1.96). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 mental health at 6 months; Group 1: mean -0.21 (SD 13.86); n=113, Group 2: mean -2.91 (SD 16.69); n=78; SF-36 mental health 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: -0.21 (-2.77, 2.34). Reported control: -2.91 (-6.62, 0.79). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5,

Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 energy at 6 months; Group 1: mean 2.47 (SD 16.76); n=113, Group 2: mean 0.56 (SD 20.16); n=78; SF-36 energy subscale 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 2.47 (-0.62, 5.56). Reported control: 0.56 (-3.91, 5.04). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 bodily pain at 6 months; Group 1: mean 4.97 (SD 23.48); n=113, Group 2: mean 0.16 (SD 25.39); n=78; SF-36 bodily pain 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 4.97 (0.64, 9.30). Reported control: 0.16 (-5.47, 5.80). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 health perception at 6 months; Group 1: mean 1.93 (SD 14.54); n=113, Group 2: mean -0.7 (SD 14.44); n=78; SF-36 health perception 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 1.93 (-0.75, 4.61). Reported control: -0.70 (-3.91, 2.50). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 role limitation physical at 6 months; Group 1: mean 3.19 (SD 38.07); n=113, Group 2: mean -7.59 (SD 40.04); n=78; SF-36 role limitation physical 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 3.19 (-3.83, 10.21). Reported control: -7.59 (-16.47, 1.30). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 role limitation emotional at 6 months; Group 1: mean 1.85 (SD 46.15); n=113, Group 2: mean 0.48 (SD 62.34); n=78; SF-36 role limitation emotional 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 1.85 (-6.66, 10.36). Reported control: 0.48 (-13.35, 14.32). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: SF-36 social functioning at 6 months; Group 1: mean 1.89 (SD 25.79); n=113, Group 2: mean 1.9 (SD 41.12); n=78; SF-36 social functioning 0-100 Top=High is good outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: 1.89 (-2.87, 6.64).

Reported control: 1.90 (-7.22, 11.03). Does not report baseline values.
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: WOMAC physical function score at 6 months; Group 1: mean -3.55 (SD 9.74); n=113, Group 2: mean -0.01 (SD 7.82); n=78; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: -3.55 (-5.34, -1.75). Reported control: -0.01 (-1.75, 1.72). Baseline exercise: 20.38 (12.54). Baseline control: 19.51 (11.52).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved

Protocol outcome 3: Pain at > 3 months
- Actual outcome: WOMAC pain score at 6 months; Group 1: mean -1.45 (SD 3.2); n=113, Group 2: mean 0.42 (SD 3.02); n=78; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: -1.45 (-2.04, -0.86). Reported control: 0.42 (-1.09, 0.25). Baseline exercise: 6.45 (3.50). Baseline control: 6.75 (2.83).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved

Protocol outcome 4: Psychological distress at > 3 months
- Actual outcome: HADS anxiety score at 6 months; Group 1: mean -0.57 (SD 3.09); n=113, Group 2: mean 0.06 (SD 3.22); n=78; HADS anxiety score 0-21 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: -0.57 (-1.14, 0.00). Reported control: 0.06 (-0.66, 0.77). Baseline exercise: 7.06 (3.69). Baseline control: 6.82 (3.65).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5, Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved
- Actual outcome: HADS depression score at 6 months; Group 1: mean -0.57 (SD 2.09); n=113, Group 2: mean 0.11 (SD 2.16); n=78; HADS depression score 0-21 Top=High is poor outcome; Comments: Reports mean change and 95% confidence intervals. Reported exercise: -0.57 (-0.96, -0.19). Reported control: 0.11 (-0.37, 0.59). Baseline exercise: 4.58 (2.91). Baseline control: 4.79 (2.91).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, weight, and baseline values for biophysical parameters and WOMAC pain/function and Anxiety/Depression scores. Does not report baseline SF-36 scores.; Group 1 Number missing: 5,

Reason: 5 lost to follow up - 4 unwell, 1 refused; Group 2 Number missing: 6, Reason: 6 lost to follow up - 2 unwell, 3 refused, 1 moved	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Osteras 2014[337]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=130)
Countries and setting	Conducted in Norway; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (with follow up at 6 months)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Hand osteoarthritis meeting the American College of Rheumatology criteria for features of hand osteoarthritis or uni/bilateral osteoarthritis of the first carpometacarpal joint, and a Functional Index for Hand Osteoarthritis score of at least 5
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Meeting the American College of Rheumatology criteria for features of hand osteoarthritis or uni-/bilateral osteoarthritis in the first carpometacarpal joint and a Functional Index for Hand Osteoarthritis score of at least 5.
Exclusion criteria	Those with inflammatory rheumatic disease (e.g., rheumatoid arthritis, polymyalgia rheumatica); had received steroid injections in the past 2 months; had recently experienced severe trauma or recently underwent osteoarthritis surgery or other major surgery were excluded along with people with cognitive dysfunction or language problems
Recruitment/selection of patients	People recruited from two previous osteoarthritis cohorts: The Musculoskeletal pain in Ullensaker STudy and the Oslo Hand osteoarthritis cohort
Age, gender and ethnicity	Age - Mean (SD): 66 (8.6). Gender (M:F): 13:117. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Low morbidity score (56-58% had no other rheumatic or chronic disease. 13-15% had other rheumatic disease, 33-35% had other chronic non-rheumatic disease.). 4. Site of osteoarthritis: Hand osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 11.5 (8.1) years
Indirectness of population	No indirectness
Interventions	(n=65) Intervention 1: Exercise - Unsupervised strength exercise. Exercise program focussing on strength performed 3 times weekly as 1 set of 10 repetitions in weeks 1-2, and 15 repetitions in weeks 3-12 with moderate to vigorous intensity. The

	programme was mainly home-based, but included also four group exercise sessions (weeks 1-3, 8). This included exercises of: shoulder extension, biceps curl, shoulder flexion, make an "O-sign", roll into a fist, thumb abduction/extension, grip strength, finger stretch.. Duration 12 weeks. Concurrent medication/care: Usual care included visits from general practitioners only, and very infrequently a referral to a consultation with an occupational therapist in secondary care. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=65) Intervention 2: No treatment. No specific attention. Duration 12 weeks. Concurrent medication/care: Usual care included visits from general practitioners only, and very infrequently a referral to a consultation with an occupational therapist in secondary care. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Financial support from The Norwegian Fund for Post-Graduate Training in Physiotherapy through the FYSIPRIm project and the Norwegian Rheumatism Association Research Fund is gratefully acknowledged.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: Function index of hand osteoarthritis at 3 months; Group 1: mean 10.3 (SD 4.7); n=65, Group 2: mean 10 (SD 4.8); n=65; FIHOA 0-30 Top=High is poor outcome; Comments: Baseline exercise: 10.8 (5.0). Baseline control: 9.8 (4.7).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, marital status, occupational status, education, self-reported hip or knee osteoarthritis, years with OA, fulfilment of ACR criteria, comorbidities; Group 1 Number missing: 8, Reason: 6 did not receive intervention (cardiovascular event, severe sickness, severe neck/shoulder pain, time/work), 2 did not have time; Group 2 Number missing: 2, Reason: 1 withdrew due to sickness, 1 did not have time

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: Function index of hand osteoarthritis at 6 months; Group 1: mean 10.9 (SD 5.4); n=65, Group 2: mean 10.5 (SD 4.9); n=65; FIHOA 0-30 Top=High is poor outcome; Comments: Baseline exercise: 10.8 (5.0). Baseline control: 9.8 (4.7).
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, marital status, occupational status, education, self-reported hip or knee osteoarthritis, years with OA, fulfilment of ACR criteria, comorbidities; Group 1 Number missing: 8, Reason: 6 did not receive intervention (cardiovascular event, severe sickness, severe neck/shoulder pain, time/work), 1 withdrew due to sickness, 1 did not attend due to

moving house; Group 2 Number missing: 3, Reason: 1 withdrew due to sickness, 1 did not attend due to illness, 1 declined to participate Protocol outcome 3: Pain at </=3 months - Actual outcome: Hand pain, NRS at 3 months; Group 1: mean 3.7 (SD 2.1); n=65, Group 2: mean 4.4 (SD 2); n=65; NRS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.2 (2.1). Baseline control: 3.9 (1.8). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, marital status, occupational status, education, self-reported hip or knee osteoarthritis, years with OA, fulfilment of ACR criteria, comorbidities; Group 1 Number missing: 8, Reason: 6 did not receive intervention (cardiovascular event, severe sickness, severe neck/shoulder pain, time/work), 2 did not have time; Group 2 Number missing: 2, Reason: 1 withdrew due to sickness, 1 did not have time Protocol outcome 4: Pain at > 3 months - Actual outcome: Hand pain, NRS at 6 months; Group 1: mean 4.3 (SD 2.3); n=65, Group 2: mean 4.3 (SD 2.1); n=65; NRS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.2 (2.1). Baseline control: 3.9 (1.8). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, marital status, occupational status, education, self-reported hip or knee osteoarthritis, years with OA, fulfilment of ACR criteria, comorbidities; Group 1 Number missing: 8, Reason: 6 did not receive intervention (cardiovascular event, severe sickness, severe neck/shoulder pain, time/work), 1 withdrew due to sickness, 1 did not attend due to moving house; Group 2 Number missing: 3, Reason: 1 withdrew due to sickness, 1 did not attend due to illness, 1 declined to participate Protocol outcome 5: Serious adverse events at > 3 months - Actual outcome: Adverse events at 6 months; Group 1: 8/65, Group 2: 0/65; Comments: Exercise: Increased pain and inflammation in one finger (1), increased pain and swelling of all fingers (2), people with previous neck/shoulder problems experiencing increased neck/shoulder pain related to the three shoulder exercises (5) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, marital status, occupational status, education, self-reported hip or knee osteoarthritis, years with OA, fulfilment of ACR criteria, comorbidities; Group 1 Number missing: 8, Reason: 6 did not receive intervention (cardiovascular event, severe sickness, severe neck/shoulder pain, time/work), 1 withdrew due to sickness, 1 did not attend due to moving house; Group 2 Number missing: 3, Reason: 1 withdrew due to sickness, 1 did not attend due to illness, 1 declined to participate	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	Park 2021[347]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=81)
Countries and setting	Conducted in South Korea; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Degenerative knee osteoarthritis diagnosed by bilateral radiographic examination
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Degenerative knee osteoarthritis diagnosed by bilateral radiographic examination; grade I or II knee osteoarthritis levels (early knee osteoarthritis); age over 60 years; female sex.
Exclusion criteria	The deformity of the knee, hip or back; central or peripheral nervous system involvement; administered any medications including steroids or intra-articular injection within previous three months, or previous surgery; pacemaker use; internal metallic materials; a history of impairment of a major organ system or a psychological disorder.
Recruitment/selection of patients	People from the Seoul Seniors Tower in Korea.
Age, gender and ethnicity	Age - Mean (SD): 66.9 (4.2). Gender (M:F): 0:81. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grades I or II Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=27) Intervention 1: Exercise - Supervised strength exercise. Isometric exercise group. Eight types of isometric movements were performed during the impulse phase as per the instructor's direction. The intensity was gradually increased from 60% of 1MT from baseline to week 2, 70% of 1MT from week 3 to week 5 and 80% of 1MT from week 6 to week 8.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=27) Intervention 2: No treatment. No treatment control.. Duration 8 weeks. Concurrent medication/care: No additional

	information.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=27) Intervention 3: Other. Isometric exercise and electromyostimulation delivered by a suit.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group does not fulfill the inclusion criteria in the protocol for this review and so was not included in the analysis.
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS knee-related quality of life at 8 weeks; Group 1: mean 21.53 (SD 30.66); n=25, Group 2: mean -5.4 (SD 9.06); n=25; KOOS quality of life 0-100 Top=High is good outcome; Comments: Baseline exercise: 18.50 (9.28). Baseline control: 16.25 (4.03).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body weight, skeletal muscle mass, fat mass, fat percent, basal metabolic rate and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 moved to a strange place; Group 2 Number missing: 2, Reason: 1 medication intake, 1 did not receive allocated assessment (far from research center)

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS activities of daily living at 8 weeks; Group 1: mean 4.21 (SD 5.04); n=25, Group 2: mean -1.52 (SD 6.12); n=25; KOOS activities of daily living 0-100 Top=High is good outcome; Comments: Baseline exercise: 40.26 (5.05). Baseline control: 42.84 (4.88).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body weight, skeletal muscle mass, fat mass, fat percent, basal metabolic rate and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 moved to a strange place; Group 2 Number missing: 2, Reason: 1 medication intake, 1 did not receive allocated assessment (far from research center)

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS pain at 8 weeks; Group 1: mean 3.57 (SD 7); n=25, Group 2: mean -0.63 (SD 9.78); n=25; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline exercise: 72.40 (5.82). Baseline control: 73.82 (7.36).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body weight, skeletal muscle mass, fat mass, fat percent, basal metabolic rate and baseline values of outcomes; Group 1 Number missing: 2, Reason: 2 moved to a strange place; Group 2 Number missing: 2, Reason: 1 medication intake, 1 did not receive allocated assessment (far from research center)

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Patrick 2001[348]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=249)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 20 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinically confirmed diagnosis of osteoarthritis from a physician
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Clinically confirmed diagnosis of osteoarthritis from a physician; aged 55-75; not currently exercising, defined as engaging in an average of less than 60 minutes of exercise per week during the last month; permission by the subject's primary physician to participate in the aquatic class; not currently enrolled in another medical study; living in an area where Arthritis foundation aquatic programs were offered; willingness to be randomized and to commit to the 5-month study period
Exclusion criteria	People scheduled for joint replacement surgery during the study peirod
Recruitment/selection of patients	People were recruited through direct invitation letters to Arthritis Foundation members, notices in their newsletter, network television coverage of the study, physician referrals, public service announcements, and newspaper advertisements.
Age, gender and ethnicity	Age - Other: Mean: 65.7. Gender (M:F): 34:215. Ethnicity: White = 94% (234)
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Not stated / Unclear (Unclear).
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=125) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Arthritis Foundation certified aquatic class - People engage in gentle upper- and lower-body ativities to help increase joint flexibility and range of motion, and maintain muscle strength. Treatment group participants were asked to attend class at least twice weekly for the 20-week study period. The number of classes offered per week varied from 2 to 7, class length ranged from 45 minutes to 1 hour,

	and class size ranged from 6 to 40 people with an average of 16.. Duration 20 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=124) Intervention 2: No treatment. Asked to maintain their usual activity levels. Duration 20 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This work was funded by the Centers for Disease Control and Prevention, grant number U48/CCU00954)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at > 3 months
- Actual outcome: Quality of Well-Being Scale at 20 weeks; Group 1: mean 0.606 (SD 0.069); n=101, Group 2: mean 0.599 (SD 0.079); n=121; Quality of Well-Being Scale 0-1 Top=High is good outcome; Comments: Baseline exercise: 0.597 (0.068). Baseline control: 0.599 (0.065).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, income, insurance, marital status, work status, education, living alone, race and baseline values of outcomes; Group 1 Number missing: 24, Reason: Specific reasons not given; Group 2 Number missing: 3, Reason: Specific reasons not given

Protocol outcome 2: Physical function at > 3 months
- Actual outcome: Health Assessment Questionnaire Disability at 20 weeks; Group 1: mean 0.933 (SD 0.55); n=101, Group 2: mean 1.127 (SD 0.671); n=121; HAQ - Disability 0-3 Top=High is poor outcome; Comments: Baseline exercise: 1.035 (0.535). Baseline placebo: 1.047 (0.608).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, income, insurance, marital status, work status, education, living alone, race and baseline values of outcomes; Group 1 Number missing: 24, Reason: Specific reasons not given; Group 2 Number missing: 3, Reason: Specific reasons not given

Protocol outcome 3: Pain at > 3 months
- Actual outcome: Health Assessment Questionnaire Pain at 20 weeks; Group 1: mean 1.382 (SD 0.737); n=98, Group 2: mean 1.462 (SD 0.619); n=117; HAQ - Pain 0-3 Top=High is poor outcome; Comments: Baseline exercise: 1.533 (0.602). Baseline control: 1.440 (0.610).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, income, insurance,

marital status, work status, education, living alone, race and baseline values of outcomes; Group 1 Number missing: 27, Reason: Specific reasons not given; Group 2 Number missing: 7, Reason: Specific reasons not given Protocol outcome 4: Psychological distress at > 3 months - Actual outcome: Centre for Epidemiological Studies Depression Scale at 20 weeks; Group 1: mean 6.956 (SD 4.729); n=101, Group 2: mean 8.092 (SD 6.005); n=113; Centre for Epidemiological Studies Depression Scale 0-60 Top=High is poor outcome; Comments: Baseline exercise: 7.261 (5.308). Baseline control: 7.715 (4.995). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, income, insurance, marital status, work status, education, living alone, race and baseline values of outcomes; Group 1 Number missing: 24, Reason: Specific reasons not given; Group 2 Number missing: 11, Reason: Specific reasons not given	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Pazit 2018[349]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=28)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis based on the presence of clinical symptoms of knee osteoarthritis as defined by the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with a clinical diagnosis of knee osteoarthritis based on the presence of clinical symptoms of knee osteoarthritis as defined by the American College of Rheumatology criteria with: age 60-90 years; knee pain for at least 6 months and experience current average page at least 3 (on an 11 point numerical rating scale); able to ambulate independently (with no more than a single point stick). In addition, people also had to have at least one of the following criteria indicating increased risk of falling: at least 1 fall in the past 12 months; had limited their activity level due to concern about falling.
Exclusion criteria	Any uncontrolled non-musculoskeletal conditions (such as chronic obstructive airways disease and congestive heart failure); a pre-existing neurological condition that affected lower limb strength, balance or ambulation (e.g. polio, stroke); any uncontrolled musculoskeletal or orthopaedic conditions that may affect ambulation (e.g. rheumatoid arthritis); currently taking part in a structured resistance training and/or organised balance training programs at least 1 time/week; any documented medical condition or physical impairment deemed by the participant's medical practitioner to contraindicate participation; mild cognitive impairment or dementia determined by a score <25 using the Saint Louis University Mental Status. Given that mild cognitive impairments increases the risk of falls, excluding people with mild cognitive impairment eliminated additional of potential confounding factor.
Recruitment/selection of patients	People were recruited from the general community of the Western suburbs of Melbourne through advertisements in local newspapers, health-care facilities and places with high circulation of senior citizens. Victoria University staff were also recruited through advertisements in university publications and posters displayed on notice boards, as well as global e-mails to staff and students, and social media.

Age, gender and ethnicity	Age - Mean (SD): 67.68 (6.68). Gender (M:F): 13:15. Ethnicity: No stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: High morbidity score (Hypertension: 17, Hypercholesterolaemia: 5, Diabetes Mellitus = 1, Depression = 2). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: At least 6 months
Indirectness of population	No indirectness
Interventions	(n=10) Intervention 1: Exercise - Supervised strength exercise. High speed resistance training including 6-8 exercises targeting the lower limbs (leg press, sit to stand, squat, step-up, calf raises lunges, going up stairs) supervised by a qualified Exercise Physiologist. They were asked to complete it in an explosive manner such that all repetitions for each shortening phase was performed as quickly as possible while the lengthening phase of the muscle was controlled over 2-3 seconds. Progression took place in three phases: phase 1 (week 1-2), two sets of 8-12 repetitions performed with 20-40% 1RM, phase 2 (week 3-5), two sets of 5-8 repetitions performed with 40-60% 1RM, phase 3 (week 6-8) two-three sets of 2-5 repetitions performed with 60-80% 1RM.. Duration 8 weeks. Concurrent medication/care: Not stated. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=10) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). High speed resistance training with balance training including 6-8 exercises targeting the lower limbs (leg press, sit to stand, squat, step-up, calf raises lunges, going up stairs) supervised by a qualified Exercise Physiologist. They were asked to complete it in an explosive manner such that all repetitions for each shortening phase was performed as quickly as possible while the lengthening phase of the muscle was controlled over 2-3 seconds. Progression took place in three phases: phase 1 (week 1-2), two sets of 8-12 repetitions performed with 20-40% 1RM, phase 2 (week 3-5), two sets of 5-8 repetitions performed with 40-60% 1RM, phase 3 (week 6-8) two-three sets of 2-5 repetitions performed with 60-80% 1RM. In addition there were six balance exercises (walking forward and backward, single leg standing, single leg tapping, side stepping and backward walking). Progression included decrease in base of support, decreased hand support, change of surface and reduced sensory input (e.g. eyes closed).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength and balance). (n=10) Intervention 3: No treatment. People from the control group were advised to continue with their usual activities, defined as any normal day-to-day activities and or any current usage of health services. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by Arthritis Australia)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Quality of life (Assessment of Quality of Life Scale) at 8 weeks; Group 1: mean 0.72 (SD 0.19); n=10, Group 2: mean 0.63 (SD 0.12); n=10; Assessment of Quality of Life scale 0-1 Top=High is good outcome; Comments: Baseline strength: 0.71 (0.18). Baseline control: 0.65 (0.10).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 unrelated cardiac issue; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 277.8 (SD 237); n=10, Group 2: mean 565.7 (SD 282.5); n=10; WOMAC function 0-1800 Top=High is poor outcome; Comments: Numbers written for WOMAC function don't make sense, and is likely an error (likely stiffness subscale is reported). Therefore, the values used here are those reported for WOMAC stiffness. Baseline strength: 732.3 (386.8). Baseline control: 534.2 (232.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 unrelated cardiac issue; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 117 (SD 132.6); n=10, Group 2: mean 249.7 (SD 309.3); n=10; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline strength: 181.6 (111.4)). Baseline control: 153.7 (91.1).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 unrelated cardiac issue; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury

Protocol outcome 4: Serious adverse events at </=3 months
- Actual outcome: Serious adverse events at 8 weeks; Group 1: 0/10, Group 2: 0/10
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 unrelated cardiac issue; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: Quality of life (Assessment of Quality of Life Scale) at 8 weeks; Group 1: mean 0.71 (SD 0.16); n=10, Group 2: mean 0.72 (SD 0.19); n=10; Assessment of Quality of Life scale 0-1 Top=High is good outcome; Comments: Baseline mixed: 0.71 (0.16). Baseline strength: 0.71 (0.18).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated cardiac issue

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 189.3 (SD 185.1); n=10, Group 2: mean 277.8 (SD 237); n=10; WOMAC function 0-1800 Top=High is poor outcome; Comments: Numbers written for WOMAC function don't make sense, and is likely an error (likely stiffness subscale is reported). Therefore, the values used here are those reported for WOMAC stiffness. Baseline mixed: 666.8 (438.5). Baseline strength: 732.3 (386.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated cardiac issue

Protocol outcome 3: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 97.3 (SD 127.1); n=10, Group 2: mean 117 (SD 132.6); n=10; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline mixed: 127.3 (107.8). Baseline strength: 181.6 (111.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated cardiac issue

Protocol outcome 4: Serious adverse events at </=3 months
- Actual outcome: Serious adverse events at 8 weeks; Group 1: 0/10, Group 2: 0/10

Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - ; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, height, weight, BMI, falls history, bilateral knee osteoarthritis, physical activity, other musculoskeletal conditions, other comorbidities, medication use and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated cardiac issue RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT Protocol outcome 1: Health related quality of life at </=3 months - Actual outcome: Quality of life (Assessment of Quality of Life Scale) at 8 weeks; Group 1: mean 0.71 (SD 0.16); n=10, Group 2: mean 0.63 (SD 0.12); n=10; Assessment of Quality of Life 0-1 Top=High is good outcome; Comments: Baseline mixed: 0.71 (0.16). Baseline control: 0.65 (0.10). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different in WOMAC pain and function subscales. Otherwise similar.; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC function at 8 weeks; Group 1: mean 277.8 (SD 237); n=10, Group 2: mean 565.7 (SD 282.5); n=10; WOMAC function 0-1800 Top=High is poor outcome; Comments: Numbers written for WOMAC function don't make sense, and is likely an error (likely stiffness subscale is reported). Therefore, the values used here are those reported for WOMAC stiffness. Baseline mixed: 666.8 (438.5). Baseline control: 534.2 (232.6). Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different in WOMAC pain and function subscales. Otherwise similar.; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 97.3 (SD 127.1); n=10, Group 2: mean 249.7 (SD 309.3); n=10; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline mixed: 127.3 (107.8). Baseline control: 153.7 (91.1). Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different in WOMAC pain and function subscales. Otherwise similar.; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Serious adverse events at 8 weeks; Group 1: 0/10, Group 2: 0/10 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Different in WOMAC pain and function subscales. Otherwise similar.; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 unrelated meniscus injury	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months;

	Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Peloquin 1999[350]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=124)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis confirmed by radiographs
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age at least 50; having an independent, non-institutional lifestyle; a stable regimen for using analgesics or nonsteroidal anti-inflammatory drugs for at least 2 weeks before the beginning of the intervention; a diagnosis of minimal to moderate idiopathic osteoarthritis of one or both knee joints; <15 degrees of fixed-flexion deformity; <10 degrees of genu varum or genu valgum; no joint blocking
Exclusion criteria	contraindication to participating in a supervised exercise program; expecting to be absent from the city for more than 2 weeks; having intra-articular steroid or viscoelastic device injections within the 2 months preceding the intervention period
Recruitment/selection of patients	All people were volunteers from the Sherbrooke metropolitan area who responded to various advertising media
Age, gender and ethnicity	Age - Mean (SD): 66.05 (7.89). Gender (M:F): 37:87. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Grade 1-3, median grade 2 Duration of symptoms (mean [SD]): 7.11 (7.03) years.
Indirectness of population	No indirectness
Interventions	(n=59) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Aerobic, muscle strengthening and stretching exercises delivered over 12 weeks. Aerobic exercises progressing from 3x4 minute sessions with 1 minute rests up to 1x17 minutes session. Muscle strengthening exercises increasing from 1x3 reps of isometric (quadriceps and hamstrings) exercises up to isotonic 9-11exercises of 1x10 maximum repetitions. Stretching

	exercises increasing from 3x15s/exercise for 2-5 exercises up to 2x15s/exercise for 9-12 exercises.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength, aerobic). (n=65) Intervention 2: No treatment. Maintain usual care/activities. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was funded by the Canadian Fitness and Lifestyle Research Institute, Grand #901R008)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: AIMS2 mobility level at 12 weeks; Group 1: mean 1.08 (SD 1.11); n=59, Group 2: mean 1.58 (SD 1.33); n=65; AIMS2 mobility level 0-10 Top=High is poor outcome; Comments: Baseline exercise: 1.30 (1.33). Baseline control: 1.57 (1.45).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 walking and bending at 12 weeks; Group 1: mean 1.64 (SD 1.89); n=59, Group 2: mean 2.89 (SD 2.78); n=65; AIMS2 walking and bending 0-10 Top=High is poor outcome; Comments: Baseline exercise: 3.14 (2.43). Baseline control: 3.43 (2.61).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 hand and finger function at 12 weeks; Group 1: mean 0.52 (SD 1.08); n=59, Group 2: mean 0.62 (SD 1.29); n=65; AIMS2 hand and finger function 0-10 Top=High is poor outcome; Comments: Baseline exercise: 0.87 (1.53). Baseline control: 0.61 (1.23).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3,

Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 arm function at 12 weeks; Group 1: mean 0.26 (SD 0.61); n=59, Group 2: mean 0.39 (SD 1.1); n=65; AIMS2 arm function 0-10 Top=High is poor outcome; Comments: Baseline exercise: 0.37 (0.73). Baseline control: 0.43 (0.77).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 self-care tasks at 12 weeks; Group 1: mean 0.05 (SD 0.33); n=59, Group 2: mean 0.06 (SD 0.39); n=65; AIMS2 self-care tasks 0-10 Top=High is poor outcome; Comments: Baseline exercise: 0.04 (0.20). Baseline control: 0.19 (0.54).
Risk of bias: All domain - Very high, Selection – Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 household tasks at 12 weeks; Group 1: mean 0.11 (SD 0.45); n=59, Group 2: mean 0.35 (SD 1.23); n=65; AIMS2 household tasks 0-10 Top=High is poor outcome; Comments: Baseline exercise: 0.12 (0.49). Baseline control: 0.82 (1.99).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 social activity at 12 weeks; Group 1: mean 5.34 (SD 1.65); n=59, Group 2: mean 5.42 (SD 1.48); n=65; AIMS2 social activity 0-10 Top=High is poor outcome; Comments: Baseline exercise: 5.42 (1.60). Baseline control: 5.45 (1.54).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 support from family and friends at 12 weeks; Group 1: mean 1.85 (SD 2.26); n=59, Group 2: mean 1.93 (SD 1.88); n=65; AIMS2 support from family and friends 0-10 Top=High is poor outcome; Comments: Baseline exercise: 2.36 (2.38). Baseline control: 2.47 (2.25).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.
- Actual outcome: AIMS2 arthritis pain at 12 weeks; Group 1: mean 3.09 (SD 1.54); n=59, Group 2: mean 3.94 (SD 2.22); n=65; AIMS2 arthritis pain 0-10 Top=High is poor outcome; Comments: Baseline exercise: 4.53 (2.02). Baseline control: 4.53 (2.20).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above. - Actual outcome: AIMS2 work at 12 weeks; Group 1: mean 0.89 (SD 1.13); n=59, Group 2: mean 1.28 (SD 1.64); n=65; AIMS2 work 0-10 Top=High is poor outcome; Comments: Baseline exercise: 1.90 (2.08). Baseline control: 1.39 (1.49). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above. - Actual outcome: AIMS2 mood at 12 weeks; Group 1: mean 1.54 (SD 1.46); n=59, Group 2: mean 1.7 (SD 1.57); n=65; AIMS2 mood 0-10 Top=High is poor outcome; Comments: Baseline exercise: 1.72 (1.36). Baseline control: 1.77 (1.40). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above. - Actual outcome: AIMS2 level of tension at 12 weeks; Group 1: mean 3.03 (SD 1.95); n=59, Group 2: mean 3.45 (SD 2.02); n=65; AIMS2 level of tension 0-10 Top=High is poor outcome; Comments: Baseline exercise: 3.64 (1.88). Baseline control: 3.74 (1.81). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 10, Reason: 10 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time. One had to quit after knee inflammation because of the exercises performed.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above. Protocol outcome 2: Serious adverse events at </=3 months - Actual outcome: Increased inflammation after exercise at 12 weeks; Group 1: 1/59, Group 2: 0/65 Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Self care tasks different at baseline, otherwise similar; Group 1 Number missing: 9, Reason: 9 dropped out. The most frequently cited reasons for dropping out were the occurrence of a medical problem unrelated to the interventions and a lack of time.; Group 2 Number missing: 3, Reason: 3 dropped out. Reasons as above.	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at </=3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Petrella 2000[355]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=179)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographic evidence of knee osteoarthritis in the tibial-femoral compartment (grade 1-3)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People in a fasted state who had stopped NSAID medications for 48 hours prior to study entry with the following: age >65 years; pain in one knee on most days; radiographic evidence of osteoarthritis in the tibial femoral compartment (grade 1-3); and difficulty with performing activities of daily living including walking one city block, rising from a chair, getting out of bed, or performing shopping, cleaning or self-care activities
Exclusion criteria	Participating in another research study; had comorbidities precluding their safe involvement in exercise including recent stroke or myocardial infarction (in the past 3 months); unstable metabolic or cardiovascular disease; severe systemic disease; psychiatric illness; contraindication or intolerance to oxaprozin; living in a dependent environment i.e., nursing home); had arthritis other than osteoarthritis (i.e. rheumatoid arthritis); had recent (<3 years) gastrointestinal haemorrhage or gastric/duodenal ulceration; had history of inflammatory bowel disease; were taking acetylsalicylic acid for reasons other than cardioprotection, or warfarin, oral or intramuscular corticosteroids within 3 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 73.7 (4.9). Gender (M:F): 69:103. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (61-67% have other chronic comorbidities (i.e. cardiac, respiratory, metabolic)). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-3, median grade 1 Duration of symptoms: Not stated

Indirectness of population	No indirectness
Interventions	(n=91) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). The exercise program included a serioes of progressive, simple, range or motiona nd resistance exercises utilizing common items in the home. All exercises were performed at home. Each session consisted of a 10 minute warmup/lower extremity stretching followed by a specific series of repetitions, sets, frequency, and resistance. All exercise sessions were recorded by people in a diary and reviewed. Progression occurred every two weeks from 2 repetitions per session, with 3 weeks per session and 1 session per day lasting 10 minutes, up to 5 repetitions with 5 sessions per week and 3 sessions per day for 15 minutes each.. Duration 8 weeks. Concurrent medication/care: All people were given oxaprozin 1200mg/day during the study period. All people were given paracetamol 325mg to be taken every 4-6 hours as needed for rescue pain therapy.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and range of motion/flexibility). (n=88) Intervention 2: Exercise - Other unsupervised exercise (including flexibility, proprioception). Non-weight bearing joint unloading and stretches which did not include resistance of progression. Duration 8 weeks. Concurrent medication/care: All people were given oxaprozin 1200mg/day during the study period. All people were given paracetamol 325mg to be taken every 4-6 hours as needed for rescue pain therapy.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Stretching).
Funding	Study funded by industry (Supported by an unrestricted educational grant from Monsanto Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus OTHER UNSUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION)

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 14 (SD 6); n=91, Group 2: mean 5 (SD 3); n=88; WOMAC physical function 0-100 Top=High is poor outcome; Comments: Does not report baseline values
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms,

grade, prior NSAID use, osteoarthritis in other joints, regular paracetamol use, use of assistive aids, annual income >$20000, education, post secondary and other chronic comorbidities. Does not report baseline values for outcomes; Group 1 Number missing: 1, Reason: 1 withdrew before repeat testing; Group 2 Number missing: 3, Reason: 2 withdrew before repeat testing and 1 was lost to follow up Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 18 (SD 9); n=91, Group 2: mean 11 (SD 7); n=88; WOMAC pain 0-100 Top=High is poor outcome; Comments: Does not report baseline values. Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, grade, prior NSAID use, osteoarthritis in other joints, regular paracetamol use, use of assistive aids, annual income >$20000, education, post secondary and other chronic comorbidities. Does not report baseline values for outcomes; Group 1 Number missing: 1, Reason: 1 withdrew before repeat testing; Group 2 Number missing: 3, Reason: 2 withdrew before repeat testing and 1 was lost to follow up Protocol outcome 3: Serious adverse events at </=3 months - Actual outcome: Adverse events at 8 weeks; Group 1: 5/91, Group 2: 8/88; Comments: Adverse events included abdominal pain (3%), dyspepsia (5%), fatigue (2%) Risk of bias: All domain – Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, duration of symptoms, grade, prior NSAID use, osteoarthritis in other joints, regular paracetamol use, use of assistive aids, annual income >$20000, education, post secondary and other chronic comorbidities. Does not report baseline values for outcomes; Group 1 Number missing: 1, Reason: 1 withdrew before repeat testing; Group 2 Number missing: 3, Reason: 2 withdrew before repeat testing and 1 was lost to follow up	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Ravaud 2004[369]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=2957)
Countries and setting	Conducted in France; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People who met the clinical and radiographic American College of Rheumatology criteria for osteoarthritis of the knee or hip
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with clinical and radiographic features of osteoarthritis meeting the American College of Rheumatology criteria; at least 6 months history of pain; pain scored by the person at least at 30mm on a 100mm visual analogue scale; pain for at least 14 days during the month preceding the study
Exclusion criteria	People with secondary arthritis as defined by Osteoarthritis Research Society International; had comorbidities that precluded their safe involvement in the exercise programme (such as recent myocardial infarction); had surgery scheduled within the 12 months following the start of the study or had serious concomitant illness (neoplasia, infectious disease, unstable metabolic or cardiovascular disease, systemic disease); had received any intra-articular injection (hyaluronic acid, corticosteroid, or joint lavage) during the 3 months preceding the study; had used slow acting anti-osteoarthritic drugs during the 2 months preceding the study; were participating in another research study
Recruitment/selection of patients	Open cluster RCT. Each rheumatologist was to enroll four people with osteoarthritis (three with knee osteoarthritis, one with hip)
Age, gender and ethnicity	Age - Mean (SD): 66.78 (10.39). Gender (M:F): 883:2074. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Knee or hip osteoarthritis).
Extra comments	Severity: Kellgren and Lawrence grade 2-4, median grade 3 Duration of symptoms (mean [SD]): 69.5 (75.5) months
Indirectness of population	No indirectness

Interventions	(n=735) Intervention 1: Exercise - Unsupervised strength exercise. During the initial visit, the rheumatologist gave an oral explanation of the importance of exercise for osteoarthritis. All people received a booklet illustrating the exercises and a videotape. The videotape presentation comprised tow parts: a motivational portion designed to address the interest of exercise for people with osteoarthritis and to provide positive role models for exercise; a 30 minute programme of the five exercise routines performed by a trained demonstrator. The five exercises were designed to improve joint mobility and increase muscle power. These exercises are derived from the programmes previously described for knee and hip osteoarthritis, and are different for each type of osteoarthritis. All exercises were to be performed at home. Each exercise was initially to be repeated 10 times and, if pain allowed, increased in increments of 5 repetitions each week up to a maximum of 30, The overall adherence goal was to perform the programme four times each week for 6 months according to an agreed level of resistance.. Duration 6 months. Concurrent medication/care: All people received the non-steroidal anti-inflammatory drug rofecoxib. The drug was administered once daily at 12.5mg during the first month and thereafter at 25mg if necessary. People were permitted to take paracetamol if necessary.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=760) Intervention 2: No treatment. Usual care no treatment. Duration 6 months. Concurrent medication/care: All people received the non-steroidal anti-inflammatory drug rofecoxib. The drug was administered once daily at 12.5mg during the first month and thereafter at 25mg if necessary. People were permitted to take paracetamol if necessary.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=1462) Intervention 3: Other. Exercise with standardised tools or standardised tools alone. Duration 6 months. Concurrent medication/care: All people received the non-steroidal anti-inflammatory drug rofecoxib. The drug was administered once daily at 12.5mg during the first month and thereafter at 25mg if necessary. People were permitted to take paracetamol if necessary.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These two groups were not included in the analysis as they did not fulfill

	the inclusion criteria
Funding	Study funded by industry (This study was supported by Merck Sharp and Dohme, Chibret, France)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: WOMAC function subscale at 6 months; Group 1: mean -12.4 (SD 19.2); n=735, Group 2: mean -11.1 (SD 20.2); n=760; WOMAC function subscale 0-100 Top=High is poor outcome; Comments: Baseline exercise: 45.3 (18.1). Baseline no treatment: 45.7 (18.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, prior treatment, mean duration of symptoms, radiological grade and baseline values of outcomes; Group 1 Number missing: 108, Reason: Lost to follow up = 108; Group 2 Number missing: 116, Reason: Lost to follow up = 116

Protocol outcome 2: Pain at > 3 months
- Actual outcome: Pain (100mm VAS) at 6 months; Group 1: mean -19.7 (SD 28.7); n=735, Group 2: mean -19.1 (SD 28.8); n=760; VAS 0-100 Top=High is poor outcome; Comments: Baseline exercise: 59.2 (17.7). Baseline no treatment: 59.6 (17.9).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, gender, BMI, prior treatment, mean duration of symptoms, radiological grade and baseline values of outcomes; Group 1 Number missing: 108, Reason: Lost to follow up = 108; Group 2 Number missing: 116, Reason: Lost to follow up = 116

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Rewald 2020[374]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=111)
Countries and setting	Conducted in Netherlands; Setting: Outpatient follow up

Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks (end of intervention), 24 weeks (follow up for an additional 12 weeks)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee pain between 4 and 7 on a 10-point numeric rating scale and a Kellgren-Lawrence score between 1 and 3.
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with knee osteoarthritis who rated their knee pain between 4 and 7 on a 10-point numeric rating scale and a Kellgren Lawrence score between 1 and 3; people had a clear indication for conservative treatment of osteoarthritis, including a primary care physical therapy referral; people had to be able to cycle on a stationary bicycle and had to score 8 points or lower on a stationary bicycle and had to score 8 points or lower on the Hospital Anxiety and Depression Scale.
Exclusion criteria	Contraindications for aquatic exercise therapy; planned total knee surgery; corticosteroid injection less than 3 months prior to study participation or hyaluronic acid injection less than 6 months prior to study participation; severe joint complaints elsewhere; symptomatic and radiologically proven hip osteoarthritis; inflammatory joint diseases and the inability to safely enter and exit the pool.
Recruitment/selection of patients	People were recruited from the Early Osteoarthritis Outpatient Clinic of the MUMC+ between March 2013 and October 2015.
Age, gender and ethnicity	Age - Mean (SD): 59.9 (8.7). Gender (M:F): 39:63. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (Mean count comorbidity: intervention = 2 (1.7), usual care = 1 (1.3).). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence score between 1-3. Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=55) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic exercise. People were instructed not to start additional physical therapy during the intervention period of 12 weeks. Participants exercised twice per week for 45 minutes in groups of maximally 4 people supervised by a physical therapist. Participants cycled in an upright position on the AquaCruiser II aqua bike. The depth of the therapy pool was adjusted to ensure that the legs were immersed during the whole movement. Typically, participants were immersed between the xiphoid process and the first rib in the warm water (32 degrees centigrade). The main part of the training consisted of cycling in a sitting position with good postural control. Also, out-of-the-saddle positions, leg exercises and upper body exercises were incorporated. Exercise intensity was monitored using the Borg Scale, heart rate (220-age) formula and pedaling tempo.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy

	(n=47) Intervention 2: No treatment. People were not prohibited to follow treatment that they would have also received outside of the trail. Thus, patients were free to start physical therapy or to use aids (as braces) to ease their complaints. Following physical therapy was not obliged, and was not considered as part of the study. After 24 weeks, they were offered 12 weekly sessions of aquatic cycling in a local community pool in Maastricht.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by The Netherlands Organisation for Scientific Research (grant no. 022.003.036). The Maastricht University Medical Center+ financed the cycling equipment.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS quality of life at 12 weeks; MD; 13 (95%CI 5.852 to 20.215) (SE: 3.61) ;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, Kellgren Lawrence grade and mean count comorbidity.; Group 1 Number missing: 26, Reason: 3 did not receive the allocated intervention (2 urgent treatment comorbidity, 1 unknown). 6 lost to follow up (1 unknown, 2 menisectonomy, 3 urgent treatment comorbidity) at post-intervention. 18 discontinued intervention (6 exacerbation comorbidity, 1 dizziness, 7 personal or work-related reasons, 3 menisectomy, 1 total knee arthroplasty - 3 of the people from this group are counted in the lost to follow up group). At follow up 4 lost to follow-up (1 menisectomy - counted for amongst the discontinued intervention), 2 comorbidity and 1 total knee arthroplasty (counted for in the discontinued intervention group)).; Group 2 Number missing: 9, Reason: 6 lost to follow-up post-intervention (4 withdrew, 1 urgent treatment comorbidity, 1 busy work schedule). 3 lost to follow-up at follow-up (2 urgent treatment comorbidity, 1 busy work schedule).

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS physical function at 12 weeks; MD; 7.16 (95%CI 0.83 to 13.49) (SE: 3.19) KOOS physical function 0-100 Top=High is good outcome;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, Kellgren Lawrence grade and mean count comorbidity.; Group 1 Number missing: 26, Reason: 3 did not receive the allocated intervention (2 urgent treatment comorbidity, 1 unknown). 6 lost to follow up (1 unknown, 2 menisectonomy, 3 urgent treatment comorbidity) at post-intervention. 18 discontinued intervention (6 exacerbation comorbidity, 1 dizziness, 7 personal or work-related reasons, 3 menisectomy, 1 total knee arthroplasty - 3 of the people from this group are counted in the lost to follow up group). At follow up 4 lost to follow-up (1 menisectomy - counted for amongst the discontinued intervention), 2 comorbidity and 1 total knee arthroplasty (counted for in the discontinued intervention group)).; Group 2 Number missing: 9, Reason: 6 lost to follow-up post-intervention (4 withdrew, 1 urgent treatment comorbidity, 1 busy work schedule). 3 lost to follow-up at follow-up (2 urgent treatment comorbidity, 1 busy work schedule).

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS pain at 12 weeks; MD; 8.16 (95%CI 1.67 to 14.64) (SE: 3.27) ;
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, Kellgren Lawrence grade and mean count comorbidity.; Group 1 Number missing: 26, Reason: 3 did not receive the allocated intervention (2 urgent treatment comorbidity, 1 unknown). 6 lost to follow up (1 unknown, 2 menisectonomy, 3 urgent treatment comorbidity) at post-intervention. 18 discontinued intervention (6 exacerbation comorbidity, 1 dizziness, 7 personal or work-related reasons, 3 menisectomy, 1 total knee arthroplasty - 3 of the people from this group are counted in the lost to follow up group). At follow up 4 lost to follow-up (1 menisectomy - counted for amongst the discontinued intervention), 2 comorbidity and 1 total knee arthroplasty (counted for in the discontinued intervention group)).; Group 2 Number missing: 9, Reason: 6 lost to follow-up post-intervention (4 withdrew, 1 urgent treatment comorbidity, 1 busy work schedule). 3 lost to follow-up at follow-up (2 urgent treatment comorbidity, 1 busy work schedule). Protocol outcome 4: Serious adverse events at </=3 months - Actual outcome: Adverse events at 12 weeks; Group 1: 15/55, Group 2: 0/47; Comments: Aquatic exercise: 1 hospitalised after hyperventilation with a history of cardiovascular symptoms; 4 exacerbation of symptoms, 10 increased knee pain the day after training. Usual care: No adverse events. Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, Kellgren Lawrence grade and mean count comorbidity.; Group 1 Number missing: 26, Reason: 3 did not receive the allocated intervention (2 urgent treatment comorbidity, 1 unknown). 6 lost to follow up (1 unknown, 2 menisectonomy, 3 urgent treatment comorbidity) at post-intervention. 18 discontinued intervention (6 exacerbation comorbidity, 1 dizziness, 7 personal or work-related reasons, 3 menisectomy, 1 total knee arthroplasty - 3 of the people from this group are counted in the lost to follow up group). At follow up 4 lost to follow-up (1 menisectomy - counted for amongst the discontinued intervention), 2 comorbidity and 1 total knee arthroplasty (counted for in the discontinued intervention group)).; Group 2 Number missing: 9, Reason: 6 lost to follow-up post-intervention (4 withdrew, 1 urgent treatment comorbidity, 1 busy work schedule). 3 lost to follow-up at follow-up (2 urgent treatment comorbidity, 1 busy work schedule).	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	**Rezasoltani 2020[376]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 4 weeks (end of intervention), 12 weeks (follow up total)

Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis with knee pain for at least 3 months
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Men with knee osteoarthritis; were 60 years or older; had knee pain for at least 3 months.
Exclusion criteria	A history of intra-articular corticosteroid injections within the last 3 months or hyaluronic acid injection within 6 months; oral anti-inflammatory drugs within the last week; a history of surgery on knee joint or major trauma to the lower limb causing fracture; a body mass index more than 34 kg/m²; knee joint pathologies such as osteonecrosis, severe osteoporosis and rheumatoid arthritis; systemic diseases that affect knee joint such as collagen vascular diseases or gout; addiction to narcotic; diabetes mellitus.
Recruitment/selection of patients	People recruited from the outpatient clinic of the Department of Physical Medicine and Rehabilitation at University Hospital (Imam Reza) from Aja University of Medical Sciences, Tehran, Iran.
Age, gender and ethnicity	Age - Mean (SD): 51.0 (2.93). Gender (M:F): 32:0. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of symptoms: At least 3 months.
Indirectness of population	No indirectness
Interventions	(n=16) Intervention 1: Exercise - Supervised strength exercise. Aquatic cycling exercise, 3 sessions per week for 4 weeks totally 12 sessions. The sessions were held in a community pool and guided by a physiotherapist certified in aquatic physiotherapy. The water depth was 1.2m and the temperature was kept at about 32 degrees centigrade (89 degrees F). Each session lasted approximately 50 minutes including 10 minutes of warm-up, 30 minutes of cycling and 10 minutes of cool-down exercises. Maximum height of the water was up to the xiphoid process. A water-resistant stationary aqua bike was used by the patients. Participants were instructed to exercise pedal with the intensity of 40% to 60% of their reserve heart rate. The group was informed not to use other forms of physical therapy during the study interval.. Duration 4 weeks. Concurrent medication/care: People were instructed to use paracetamol if needed and to follow lifestyle recommendations to use their knees more appropriately.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Hydrotherapy (n=16) Intervention 2: No treatment. No additional treatment. Duration 4 weeks. Concurrent medication/care: People were instructed to use paracetamol if needed and to follow lifestyle recommendations to use their knees more appropriately.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable

Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: KOOS physical function at 12 weeks; Group 1: mean 80.2 (SD 1.8); n=15, Group 2: mean 58.4 (SD 3.5); n=15; KOOS physical function 0-100 Top=High is good outcome; Comments: Baseline exercise: 55.5 (2.5). Baseline no treatment: 54.9 (1.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 decline to continue; Group 2 Number missing: 1, Reason: 1 desire to attend aquatic exercise

Protocol outcome 2: Pain at </=3 months
- Actual outcome: KOOS pain at 12 weeks; Group 1: mean 80.8 (SD 1.8); n=15, Group 2: mean 55.3 (SD 3.5); n=15; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline exercise: 54.8 (1.3). Baseline control: 55.7 (3.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, weight, height, BMI and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 decline to continue; Group 2 Number missing: 1, Reason: 1 desire to attend aquatic exercise

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Robbins 2022[377]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=215)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed by an independent rehabilitation specialist
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged from 50 and 75; symptomatic knee osteoarthritis for at least three months; visual analogue scale score above 3; radiographic knee osteoarthritis compatible with Kellgren-Lawrence grade two or higher.
Exclusion criteria	Contraindication to laser application (e.g. cancer and insulin-dependent diabetes); inability to perform the assessment or treatment; continuous use of anti-inflammatory drugs, symptomatic hip osteoarthritis and physiotherapeutic knee treatment within the last 3 months.
Recruitment/selection of patients	People who attended the Special Rehabilitation Services in Taboao da Serra-SP
Age, gender and ethnicity	Age - Mean (SD): 63.5 (6.4). Gender (M:F): 45:170. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Osteoarthritis degree 2-4, median grade 3 Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=86) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Two groups combined: Group 1 = Stretching exercises for eight weeks (n=43) involving group-based exercises (groups of 5-7 people), three times per week for 8 weeks. Sessions lasted 45 minutes including a 10-minute warm-up followed by seven stretching exercises, repeated four times for both legs and sustained for 30 seconds each. The exercises aimed to stretch the major muscles of the posterior and antero-internal hip muscle chains including paraspinal muscle, gluteus, iliopsoas, hamstrings, quadriceps, hip adductors and gastrocnemius. Five

	exercises were executed in supine, one seated and the last in an upright position. Group 2 = Stretching exercise and laser therapy (low-level laser therapy - 3J energy per point with a total dose of 27J per treatment. Gallium arsenide semiconductor with wavelength of 904nm, average power 40mW, peak power 70W, pulse duration 60ns, pulse repetition rate 9500Hz and beam area 0.1cm2. Delivered three times a week) (n=43) for the same time period.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Flexibility Comments: These two groups were combined as they fulfilled the same class effect when being compared to other treatments, that have also been combined, as in the protocol. (n=86) Intervention 2: No treatment. Two groups: group 1 = laser only (same experimental approach as in the intervention group) (n=43). Group 2 = educational booklet only (no treatment control).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: These two groups were combined as they fulfilled the same class effect when being compared to other treatments, that have also been combined, as in the protocol.
Funding	Academic or government funding (This study was funded by Fundacao de Amparo a Pesquisa do Estado de Sao Paulo (FAPESP) (2012/01827-3), Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES) (institutional) and Conselho Nacional de Desenvolvimento Cientifico e Technologico (CNPq) (248967/2013-4).)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT Protocol outcome 1: Pain at </=3 months - Actual outcome: Pain in daily life activities (NRS) at 8 weeks; Group 1: mean -3.62 (SD 2.73); n=86, Group 2: mean -1.43 (SD 2.51); n=86; NRS 0-10 Top=High is poor outcome; Comments: Combination of values. Reported Laser + stretch = 4.64 (2.69). Reported stretch = 2.59 (2.35). Reported laser = 3.68 (2.28). Reported control = 0.27 (2.1). Baseline values not reported Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, BMI, osteoarthritis degree, gender and	

medication intake frequency; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Rogers 2011[379]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=20)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Self-reported knee pain with physician diagnosed knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 45 years or over of either gender; self reported knee pain; physician diagnosed knee osteoarthritis, unilateral or bilateral; demonstrated minial knee osteoarthritis related dysfunction per WOMAc LK 3.1 score of 17 or above on 68 point physical function sub-scale; not engaged in a regular leg exercise program for minimum of 6 months
Exclusion criteria	Inability to obtain physician release for exercise; unresolved balance disorder; unresolved neurological disorder; history of knee surgery or major knee trauma injury; hip or ankle instability, excessive weaknesss, surgery or major trauma injury; hip or knee replacement; intra-articular joint injection within 4 weeks of the study
Recruitment/selection of patients	People were recruited from the Tampa Bay Florida region via announcements, advertisements, word of mouth and physician referral
Age, gender and ethnicity	Age - Mean (SD): 71.16 (11.26). Gender (M:F): 4:16. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=8) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Kinaesthesia, balance and agility neuromuscular exercises conducted as three 30 minute sessions per week with a five minute warmup and post-workout static stretching. Including: the wedding march, backward wedding march, high knees march, side stepping, semi-tandem walk, tandem walk, cross-over walk, modified

	grapevine, toe walking and heel walking. Up to three sets of up to 30 seconds of each exercise was conducted.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Neuromodulatory (n=12) Intervention 2: Exercise - Supervised strength exercise. Strength training exercises conducted as three 30 minute sessions per week with a five minute warmup and post-workout static stretching. Including: seated resistance band exercises; standing hip hyperextension with resistance band; standing wall slides (partial squats) with a small "play ball" behind the back and supine heel slides (hip and knee flexion and extension). Completed with 10-15 repetition maximum for each exercise. Duration 8 weeks. Concurrent medication/care: No additional informtion. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 6 weeks; Group 1: mean -16.5 (SD 5.69); n=4, Group 2: mean -9.37 (SD 15.94); n=8; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline other: 30.67 (9.97). Baseline strength: 30.22 (SD not reported).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, SBP, DBP, resting HR, gender and baseline values of outcomes; Group 1 Number missing: 4, Reason: reports that 2 were lost to follow up due to increased pain. But outcome reports 4 people were included in the outcome.; Group 2 Number missing: 4, Reason: Reports 1 discontinued due to "collapsing" knee sensation, and 2 discontinued due to reasons unrelated to the study. However, study reports 4 were not included.

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean -2.67 (SD 4.41); n=6, Group 2: mean -4 (SD 4.79); n=9; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline other: 6.83 (4.07). Baseline strength: 9.44 (3.28).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, BMI, SBP, DBP, resting HR, gender and baseline values of outcomes; Group 1 Number missing: 2, Reason: reports that 2 were lost to follow up due to increased pain.; Group 2 Number

missing: 3, Reason: Reports 1 discontinued due to "collapsing" knee sensation, and 2 discontinued due to reasons unrelated to the study.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Rogers 2012[380]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=44)
Countries and setting	Conducted in South Africa; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: met the American College of Rheumatology diagnostic criteria for unilateral or bilateral symptomatic knee osteoarthritis as confirmed by the person's physician
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 50 or older; self-reported knee pain on most days of the previous month; met American College of Rheumatology diagnostic criteria for unilateral or bilateral symptomatic knee osteoarthritis as confirmed by the participant's physician; not engaged in lower extremity exercise program for a minimum of six months prior to enrollment; minimum disability score of 17 points on the Physician Function sub-scale of the WOMAC.
Exclusion criteria	Rheumatic disease other than osteoarthritis; high risk health status for exercise; inability to obtain physician release for exercise; unresolved balance or neurological disorder; history of major knee surgery; major knee trauma; hip or knee arthroplasty; hip or ankle instability or excessive weakness; intra-articular joint injection within 4 weeks of beginning the study
Recruitment/selection of patients	People recruited from the Tampa Bay, Florida, USA community via newspaper announcements and advertisements, posted fliers, work of mouth and internet postings
Age, gender and ethnicity	Age - Mean (SD): 70.4 (9.8). Gender (M:F): 20:13. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=11) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Kinaesthesia, balance and agility exercise training with resistance exercise training. Training for 8 weeks three times a week for 30-40 minutes. Kinaesthesia, balance and agility exercises included leg static and dynamic balancing. Agility exercises preceded the balance exercises and were progressed by adding repetitions. People gradually began with 15 steps and progressed to a maximum of 75 steps per agility exercise. Balance exercises were conducted on either the floor or on Thera-Band stability trainer pads of two difficulty levels. People completed up to 3 sets up to 30 seconds per set. Both legs were trained. For static balance, the aim was to stay steady for as long as possible (up to 30 seconds) while dynamic balance required the addition of small, rapid bouncing movements. People were taught to flex and extend the knee about 5 to 10 degrees maximum during dynamic balance. Resistance training participants were trained to use the Thera-Band to perform a single 15-repetition set of lower extremity exercises with each leg. The program utilized primarily seated, open chain exercises to train the major muscle groups without challenging balance or agility. Exercises were progressed by adding greater stretch to the prescribed band to given greater resistance or by moving up to the next strength of resistance band.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Other (Neuromodulatory and strength). (n=11) Intervention 2: Exercise - Other supervised exercise (including flexibility, proprioception). Kinaesthesia, balance and agility exercise training. Training for 8 weeks three times a week for 30-40 minutes. Kinaesthesia, balance and agility exercises included leg static and dynamic balancing. Agility exercises preceded the balance exercises and were progressed by adding repetitions. People gradually began with 15 steps and progressed to a maximum of 75 steps per agility exercise. Balance exercises were conducted on either the floor or on Thera-Band stability trainer pads of two difficulty levels. People completed up to 3 sets up to 30 seconds per set. Both legs were trained. For static balance, the aim was to stay steady for as long as possible (up to 30 seconds) while dynamic balance required the addition of small, rapid bouncing movements. People were taught to flex and extend the knee about 5 to 10 degrees maximum during dynamic balance. . Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Neuromodulatory

	(n=11) Intervention 3: Exercise - Supervised strength exercise. Kinaesthesia, balance and agility exercise training with resistance exercise training. Training for 8 weeks three times a week for 30-40 minutes. Resistance training participants were trained to use the Thera-Band to perform a single 15-repetition set of lower extremity exercises with each leg. The program utilized primarily seated, open chain exercises to train the major muscle groups without challenging balance or agility. Exercises were progressed by adding greater stretch to the prescribed band to given greater resistance or by moving up to the next strength of resistance band.. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Not applicable (n=11) Intervention 4: Other. Inert skin lotion applied to the affected knee once daily. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Topical treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group was not included as it did not fulfill the inclusion criteria
Funding	Study funded by industry (This research was supported by a product grant from The Thera-Band Academy, which provided elastic resistance bands and stability trainers)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION)

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 13.89 (SD 9.44); n=11, Group 2: mean 20 (SD 9.2); n=11; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 30.11 (7.67). Baseline other: 27.50 (8.25).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and baseline values of outcomes; Group 1 Number missing: 2, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1)

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 5 (SD 3.35); n=11, Group 2: mean 4.87 (SD 3.6); n=11; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline mixed: 8.33 (2.18). Baseline other: 6.87 (2.75).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and baseline values of outcomes; Group 1 Number missing: 2, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 13.89 (SD 9.44); n=11, Group 2: mean 16.25 (SD 12.53); n=11; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 30.11 (7.67). Baseline strength: 29.75 (6.82).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and baseline values of outcomes; Group 1 Number missing: 2, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1)

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 5 (SD 3.35); n=11, Group 2: mean 4.25 (SD 3.45); n=11; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline mixed: 8.33 (2.18). Baseline strength: 8.00 (2.20).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and baseline values of outcomes; Group 1 Number missing: 2, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 20 (SD 9.2); n=11, Group 2: mean 16.25 (SD 12.53); n=11; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline mixed: 30.11 (7.67). Baseline other: 27.50 (8.25). Baseline strength: 29.75 (6.82).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and

baseline values of outcomes; Group 1 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1) Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 4.87 (SD 3.6); n=11, Group 2: mean 4.25 (SD 3.45); n=11; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Baseline other: 6.87 (2.75). Baseline strength: 8.00 (2.20). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, weight, BMI, gender and baseline values of outcomes; Group 1 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1); Group 2 Number missing: 3, Reason: Overall reasons: injury/illness unrelated to study (4), no-show for follow-up (1), joined an exercise program (1), caring for ill family member (1), out of state emergency (1), and other rheumatic disease diagnosed during study (1)	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Rogind 1998[382]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=23)
Countries and setting	Conducted in Denmark; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 year follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fulfilling the American College of Rheumatology criteria of osteoarthritis of the knee and the radiograph of the knee had to be rated at least 3 on the Kellgren scale
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People had to fulfill the American College of Rheumatology criteria for osteoarthritis of the knee they appointed as the most affected knee, and the radiograph of this knee had to be rated at least 3 on the Kellgren Lawrence scale. All radiographic evaluations of the knee changes were weight-bearing, performed with the person standing.
Exclusion criteria	Rheumatoid arthritis or inflammatory joint disease; knee arthroplasty or planned knee arthroplasty in the study period; intra-articular steroid injection within 2 weeks of the screening visit; medical or surgical condition contraindicating training during the intervention period; malaignment of the knees (varus/valgus) larger than 15 degrees; osteoarthritis of the hip; recent (3 months) fracture of upper or lower extremity; lack of understanding of the study (dementia, language problems); neurologic illness (stroke, polyneuropathy); and abuse of drugs or alcohol
Recruitment/selection of patients	People were recruited from the outpatient clinic
Age, gender and ethnicity	Age - Mean (SD): 71.2 (7.4). Gender (M:F): 2:21. Ethnicity: Not started
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: At least grade 3 on the Kellgren Lawrence scale Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=12) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Training focused on general fitness, balance, coordination, stretching, and lower extremity muscle strength, including a daily home

	exercise program. People were trained by physiotherapists and then instructed to continue the training at home apart from 2 days per week when they attended a training session in the project. They were also allowed to take a 1 day break per week. This was continued for 3 months.. Duration 3 months. Concurrent medication/care: As far as possible the medication was kept constant, apart from small changes in mild analgesics (paracetamol). No intra-articular or periarticular injections were given during the entire study period. . Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Coordination, flexibility, strength). (n=13) Intervention 2: No treatment. No exercise intervention. Duration 3 months. Concurrent medication/care: As far as possible the medication was kept constant, apart from small changes in mild analgesics (paracetamol). No intra-articular or periarticular injections were given during the entire study period. . Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by grants from Helsefonden and Kommunehospitalets Jubilaeumsfond)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Pain weight bearing (11 point NRS) at 3 months; Group 1: mean 4 (SD 2.2); n=12, Group 2: mean 6 (SD 2.3); n=13; NRS 0-10 Top=High is poor outcome; Comments: Reports means and 95% confidence intervals. Reported exercise: 4.0 (3.0-5.5). Reported control: 6.0 (4.8-7.3). Baseline exercise: 7.0 (4.5-7.5). Baseline control: 5.0 (5.0-6.5).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcome different at baseline (worse for intervention group than control group); Group 1 Number missing: 1, Reason: Reports overall causes of withdrawal. 1 hip fracture, 1 died of cancer; Group 2 Number missing: 1, Reason: Reports overall causes of withdrawal. 1 hip fracture, 1 died of cancer

Protocol outcome 2: Pain at > 3 months
- Actual outcome: Pain weight bearing (11 point NRS) at 12 months; Group 1: mean 4 (SD 2.7); n=12, Group 2: mean 7 (SD 3.2); n=13; NRS 0-10 Top=High is poor outcome; Comments: Reports means and 95% confidence intervals. Reported exercise: 4.0 (3.0-6.0). Reported control: 7.0 (4.0-7.5). Baseline exercise: 7.0 (4.5-7.5). Baseline control: 5.0 (5.0-6.5).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Outcome different at baseline (worse for intervention group than control group); Group 1 Number missing: 1, Reason: Reports overall causes of withdrawal. 1 hip fracture, 1 died of cancer; Group 2 Number missing: 1, Reason: Reports overall causes of withdrawal. 1 hip fracture, 1 died of cancer	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Rosedale 2014[384]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=180)
Countries and setting	Conducted in Canada; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with knee pain for greater than 4 months and radiologically confirmed diagnosis of knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable:
Inclusion criteria	All people were from waiting lists of 5 orthopaedic surgeons specialising in hip and knee joint replacement. People were required to have had knee pain for greater than 4 months and to have been referred to the orthopaedic clinic with a radiologically confirmed diagnosis of knee osteoarthritis
Exclusion criteria	Unable to attend exercise-based physiotherapy 2 to 3 times per week over a 2 week period; had neurological conditions affecting the lower extremitieis; were unable to understand written or spoken English; were unable to provide informed consent
Recruitment/selection of patients	People were recruited from outpatient orthopaedic clinics at a tertiary health care centre.
Age, gender and ethnicity	Age - Mean (SD): 65.3 (10.4). Gender (M:F): 69:89. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (Median number of comorbidities: 3). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=120) Intervention 1: Exercise - Supervised strength exercise. Specific strength based exercises with advise on aerobic exercise. People with mechanical derangement therapy classification of derangement were given specific end-range exercises in the direction in which the person responded. Exercises were prescribed as 10 repetitions every 2 to 3 hours. People classified as nonresponders had quadriceps strengthening exercises and advice on aerobic exercises. . Duration 3 months. Concurrent medication/care: No additional information. Indirectness: No

	indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=60) Intervention 2: No treatment. Waiting list control. Duration 3 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: KOOS function at 3 months; Group 1: mean 61 (SD 17); n=120, Group 2: mean 52 (SD 16); n=60; KOOS function 0-100 Top=High is good outcome; Comments: Baseline exercise: 56 (17). Baseline control: 51 (18). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, mass, height, BMI, sex, median comorbidities, and baseline values of outcomes; Group 1 Number missing: 20, Reason: 4 ill health, 8 confounding intervention, 4 personal reasons, 1 collection error, 3 unable to contact; Group 2 Number missing: 14, Reason: 7 confounding intervention, 2 diagnostic error, 3 enable to contact, 2 personal reasons Protocol outcome 2: Pain at </=3 months - Actual outcome: KOOS pain at 3 months; Group 1: mean 56 (SD 17); n=120, Group 2: mean 46 (SD 16); n=60; KOOS pain 0-100 Top=High is good outcome; Comments: Baseline exercise: 51 (17). Baseline control: 46 (17). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, mass, height, BMI, sex, median comorbidities, and baseline values of outcomes; Group 1 Number missing: 20, Reason: 4 ill health, 8 confounding intervention, 4 personal reasons, 1 collection error, 3 unable to contact; Group 2 Number missing: 14, Reason: 7 confounding intervention, 2 diagnostic error, 3 enable to contact, 2 personal reasons	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Saccomanno 2016[386]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=165)
Countries and setting	Conducted in Italy; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology diagnostic criteria with knee malalignment confirmed by radiographic examinations
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 18 years or older in good general health with knee osteoarthritis according to the American College of Rheumatology diagnostic criteria with knee malalignment (varus or valgus deformity) and osteoarthritis were confirmed by radiographic examinations in different views: weight-bearing anteroposterior, weight-bearing posteroanterior according to Rosenberg, standard lateral view and axial patella view at 30 degrees of flexion. Radiographic evidence was graded according to the Kellgren and Lawrence classification for the tibio-femoral osteoarthritis and according to Iwano et al. for the patello-femoral osteoarthritis.
Exclusion criteria	People with inability or unwillingness to sign informed consent; intra-articular injections with steroids or hyaluronic acid in prior 6 months; physiotherapy for knee problems in prior 6 months; congenital or acquired inflammatory or neurological (systemic or local) diseases involving the knee; chronic treatment with steroids or NSAIDs and cognitive or psychiatric disorders
Recruitment/selection of patients	Recruited among people referred for knee pain to the outpatient clinics of the Orthopaedic Institute of the Department of Geriatrics, Neuroscience and Orthopaedics, "Agostino Gemelli" University Hospital at the Catholic University of Rome, Italy
Age, gender and ethnicity	Age - Mean (SD): 61.8 (11.2). Gender (M:F): 44:113. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis

Extra comments	Severity: Kellgren Lawrence grade 1-3, median grade 1 Duration of symptoms (median [IQR]): between 24-36 (10-80).
Indirectness of population	No indirectness
Interventions	(n=55) Intervention 1: Pharmacological treatment - Intra-articular hyaluronic acid. Three intra-articular injections (one injection every 2 weeks) of high molecular weight hyaluronic acid (Orthovisc 2mL, 15mg/mL).. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Intrarticular treatment 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=55) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Rehabilitation exercises with a detailed programme of exercises for a total of 20 treatment sessions in a month (5 sessions per week). This included isometric and isotonic exercises, stretching and proprioceptive exercises. Duration 1 month. Concurrent medication/care: People were asked to refrain from any additional pharmacological or physical treatment of pain management. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, proprioception, stretching). (n=55) Intervention 3: Other. Exercise and intra-articular hyaluronic acid injections. Duration 6 weeks. Concurrent medication/care: People were asked to refrain from any additional pharmacological or physical treatment of pain management. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, proprioception, stretching). Comments: This group was not included in the analysis as they did not fulfill the inclusion criteria of this review
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus INTRA-ARTICULAR HYALURONIC ACID

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 3 months; Group 1: mean 596.5 (SD 298.9); n=51, Group 2: mean 685.7 (SD 360); n=53; WOMAC function 0-1800 Top=High is poor outcome; Comments: Baseline exercise: 706.9 (254). Baseline HA: 842.4 (384.9).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, timing of symptoms, previous surgery, compartment involved, Kellgren-Lawrence grade, Iwano stage and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 missed follow up; Group 2 Number missing: 6, Reason: 6 missed follow up Protocol outcome 2: Physical function at > 3 months - Actual outcome: WOMAC function at 6 months; Group 1: mean 618.5 (SD 310.4); n=51, Group 2: mean 691.4 (SD 363.8); n=53; WOMAC function 0-1800 Top=High is poor outcome; Comments: Baseline exercise: 706.9 (254). Baseline HA: 842.4 (384.9). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, timing of symptoms, previous surgery, compartment involved, Kellgren-Lawrence grade, Iwano stage and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 missed follow up; Group 2 Number missing: 5, Reason: 5 missed follow up Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain at 3 months; Group 1: mean 154.6 (SD 92); n=51, Group 2: mean 177.7 (SD 100.5); n=53; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline exercise: 216 (97.5). Baseline HA: 241.2 (101.9). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, timing of symptoms, previous surgery, compartment involved, Kellgren-Lawrence grade, Iwano stage and baseline values of outcomes; Group 1 Number missing: 6, Reason: 6 missed follow up; Group 2 Number missing: 6, Reason: 6 missed follow up Protocol outcome 4: Pain at > 3 months - Actual outcome: WOMAC pain at 6 months; Group 1: mean 161.6 (SD 90.2); n=51, Group 2: mean 181.5 (SD 98); n=53; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline exercise: 216 (97.5). Baseline HA: 241.2 (101.9). Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, timing of symptoms, previous surgery, compartment involved, Kellgren-Lawrence grade, Iwano stage and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 missed follow up; Group 2 Number missing: 5, Reason: 5 missed follow up Protocol outcome 5: Serious adverse events at > 3 months - Actual outcome: Adverse events at 6 months; Group 1: 0/55, Group 2: 0/55 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, BMI, timing of symptoms, previous surgery, compartment involved, Kellgren-Lawrence grade, Iwano stage and baseline values of outcomes; Group 1 Number missing: 5, Reason: 5 missed follow up; Group 2 Number missing: 5, Reason: 5 missed follow up	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months;

	Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	Salacinski 2012[387]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=37)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Mild-to-moderate osteoarthritis of the knee with grades 1-3 Kellgren Lawrence changes on radiography
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Men and women with mild to moderate knee osteoarthritis and Kellgren Lawrence grades 1-3 changes on radiography who reported knee pain on most days of the previous month. They also required the people to have at least 90 degrees of knee range of motion, stable baseline blood pressure according to the American College of Sports Medicine guidelines for exercise, and no knee swelling
Exclusion criteria	Severe patellofemoral pain that would not allow participation in the stationary cycling regimen; injection of viscosupplements in the knee within the previous 3 months; if they had any medical condition that would prohibit them from safely participating in an aerobic exercise program of moderate intensity
Recruitment/selection of patients	People were recruited by newspaper advertisements, posters and from physician practices
Age, gender and ethnicity	Age - Mean (SD): 57.7 (9.8). Gender (M:F): 10:27. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Mild-to-moderate, Kellgren Lawrence grades 1-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=19) Intervention 1: Exercise - Supervised aerobic exercise . Facility based cycling exrcises. All instructors were Spinning instructors certified by Mad Dogg Athletics, Inc. Reduced intensity pedalling (adapted to the needs ofpeople with osteoarthritis). Conducted over 12 weeks with at least 2 supervised group sessions per week to maximise aerobic fitness while limiting direct knee joint stress. People wore heart rate

	monitors and were instructed to maintain and average of 70-75% of their maximal heart rate. The sessions were composed of warm up, aerobic loading and a cool down. The instructors progressively increased from 40 to 60 minutes and included light intensity warm up/cool down, alternating efforts of fast-cadence pedaling and simulated hill climbs, and stretching. People were advised not to leave the saddle or cycle in the standing position, to avoid aggravating their knee osteoarthritis symptoms.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=18) Intervention 2: No treatment. No exercise intervention. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Study funded by industry (This work was supported by grants from the PNC Bank Arthritis Research Fund (Pittsburg, PA) and Mad Dogg Athletics, Inc.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS knee-related quality of life at 12 weeks; Group 1: mean 10.6 (SD 13.2); n=13, Group 2: mean 3.8 (SD 21.7); n=15; KOOS quality of life 0-100 Top=High is good outcome; Comments: Reports change scores (95% confidence intervals). Reported exercise: 10.6 (3.4, 17.8). Reported control: 3.8 (-7.2, 14.8). Baseline exercise: 72.7 (12.6). Baseline control: 67.8 (18.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, VO2max, muscle strength, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 2 discontinued due to knee pain, 1 employment conflict, 1 wrist pain, 1 nonrelated medical, 1 lost to follow-up; Group 2 Number missing: 3, Reason: 3 wanted intervention

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS function in daily living at 12 weeks; Group 1: mean 11.9 (SD 22.1); n=13, Group 2: mean 0.8 (SD 14.2); n=15; KOOS function 0-100 Top=High is good outcome; Comments: Reports change scores (95% confidence intervals). Reported exercise: 11.9 (-0.1, 23.9). Reported control: 0.8 (-6.4, 8.0). Baseline exercise: 72.3 (17.9). Baseline control: 70.3 (15.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, VO2max, muscle strength, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 2 discontinued due to knee pain, 1 employment conflict, 1 wrist pain, 1 nonrelated medical, 1 lost to follow-up; Group 2 Number missing: 3, Reason: 3 wanted intervention

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS pain at 12 weeks; Group 1: mean 12.4 (SD 14); n=13, Group 2: mean -0.9 (SD 13.8); n=15; KOOS pain 0-100 Top=High is good outcome; Comments: Reports change scores (95% confidence intervals). Reported exercise: 12.4 (4.8, 20.0). Reported control: -0.9 (-6.9, 7.1). Baseline exercise: 63.9 (13.8). Baseline control: 64.8 (16.0).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, VO2max, muscle strength, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 2 discontinued due to knee pain, 1 employment conflict, 1 wrist pain, 1 nonrelated medical, 1 lost to follow-up; Group 2 Number missing: 3, Reason: 3 wanted intervention

Protocol outcome 4: Serious adverse events at </=3 months
- Actual outcome: Knee pain and wrist pain at 12 weeks; Group 1: 3/19, Group 2: 0/18; Comments: Exercise: 1 wrist pain, 2 knee pain
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports gender, age, BMI, VO2max, muscle strength, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 discontinued due to employment conflict, 1 nonrelated medical, 1 lost to follow-up; Group 2 Number missing: 3, Reason: 3 wanted intervention

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at > 3 months

Study	Salli 2010[388]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=100)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 20 weeks (treatment for 8 weeks only)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with clinically and radiologically diagnosed osteoarthritis in both knees according to the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with clinically and radiologically diagnosed osteoarthritis in both knees according to the American College of Rheumatology criteria who led sedentary lifestyles and had participated in no regular exercise programs but with no contraindication to exercise
Exclusion criteria	People with severe knee trauma; secondary osteoarthritis or ligament damage; who had undergone any orthopaedic intervention or intra-articular knee injection in the last 6 months; with lumbar and hip pathologies or knee joint defomities; with findings of inflammation such as effusion of the knees and increase of temperature; and those with comorbidities
Recruitment/selection of patients	People admitted to their clinic
Age, gender and ethnicity	Age - Mean (SD): 57.06 (7.31). Gender (M:F): 13:58. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Low morbidity score (As people with comorbidities were excluded). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=47) Intervention 1: Exercise - Supervised strength exercise. Concentric-eccentric type isokinetic exercises or isometric exercises. Performed 3 days a week for 8 weeks.. Duration 8 weeks (then no treatment for 12 weeks). Concurrent medication/care: People in all groups received 500mg paracetamol tablets as required, up to 3 grams per day.. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable Comments: These two groups were combined together due to class effect as agreed in the protocol (n=24) Intervention 2: No treatment. No exercise intervention. Duration 20 weeks. Concurrent medication/care: People in all groups received 500mg paracetamol tablets as required, up to 3 grams per day.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical component at 8 weeks; Group 1: mean 59.2 (SD 16.3); n=47, Group 2: mean 38.4 (SD 9.5); n=24; SF-36 physical component 0-100 Top=High is good outcome; Comments: Reported concentric-eccentric: 66.7 (15.4). Reported isometric: 52.1 (13.8). Baseline concentric-eccentric: 41.1 (16.7). Baseline isometric: 39.2 (15.3). Baseline control: 37.2 (10.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.
- Actual outcome: SF-36 mental component at 8 weeks; Group 1: mean 68.3 (SD 12.4); n=47, Group 2: mean 50.9 (SD 12.5); n=24; SF-36 mental component 0-100 Top=High is good outcome; Comments: Reported concentric-eccentric: 70.1 (13.1). Reported isometric: 66.5 (11.3). Baseline concentric-eccentric: 53.8 (17.4). Baseline isometric: 50.9 (14.2). Baseline control: 46.6 (13.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: SF-36 physical component at 20 weeks; Group 1: mean 58.4 (SD 15.8); n=47, Group 2: mean 40.8 (SD 10.9); n=24; SF-36 physical component 0-100 Top=High is good outcome; Comments: Reported concentric-eccentric: 65.2 (15.6). Reported isometric: 51.9 (13.1). Baseline concentric-

eccentric: 41.1 (16.7). Baseline isometric: 39.2 (15.3). Baseline control: 37.2 (10.4).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 2, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms. However, all 25 were included in the 20 week evaluation for the concentric-eccentric group, with 1 more dropping out in the isometric group.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.
- Actual outcome: SF-36 mental component at 20 weeks; Group 1: mean 65.5 (SD 12.2); n=47, Group 2: mean 51.2 (SD 12.8); n=24; SF-36 mental component 0-100 Top=High is good outcome; Comments: Reported concentric-eccentric: 68.8 (12.7). Reported isometric: 62.4 (10.9). Baseline concentric-eccentric: 53.8 (17.4). Baseline isometric: 50.9 (14.2). Baseline control: 46.6 (13.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 2, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms. However, all 25 were included in the 20 week evaluation for the concentric-eccentric group, with 1 more dropping out in the isometric group.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 8 weeks; Group 1: mean 20.8 (SD 10.2); n=47, Group 2: mean 32.6 (SD 11.6); n=24; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reported concentric-eccentric: 15.5 (9.4). Reported isometric: 25.8 (8.3). Baseline concentric-eccentric: 29.4 (10.8). Baseline isometric: 35.2 (8.5). Baseline control: 33.5 (12.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: WOMAC physical function at 20 weeks; Group 1: mean 19.6 (SD 10.5); n=47, Group 2: mean 32.7 (SD 11.3); n=24; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reported concentric-eccentric: 13.3 (8.7). Reported isometric: 25.7 (8.3). Baseline concentric-eccentric: 29.4 (10.8). Baseline isometric: 35.2 (8.5). Baseline control: 33.5 (12.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 2, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms.

However, all 25 were included in the 20 week evaluation for the concentric-eccentric group, with 1 more dropping out in the isometric group.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm. Protocol outcome 5: Pain at </=3 months - Actual outcome: VAS motion at 8 weeks; Group 1: mean 3.4 (SD 1.9); n=47, Group 2: mean 6.5 (SD 1.8); n=24; VAS motion 0-10 Top=High is poor outcome; Comments: Reported concentric-eccentric: 2.8 (1.7). Reported isometric: 3.9 (1.9). Baseline concentric-eccentric: 7.1 (1.2). Baseline isometric: 7.5 (1.4). Baseline control: 7.1 (1.3). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 3, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm. Protocol outcome 6: Pain at > 3 months - Actual outcome: VAS motion at 20 weeks; Group 1: mean 3.5 (SD 1.9); n=47, Group 2: mean 6.3 (SD 1.5); n=24; VAS motion 0-10 Top=High is poor outcome; Comments: Reported concentric-eccentric: 3.1 (1.7). Reported isometric: 3.9 (1.9). Baseline concentric-eccentric: 7.1 (1.2). Baseline isometric: 7.5 (1.4). Baseline control: 7.1 (1.3). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, BMI and baseline values of outcomes; Group 1 Number missing: 2, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 3 dropped out from these arms. However, all 25 were included in the 20 week evaluation for the concentric-eccentric group, with 1 more dropping out in the isometric group.; Group 2 Number missing: 1, Reason: Stated it randomised 100 people at first (no information on which groups they went into), but only 25 in each group underwent the pre-treatment evaluation. This means 25 people were unaccounted for at first anyway. After this 1 dropped out from this arm.	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Samut 2015[390]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Kellgren Lawrence grade 2-3 knee osteoarthritis fulfilling the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Postmenopausal women and men aged over 50 years with a diagnosis of knee osteoarthritis according to the American College of Rheumatology criteria with Kellgren Lawrence grade 2-3 knee osteoarthritis and having a sedentary lifestyle (less than 60 minutes of moderate- to high intensity activity per week)
Exclusion criteria	Cooperation problems; depression; cognitive impairment; neurologic impairment/disease; orthopedic problems; inflammatory arthritis; regular exercise habits; having received physical therapy or intra-articular injection in the last 3 months; cardiovascular problems; end-stage disease; immunosuppressive drug usage; having an infection or inflammatory condition; pregnancy; and malignant disease.
Recruitment/selection of patients	People attending the department of physical medicine and rehabilitation
Age, gender and ethnicity	Age - Mean (SD): 60.4 (7.8). Gender (M:F): 4:38. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Supervised strength exercise. Isokinetic exercise performed 3 days/week for 6 weeks in Biodex isokientic system. This included a 5 min warm up period on a treadmill followed by 5 concentric flexion and extension at angular velocities of 60 degrees/s, 90 degrees/s, 120 degrees/s, and 180 degrees/s. One set of contraction was performed in the first session which was increased to 6 sets by 1 increment in each of the following sessions and continued as 6 sets until the

	send of the study. 20s of rest was allowed between sets and 2 mins of rest was allowed between legs. No other physical therapy modality was applied to the people.. Duration 6 weeks. Concurrent medication/care: All groups were allowed to take paracetamol whenever needed. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=14) Intervention 2: Exercise - Supervised aerobic exercise . Aerobic exercise performed 3 days/week for 6 weeks on a treadmill. After 5 min of warm-up on the treadmill, exercise intensity was adjusted for 65-70% of age-related heart rate for the first 4 weeks and 70-75% for the next 2 weeks and continued that way for the rest of the study. Exercise sessions were ended with a 5-min cool-down period. No other physical therapy modality was applied to the subjects.. Duration 6 weeks. Concurrent medication/care: All groups were allowed to take paracetamol whenever needed. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=13) Intervention 3: No treatment. No exercise intervention. Duration 6 weeks. Concurrent medication/care: All groups were allowed to take paracetamol whenever needed. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This study was supported in part by the Hacettepe University Scientific Research Center)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus SUPERVISED AEROBIC EXERCISE

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 6 weeks; Group 1: mean 16.08 (SD 11.27); n=15, Group 2: mean 14.57 (SD 11.74); n=14; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline strength: 33.85 (7.12). Baseline aerobic: 26.29 (12.09).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 2, Reason: 2 dropped out due to lack of effect; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months

- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 4 (SD 3); n=15, Group 2: mean 3.29 (SD 2.4); n=14; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline strength: 9.15 (3.78). Baseline aerobic: 7.00 (3.16).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 2, Reason: 2 dropped out due to lack of effect; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 6 weeks; Group 1: mean 16.08 (SD 11.27); n=15, Group 2: mean 29.92 (SD 11.25); n=13; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline strength: 33.85 (7.12). Baseline control: 30.00 (10.73).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 2, Reason: 2 dropped out due to lack of effect; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 4 (SD 3); n=15, Group 2: mean 7.31 (SD 2.84); n=13; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline strength: 9.15 (3.78). Baseline control: 7.92 (3.01).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 2, Reason: 2 dropped out due to lack of effect; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 6 weeks; Group 1: mean 14.57 (SD 11.74); n=14, Group 2: mean 29.92 (SD 11.25); n=13; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline aerobic: 26.29 (12.09). Baseline control: 30.00 (10.73).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean 3.29 (SD 2.4); n=14, Group 2: mean 7.31 (SD 2.84); n=13; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline aerobic: 7.00 (3.16). Baseline control: 7.92 (3.01).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: WOMAC pain and function subscales are different for the aerobic and strength exercise groups; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Sayers 2012[393]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=45)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology clinical classification
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Meeting criteria of the American College of Rheumatology clinical classification of knee osteoarthritis, which consisted of knee pain and inclusions of 3 of the following 6 criteria: age >50 years, crepitus on active motion, less than 30 minutes of stiffness upon waking in the morning, bony tenderness, bony enlargement, and no palpable warmth of synovium. Also required to have evidence of pain or function deficit on WOMAC (a minimum of 1 response of 'moderate' or 2 responses of 'minimal' for pain, 2 responses of 'moderate' or 4 responses of 'mild' for physical function).
Exclusion criteria	History of heart disease; severe visual impairment; presence of neurologic disease; pulmonary disease requiring the use of oxygen; uncontrolled hypertension; hip fracture or lower extremity joint replacement in the past 6 months; and current participation in structured exercise
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 67.1 (7.3). Gender (M:F): 8:25. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence mean grade 1-2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise - Supervised strength exercise. Power therapy (including low intensity and explosive high intensity) performed 3 times per week for 12 weeks. High intensity performed 3 sets of 12-14 repetitions at 40% of 1RM while low intensity performed 3 sets of 8-10 repetitions at 80% of 1RM. The high intensity

	group performed the concentric phase with an explosive movement at high speed, paused for 1 second, and performed the eccentric portion over 2 seconds. The low intensity performed each action at slow velocity (2 seconds for the concentric phase, pause for 1 second, then 2 seconds for the eccentric phase).. Duration 12 weeks. Concurrent medication/care: The exercises started with 12 stretches including the back, trunk and lower extremity stretches. Each stretch was initiated and held for 30 seconds by a trained physical therapy study research assistant. Following the stretching protocol, a 5 minute warm up on a cycle ergometer was performed before starting the exercise.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable Comments: The two groups were combined due to class effect as agreed in the protocol (n=15) Intervention 2: No treatment. Stretches only. Duration 12 weeks. Concurrent medication/care: The exercises started with 12 stretches including the back, trunk and lower extremity stretches. Each stretch was initiated and held for 30 seconds by a trained physical therapy study research assistant. Following the stretching protocol, a 5 minute warm up on a cycle ergometer was performed before starting the exercise.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by the American College of Rheumatology and the Arthritis Foundation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 29.7 (SD 10.2); n=22, Group 2: mean 34.8 (SD 13.9); n=11; WOMAC function 0-68 Top=High is poor outcome; Comments: Reported high intensity: 26.5 (6.1). Reported low intensity: 33.5 (12.6). Baseline high intensity: 41.4 (9.7). Baseline low intensity: 41.9 (9.8). Baseline control: 39.5 (11.0).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, height, weight, BMI, geriatric depression scale, mini-mental state examination, number of prescribed medications, Kellgren Lawrence scores for each knee and baseline values of outcomes; Group 1 Number missing: 11, Reason: 8 withdrew during baseline. 3 withdrew during treatment but were included in the final analyses.; Group 2 Number missing: 5, Reason: 4 withdrew during baseline. 1 withdrew during treatment but was included in the final analyses.

Protocol outcome 2: Pain at </=3 months - Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 9.8 (SD 3.1); n=22, Group 2: mean 10.2 (SD 2.5); n=11; WOMAC pain subscale 0-20 Top=High is poor outcome; Comments: Reported high intensity: 9.3 (3.2). Reported low intensity: 10.4 (2.8). Baseline high intensity: 11.5 (2.8). Baseline low intensity: 12.2 (3.4). Baseline control: 11.7 (2.6). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, height, weight, BMI, geriatric depression scale, mini-mental state examination, number of prescribed medications, Kellgren Lawrence scores for each knee and baseline values of outcomes; Group 1 Number missing: 11, Reason: 8 withdrew during baseline. 3 withdrew during treatment but were included in the final analyses.; Group 2 Number missing: 5, Reason: 4 withdrew during baseline. 1 withdrew during treatment but was included in the final analyses.	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Sedaghatnezhad 2021[400]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in Iran; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks (end of treatment) and 20 days after the end of treatment
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology (including radiographic findings)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People referred to the private physical therapy centre with a diagnosis of knee osteoarthritis according to the American College of Rheumatology criteria; grade II or III knee osteoarthritis based on the Kellgren and Lawrence scale on the knee x-ray; aged between 40 and 65 years; body mass index <30kg/m² to avoid different biomechanical patterns of knee joint loading during gait;

	a pain of at least 30 on an 100mm visual analogue scale during the last week of participation and having knee extension limitation of at least 2 degrees.
Exclusion criteria	If they had received steroid injections in the past 3 months and/or physical therapy for knee problems in the past 6 months; if they could not walk unassisted; if they had history of fractures, dislocation, knee joint surgeries, knee ligamentous injury, other types of arthritis, heel spur that prevented the patient from uphill walking, discopathy or trauma that affect lower extremities functions; sudden onset of pain.
Recruitment/selection of patients	A non-probability and convenience sampling method. People referred to a private physical therapy centre with a diagnosis of knee osteoarthritis between December 2017 and September 2018.
Age, gender and ethnicity	Age - Mean (SD): 56.7 (8.0). Gender (M:F): 5:25. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade II-III Duration of symptoms: Not stated/unclear. IRCT20171115034920N1.
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Walking on an uphill treadmill. Initially for familiarisation: walking for two minutes at 1.1 m/s speed and +8 degree slope on the treadmill. After this, they did a warm up at 1.1 m/s speed at 0 degrees then increasing this to +8 degrees after 2 minutes. This then continued for a total of 30 minutes (15 minutes before physical therapy, 15 minutes after physical therapy). Physical therapy (available to all participants) included a strengthening exercise program (see concomitant treatment).. Duration 2 weeks. Concurrent medication/care: Everyone received the following: a 201B ultrasound used for continuous ultrasound therapy (using a 1MHz head set to 1W/cm² applied for 6 minutes - 3 minutes on the anteromedial and 3 on the posterior of the knee); a transcutaneous nerve stimulation unit giving therapy for 20 minutes at 100Hz for a pulse duration of 50 microseconds; two hot packs on the anterior and posterior aspects of the knees. This was followed by a muscle strengthening program performed individually in two sets, repeated from 10 up to 30 times between the first and fifth sessions, and then 30 for the remaining five sessions. Exercises included supine quadriceps setting, side lying hip abduction and standing heel raising on two legs.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strengthening and aerobic). (n=15) Intervention 2: Exercise - Supervised strength exercise. Strengthening exercise component (and other physical therapy available to all participants) only.. Duration 2 weeks. Concurrent medication/care: Everyone received the following: a 201B ultrasound used for continuous ultrasound therapy (using a 1MHz head set to 1W/cm² applied for 6 minutes - 3 minutes on the anteromedial and 3 on the posterior of the knee); a transcutaneous nerve stimulation unit giving therapy for 20 minutes at 100Hz for a pulse duration of 50 microseconds; two hot packs on the anterior and posterior aspects of the knees. This was followed by a muscle strengthening program performed individually in two sets, repeated from 10 up to 30 times between the first and fifth sessions, and then 30 for the remaining five sessions. Exercises included supine quadriceps setting, side lying hip

	abduction and standing heel raising on two legs.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	Academic or government funding (Financially supported by the Deputy of Research affair of the University of Social Welfare and Rehabilitation Sciences)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Visual analogue scale at 34 days; Group 1: mean 13.67 (SD 10.06); n=15, Group 2: mean 25.89 (SD 16.69); n=15; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline supervised mixed modality: 42.67 (10.46). Baseline supervised strength: 39 (14.81).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, weight, height, BMI, disease severity, gender and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Segal 2015[401]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=48)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: Intervention for 3 months, follow up for 12 months in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Symptomatic knee osteoarthritis, defined by a definite osteophyte or joint space narrowing in either tibiofemoral compartment or posteroanterior knee radiographs and an affirmative response to "Have you had pain or stiffness in one or both knees on most of the past 30 days" on both the telephone screen and screening visit and mobility disability {LLFDI advanced lower limb function score below 32 points})
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	men and women age 60 years or older with symptomatic knee osteoarthritis. All people were able to walk without an assistive device and ascend at least two stairs.
Exclusion criteria	Conditions other than knee osteoarthritis, which could affect walking, were exclusionary (e.g. amputation, severe back pain, severe peripheral vascular or heart disease and neurological or developmental disease including multiple sclerosis, Parkinson's disease, myositis, rickets, or lower limb musculoskeletal surgery in the past 6 months). In addition, participants who had undergone corticosteroid injection into either a peripheral joint or into the spine in the past 3 months (which could threaten internal validity of assessing the independent effect of the intervention) or who anticipated inability to return for follow-up; medical conditions that preclude safe participation in the study protocol, including but not limited to acute or terminal illness or unstable cardiovascular condition; report of medical condition that may impair ability to participate including but not limited to pulmonary disease requiring the use of supplemental oxygen; inability or unwillingness to comply with the study protocol or be randomize; inability to obtain written clearance for participation int he study by a physician; concurrent participation in another observation or interventional research study; current consumption of more than 14 alcoholic drinks per week; judgment of the principal investigator that participation would endanger the safety of an individual.
Recruitment/selection of patients	People recruited at the University of Iowa. Achieved through targeted mailings to people with iCD-9 codes relevant to knee osteoarthritis (715.96, 715.16, 715.36) while

	excluding those with a code indicating lower limb surgery in the past 6 months. Orthopaedics, rheumatology and internal medicine clinics within a 40-mile radius also were targeted with fliers and mailings. In addition, study notices were posted in local senior centers and assisted living centers.
Age, gender and ethnicity	Age - Mean (SD): 69.3 (7.0). Gender (M:F): 16:32. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-4, median grade 3 Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=36) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Gait training intervention completed in 24 biweekly 45 minute sessions directed by a physical therapist that was comprised of guided strategies to optimize knee movements during treadmill walking, using computerized motion analysis with visual biofeedback. Following the initial 3 month intervention, people were encouraged to continue the intervention at home through scripted telehpone-based motivational interviewing and a tracking component.. Duration 3 months of intervention, 12 months follow up in total. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Neuromodulatory (n=22) Intervention 2: No treatment. Usual care, which could include: a yearly visit with their physician, use of pain medications for knee symptoms, knee surgery and/or physical therapy (available to both groups). They had additional telephone follow up to match the follow up of the intervention group.. Duration 12 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (This research was supported by a Paul B. Beeson Career Development Award in Aging Research (K23AG030945))
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT	

Protocol outcome 1: Pain at </=3 months - Actual outcome: KOOS pain at 3 months; Group 1: mean 8.2 (SD 14.7); n=29, Group 2: mean 1.1 (SD 14.7); n=19; KOOS pain 0-100 Top=High is good outcome; Comments: Reported change scores and 95% confidence intervals. Reported exercise: 8.2 (2.8, 13.5). Reported control: 1.1 (-5.5, 7.7). Baseline exercise: 62.7 (10.8). Baseline control: 59.8 (13.1). Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, CESD score, PASE score, osteoarthritis severity, LLFDI score, KOOS pain, KOOS symptoms, LDCW, stair -climb and chair stand time; Group 1 Number missing: 7, Reason: 3 did not receive the intervention (lack of time/driving distance, dementia). 4 discontinued intervention (family emergency, epidural steroid use, driving distance); Group 2 Number missing: 3, Reason: 1 did not receive the intervention (lack of knee osteoarthritis). 1 lost to follow up, 1 discontinued intervention (family emergency) Protocol outcome 2: Pain at > 3 months - Actual outcome: KOOS pain at 12 months; Group 1: mean 10.1 (SD 14.2); n=24, Group 2: mean 2.8 (SD 15.7); n=18; KOOS pain 0-100 Top=High is good outcome; Comments: Reported change scores and 95% confidence intervals. Reported exercise: 10.1 (4.4, 15.8). Reported control: 2.8 (-4.4, 10.1). Baseline exercise: 62.7 (10.8). Baseline control: 59.8 (13.1). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reports age, sex, CESD score, PASE score, osteoarthritis severity, LLFDI score, KOOS pain, KOOS symptoms, LDCW, stair -climb and chair stand time; Group 1 Number missing: 12, Reason: 3 did not receive the intervention (lack of time/driving distance, dementia). 9 discontinued intervention (family emergency, epidural steroid use, driving distance, distance/time, death, foot, hip or knee surgery); Group 2 Number missing: 4, Reason: 1 did not receive the intervention (lack of knee osteoarthritis)., 1 lost to follow up, 2 discontinued intervention (family emergency, terminal illness)	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Sekir 2005[403]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=22)
Countries and setting	Conducted in Turkey; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with bilateral complaints of knee osteoarthritis, who had grade 2 or 3 osteoarthritis, as judged by the criteria of the American College of Rheumatology based on weight-bearing radiographs
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with bilateral complaints of knee osteoarthritis, who had grade 2 or 3 osteoarthritis, as judged by the criteria of the American College of Rheumatology based on weight-bearing radiographsPeople with bilateral complaints of knee osteoarthritis, who had grade 2 or 3 osteoarthritis, as judged by the criteria of the American College of Rheumatology based on weight-bearing radiographs.
Exclusion criteria	None of the people had any neurological disorder (e,.g. Parkinson's, Alzheimer's) and/or a vestibular disorder; previous surgery on either knee,; symptomatic disease of the hip, ankle or foot; receving intra-articular steroid or hyaluronic acid injections in the previous 6 months; receiving physiotherapy treatment; knee cruciate ligament injury
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 60.4 (8.7). Gender (M:F): 6:16. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=12) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Multistation exercise program including balance and proprioception exercises. This including walking forward through 6 boxes, stair-up and down regular 3 steps staircase, leaning with heels off the floor with and without hand behind the back, one-legged stand, walking heel-to-toe, rising from a standard chair, and leg

	raises. Performed twice a week.. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Proprioception (n=10) Intervention 2: No treatment. No exercise intervention. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Pain at </=3 months
- Actual outcome: VAS total pain at 6 weeks; Group 1: mean 16.6 (SD 14.4); n=12, Group 2: mean 34.2 (SD 14.4); n=10; VAS total pain 0-70 (made of 7x0-10 subscale scores) Top=High is poor outcome; Comments: Reported means, IQRs and p-values. Reported exercise: 16.6 (6.0, 25.0), p-value = <0.01. Reported control: 34.2 (28.4, 42.9). Baseline exercise: 37.6 (27.1, 51.2). Baseline control: 40.8 (36.1, 47.0).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, body mass and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Shahine 2020[405]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=66)
Countries and setting	Conducted in Egypt; Setting: Outpatient follow up

Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosed with knee osteoarthritis for at least 1 year
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 60 years old and above; diagnosed with knee osteoarthritis lasting more than one year; agree for study participation; able to or have caregiver who can speak, read and write; available for telephone follow-up
Exclusion criteria	Older people using assistive devices during ambulation; older people with a history of traumatic hip, knee or ankle injury or surgery within the last year; older people who had cardiovascular disease; older people undergoing haemodialysis; older people who have osteoarthritis complications; older people with other types of osteoarthritis.
Recruitment/selection of patients	Rheumatology rehabilitation department outpatient clinic affiliated to Mansoura University Hospital between April 2019 and Augsut 2019.
Age, gender and ethnicity	Age - Mean (SD): 66.2 (5.5). Gender (M:F): 29:37. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: High morbidity score (No comorbidities = 19. One comorbidity = 30. Two comorbidities = 14. More than 3 = 3.). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of symptoms: At least one year. Between 1 year and 10+ years, median 5-10 years.
Indirectness of population	No indirectness
Interventions	(n=33) Intervention 1: Exercise - Unsupervised aerobic exercise . Routine care, educational sessions about pedometer self monitoring, aerobic weekly step count goals and weekly telephone follow up. People were given an individualised step count goal every week to gradually increase by 10% of baseline steps/day for weeks 2-12. People were taught to walk at a cadence of 100 steps/minute, which elicits a moderate, noticeable increase in depth and rate of breathing, while the person can talk with slight effort. . Duration 12 weeks. Concurrent medication/care: Everyone received a coloured Arabic booklet as a disease guide and a pedometer.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable (n=33) Intervention 2: No treatment. Usual routine care only. Duration 12 weeks. Concurrent medication/care: Everyone received a coloured Arabic booklet as a disease guide and a pedometer.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Partial funding from Mansoura University.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED AEROBIC EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 30 (SD 14.8); n=33, Group 2: mean 79.9 (SD 15.5); n=33; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline exercise: 80 (12.9). Baseline no treatment: 79 (15).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, social status, education, working before retirement, income source, living condition, duration of symptoms, knee effected, physical activity performance, therapeutic regimen, number of associated diseases and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 17.5 (SD 18.1); n=33, Group 2: mean 77 (SD 17.3); n=33; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline exercise: 74.9 (13.7). Baseline no treatment: 75.9 (16.7).
Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported sex, age, social status, education, working before retirement, income source, living condition, duration of symptoms, knee effected, physical activity performance, therapeutic regimen, number of associated diseases and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Silva 2008[408]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=64)
Countries and setting	Conducted in Brazil; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 18 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical and radiographic diagnosis of osteoarthritis of the knee according to the American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with osteoarthritis of the knee fulfilling the clincial and radiographic criteria of the American College of Rheumatology and knee pain ranging from 30-90mm on a VAS
Exclusion criteria	Neurological diseases of the lower limbs; symptomatic heart disease; symptomatic disease affecting the extremities other than osteoarthritis of the knee; symptomatic lung disease; severe systemic disease that could interfere with the assessment; psychiatric disorder; epilepsy; skin disease; inability to walk; people who received intra-articular injections of steroids in the preceding 3 months; those who had physical therapy intervention for their knee in the preceding 6 months or practiced regular physical activity (3 times a week or more) for more than 1 month
Recruitment/selection of patients	People selected from the Rheumatology Outpatient Clinics at Sao Paulo Hospital
Age, gender and ethnicity	Age - Mean (SD): 59 (6.8). Gender (M:F): 5:59. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=32) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Hydrotherapy including stretching, isometric strengthening, isotonic strengthening, and gait training exercises performed in groups of 5 to 8 people in 50 minute sessions 3 times a week for 18 weeks. Duration 18 weeks. Concurrent medication/care: People were instructed to take 50mg sodium diclofenac tablets as

	required, not surpassing a maximum dose of 150mg per day. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=32) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Land based therapy based on strengthening and gait training exercises including stretching, isometric strengthening, isotonic strengthening, and gait training exercises performed in groups of 5 to 8 people in 50 minute sessions 3 times a week for 18 weeks. Duration 18 weeks. Concurrent medication/care: People were instructed to take 50mg sodium diclofenac tablets as required, not surpassing a maximum dose of 150mg per day. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Strength and proprioception).
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED)

Protocol outcome 1: Pain at </=3 months
- Actual outcome: VAS at 9 weeks; Group 1: mean 37 (SD 18.1); n=32, Group 2: mean 38.4 (SD 27.5); n=32; VAS 0-100 Top=High is poor outcome; Comments: Baseline hydrotherapy: 61.9 (15.7). Baseline mixed: 68.2 (15.5).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body weight, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 31 of the 32 completed the protocol, 1 dropped out due to work; Group 2 Number missing: 6, Reason: 26 completed the study, with 6 drop outs - 3 due to allocation, 1 due to transportation problems, 1 had fibromyalgia, 1 left Sao Paulo due to personal problems

Protocol outcome 2: Pain at > 3 months
- Actual outcome: VAS at 18 weeks; Group 1: mean 26.7 (SD 23.1); n=32, Group 2: mean 37.3 (SD 27.5); n=32; VAS 0-100 Top=High is poor outcome; Comments: Baseline hydrotherapy: 61.9 (15.7). Baseline mixed: 68.2 (15.5).
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, height, body weight, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 31 of the 32 completed the protocol, 1 dropped out due to work; Group 2 Number missing: 6, Reason: 26 completed the study, with 6 drop outs - 3 due to allocation, 1 due to transportation problems, 1 had fibromyalgia, 1 left Sao Paulo due to

personal problems	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Song 2003[415] (Song 2007[414])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=72)
Countries and setting	Conducted in South Korea; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical and radiographic evidence of knee osteoarthritis according to the American College of Rheumatology criteria with a Kellgren Lawrence grade of at least 2
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 55 years or older; clinical and radiographic evidence of osteoarthritis
Exclusion criteria	Chronic disease of disability that would prevent completion of the program or survey communications such as ischaemic heart disease or cerebrovascular attack; participation in any regular exercise program during the previous year
Recruitment/selection of patients	People from an arthritis outpatient clinic of a university hospital were reviewed by their primary physician according to the inclusion criteria for the study
Age, gender and ethnicity	Age - Mean (SD): 63.7 (5.9). Gender (M:F): 0:72. Ethnicity: Not stated
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms (mean [SD]): 9.8 (7.2) years.
Indirectness of population	No indirectness
Interventions	(n=22) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Sun style tai chi exercise taught over 12 weeks. Consisted of warm-up exercise, 12 main movements and cool-down exercise. in this study the warm-up and cool-down exercises involved stretching and relaxing the head, neck, upper and lower body, and the whole body, and they were repeated 3-5 times, alternating sides where appropriate. People performed each exercise slowly and walked or moved at their own pace while simultaneously breathing in or out. One set of basic and advanced movements took about 2 minutes, with the people performing 10-15 sets of these at a session. Each person was given an instructional audiotape with the

	background music to practice the tai chi exercise at home.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Tai Chi). (n=21) Intervention 2: No treatment. No intervention. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported by the Korea Research Foundation (grant no. 2000-042-F00100), Seoul, Korea)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: Physical functioning (WOMAC) at 12 weeks; Group 1: mean -11.09 (SD 12); n=22, Group 2: mean -1.33 (SD 10.6); n=21; WOMAC physical functioning 0-68 Top=High is poor outcome; Comments: Baseline exercise: 37.59 (10.6). Baseline control: 37.95 (12.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Pain outcome is significantly different between the two groups. Otherwise similar.; Group 1 Number missing: -, Reason: Reported that there were originally 72 participants in the study and that 41% dropped out.; Group 2 Number missing: -, Reason: Reported that there were originally 72 participants in the study and that 41% dropped out.

Protocol outcome 2: Pain at </=3 months
- Actual outcome: Joint pain (WOMAC) at 12 weeks; Group 1: mean -2.45 (SD 3.9); n=22, Group 2: mean 0.61 (SD 5.1); n=21; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.91 (4.1). Baseline control: 8.90 (5.1).
Risk of bias: All domain - Very high, Selection - Very high, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Pain outcome is significantly different between the two groups; Overall rate reported only Group 1 Number missing: -, Reason: Reported that there were originally 72 participants in the study and that 41% dropped out.; Group 2 Number missing: -, Reason: Reported that there were originally 72 participants in the study and that 41% dropped out.

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Takacs 2017[430]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Canada; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 weeks + 1 week
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Individuals aged 50 to 80 years with radiographically confirmed tibiofemoral knee OA (Kellgren and Lawrence [KL] grade ≥2) and knee pain
Exclusion criteria	Inflammatory arthritic condition, history of knee or hip replacement, recent corticosteroid use, knee injections, or arthroscopic surgery (within the last 6mo); inability to ambulate without a gait aid; planning to start an exercise program within 3 months; or unable to attend 8 sessions at the university. Individuals with any neurologic, musculoskeletal, or other condition affecting their lower extremity movement ability, balance, or maximal strength were also excluded
Recruitment/selection of patients	Participants were recruited from a laboratory participant database and via advertisements in print media
Age, gender and ethnicity	Age - Mean (SD): Exercise group: 66.1 (8.7); control group: 67.1 (5.4). Gender (M:F): 8/32. Ethnicity: Not reported
Further population details	1. Age: Mixed age group 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren-Lawrence grade ≥2 Duration: not reported
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Dynamic balance training consisted of progressive exercise training over 3 phases, with exercises emphasizing dynamic balance control, eccentric lower limb muscle strength, and core stability. Exercises included sitting rotation, chair sit/squat, calf raise, side stepping, stepping pattern, standing rotation, step down, toe walking, lateral step up, stepping rotation, lunge, mini-hop, skate stepping and cone walking. Exercises were performed as 2to 3 sets of 8 to 12

	repetitions, and were individually progressed by an experienced kinesiologist. Exercises were progressed through the phases when the following conditions were met: (1) the participant was able to complete 3sets of 12 repetitions for each exercise, (2) the self-reported difficulty for all exercises dropped below 3 (out of 10), (3) the kinesiologist deemed the exercise had been mastered, and (4) knee pain was <6 (out of 10) during the performance of all exercises. Participants were asked to perform all exercises4 times per week, for a total of 40 exercise sessions over 10 weeks. Participants completed 6 supervised training sessions at the university (during weeks 1, 2,3, 5, 7, and 9) that were included in the total number of sessions for each week. All other training sessions were performed at home.. Duration 10 weeks. Concurrent medication/care: Co-interventions included prescription pain medication (n=1). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated / Unclear 3. Type of exercise: Other (Strength, balance). (n=20) Intervention 2: No treatment. The non-intervention group attended the same 2 testing sessions (baseline and follow-up, 12wk apart) as the training group, with no other visits to the university or contact with the study team. Participants were asked to maintain their usual level of activity and refrain from trying new treatment programs or medications. Participants were asked to record any changes to their usual activity routine and any new treatments or medications in a weekly log book.. Duration 10 weeks. Concurrent medication/care: Co interventions included physiotherapy (n=2); hydrotherapy (n=1), and exercise circuit training (n=1). Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: WOMAC physical function subscale at End of treatment (11 weeks); Group 1: mean 20 (SD 11); n=17, Group 2: mean 28 (SD 10); n=19; WOMAC 0-68 Top=High is poor outcome; Comments: Baseline scores: exercise group 30 (17); control group 30 (10) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in number of men/women in each group (5% men vs 35%); Group 1 Number missing: 3; Group 2 Number missing: 1	

Protocol outcome 2: Pain at </=3 months - Actual outcome: Pain at End of treatment (11 weeks); Group 1: mean 2.8 (SD 1.7); n=17, Group 2: mean 4.6 (SD 2.3); n=19; Numerical rating scale 0-10 Top=High is poor outcome; Comments: Baseline scores: exercise group 5 (1.8); control group 4.8 (2.2) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in number of men/women in each group (5% men vs 35%); Group 1 Number missing: 3; Group 2 Number missing: 1	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Teirlinck 2016[439]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=203)
Countries and setting	Conducted in Netherlands; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months + 12 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Fulfilling the clinical criteria for hip osteoarthritis of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable:
Inclusion criteria	Aged ≥45 years, and suffered from a new episode of non-traumatic hip complaints fulfilling the clinical criteria for hip OA of the American College of Rheumatology (ACR)
Exclusion criteria	Exercise therapy in the past 3 months; hip pain score <2 on an 11-point numeric rating scale (NRS: 0 = no pain); high level of physical function (score of <2 on the Algo functional Index); hip surgery or on waiting list; disabling co-morbidity (e.g., severe heart failure); insufficient comprehension of the Dutch language; mentally incapable of participation
Recruitment/selection of patients	Patient registries of GPs were searched for those who had visited in the past year for a new episode of hip complaints
Age, gender and ethnicity	Age - Mean (SD): 65.5 (9.2). Gender (M:F): 84:119. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: High morbidity score (High blood pressure = 82, Heart disease = 33, Lung disease = 17, Diabetes = 26, Rheumatoid arthritis = 6). 4. Site of osteoarthritis: Hip osteoarthritis
Extra comments	Severity Kellgren Lawrence grade 0-4, median grade 2 Duration of (current episode) symptoms (median [IQR): 365 (810-189) days.
Indirectness of population	No indirectness
Interventions	(n=101) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). Usual GP care with an exercise therapy. The exercise therapy consisted of maximally 12 treatment sessions during the first 3 months of follow-up and was administered by physiotherapists. Physiotherapist advised patients about lifestyle adaptations, possible walking aids, appropriate postural loading of joints, (in)appropriate pain behaviour and more. Exercises

	consisted of strengthening and improving flexibility of muscles around the hip joint (especially extensors andabductors), leg and abdominal muscles. Aerobic exercises to improve endurance were also included. Patients were expected to perform home exercises and wereprovided a booklet describing the exercises. During booster sessions advices and exercises were repeated and possible problems and obstacles to perform the home exercises were discussed. Usual GP care included a booklet about hip OA, and could include education, counselling, prescription of pain medication, additional diagnostic tests or referral to an orthopedic surgeon. Duration 3 months. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, flexibility, aerobic). (n=102) Intervention 2: No treatment. Usual GP care, including a brochure with information about hip OA. GP care could include education, counselling, prescription of pain medication, additional diagnostic tests or referral to an orthopedic surgeon. In the control group, referral to a physical therapist was discouraged, but was not restricted. Duration 3 months. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The Netherlands Organization for Health Research andDevelopment and the Dutch Arthritis Foundation)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: EQ-5D Quality of life at 3 months; Group 1: mean 0.78 (SD 0.162); n=101, Group 2: mean 0.777 (SD 0.147); n=102; EuroQuol (EQ5D) - 0.329-1.0 Top=High is good outcome; Comments: Baseline scores: exercise group 0.778 (0.122); control group 0.748 (0.161)
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6

Protocol outcome 2: Health related quality of life at > 3 months
- Actual outcome: EQ-5D Quality of life at 12 months; Group 1: mean 0.784 (SD 0.198); n=101, Group 2: mean 0.784 (SD 0.151); n=102; EuroQoL (EQ5D) - 0.329-1.0 Top=High is good outcome; Comments: Baseline scores: exercise group 0.778 (0.122); control group 0.748 (0.161)
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6

Protocol outcome 3: Physical function at </=3 months - Actual outcome: Function HOOS subscale at 3 months; Group 1: mean 28.8 (SD 21.3); n=101, Group 2: mean 35.7 (SD 19); n=102; hip osteoarthritis outcome score 0-100 Top=High is poor outcome; Comments: Baseline values: exercise group 35.4 (18); control group (38) (16.6) Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6 Protocol outcome 4: Physical function at > 3 months - Actual outcome: Function HOOS subscale at 12 months; Group 1: mean 26.8 (SD 21.2); n=101, Group 2: mean 34.2 (SD 21.4); n=102; hip osteoarthritis outcome score 0-100 Top=High is poor outcome; Comments: Baseline values: exercise group 35.4 (18); control group (38) (16.6) Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6 Protocol outcome 5: Pain at </=3 months - Actual outcome: Pain HOOS subscale at 3 months; Group 1: mean 31.8 (SD 17.7); n=101, Group 2: mean 36.2 (SD 18.9); n=102; hip osteoarthritis outcome score 0-100 Top=High is poor outcome; Comments: Baseline scores: Exercise group 37.6 (16.1); control group: 38.9 (15.7) Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6 Protocol outcome 6: Pain at > 3 months - Actual outcome: Pain HOOS subscale at 12 months; Group 1: mean 31.6 (SD 19.5); n=101, Group 2: mean 34.6 (SD 19.3); n=102; hip osteoarthritis outcome score 0-100 Top=High is poor outcome; Comments: Baseline scores: Exercise group 37.6 (16.1); control group: 38.9 (15.7) Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 5; Group 2 Number missing: 6	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Thorstensson 2005[444]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=61)
Countries and setting	Conducted in Sweden; Setting: Community.
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks (6 weeks of therapy, 20 weeks additional follow up)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Radiographic diagnosis with a Kellgren and Lawrence grade of 3 or more.
Stratum	Overall
Subgroup analysis within study	Not applicable:
Inclusion criteria	Age 35-65, living in the defined geographic area, diagnosis of radiographic osteoarthritis of Kellgren and Lawrence grade III or more (i.e. definite osteophytes and joint space narrowing).
Exclusion criteria	Inflammatory joint disease; anterior cruciate ligament injury; known symptomatic injury to the menisci; hip symptoms more aggravating than the knee symptoms; about to have knee replacement surgery within 6 months; comorbidities not allowing exercise.
Recruitment/selection of patients	People registered at GP practices in south-west of Sweden.
Age, gender and ethnicity	Age - Mean (SD): 56.1 (6.1). Gender (M:F): 29:31. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated/unclear. 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity Kellgren Lawrence grade at least 3. Duration of symptoms: Not stated/unclear.
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Exercise – Supervised strength exercise. One-hour exercise sessions, twice a week for 6 weeks supervised by a physical therapist. The program consisted of weight-bearing exercises aimed at increasing postural control and endurance and strength in the lower extremity. Exercises were performed at five stations at submaximal intensity (minimum 60% of maximum heart rate). Intensity was gradually and individually increased during the six weeks by increased lever arms or range of motion. People were encouraged to exercise at their most vigorous intensity possible, without losing quality in performance or severely exacerbating pain. Pain during exercise was not considered as an obstacle as long as the person perceived it

	as “acceptable” and there were no increased symptoms after 24 hours. People received a thera band to perform daily pulley exercises at home. In addition, three exercises, which were considered as the most challenging to the individual, were chosen as daily home exercises. People were recommended to perform some kind of weight bearing submaximal activity, such as walking or their home exercises, for at least 30 minutes or two times 15 minutes each day. Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=31) Intervention 2: No treatment. No lifestyle changes. They met with a physical therapist for one hour at three times; baseline, follow up at 6 weeks and 6 months After 6 months they were offered exercise classes or instructions and a home-exercise program. Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable
Funding	Grants received from The Vardal Foundation, Sweden, The Swedish Rheumatism Association in Stockholm, The Swedish Rheumatism Association in Gothenburg, The Swedish Research Council, The Department of Research and Development at Spenshult Hospital for Rheumatid Diseases, Halmstad, Sweden.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 Physical Component Summary at 6 weeks; Group 1: mean 3 (SD 25.84); n=30, Group 2: mean 0.3 (SD 37.90); n=31; SF-36 Physical Component Summary 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 3.0 (-5.9, 13.4). Reported control: 0.3 (-15.2, 12.6). Baseline scores: exercise group 42.5 (24.4, 57.5); control group 43.8 (24.2, 57.3)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 2: Health related quality of life at </=3 months
- Actual outcome: SF-36 Mental Component Summary at 6 weeks; Group 1: mean 1.6 (SD 34.28); n=30, Group 2: mean -2.1 (SD 38.71); n=31; SF-36 Mental Component Summary 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 1.6 (-10.6, 15.0). Reported control: -2.1 (-16.9, 11.5). Baseline scores: exercise group 55.6 (40.2, 66.2); control group 56.3 (37.0, 67.0)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,

Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 3: Health related quality of life at >3 months
- Actual outcome: SF-36 Physical Component Summary at 6 months; Group 1: mean 3 (SD 29.73); n=30, Group 2: mean -0.7 (SD 33.53); n=31; SF-36 Physical Component Summary 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 3.0 (-5.9, 16.3). Reported control: -0.7 (-14.8, 9.8). Baseline scores: exercise group 42.5 (24.4, 57.5); control group 43.8 (24.2, 57.3)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 4: Health related quality of life at >3 months
- Actual outcome: SF-36 Mental Component Summary at 6 months; Group 1: mean 0.7 (SD 41.91); n=30, Group 2: mean -0.7 (SD 40.35); n=31; SF-36 Mental Component Summary 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 0.7 (-18.1, 13.2). Reported control: -0.7 (-16.8, 12.8). Baseline scores: exercise group 55.6 (40.2, 66.2); control group 56.3 (37.0, 67.0)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 3: Physical function at </=3 months
- Actual outcome: ADL KOOS subscale at 6 weeks; Group 1: mean 2.0 (SD 11.52); n=30, Group 2: mean -0.6 (SD 17.45); n=31; KOOS ADL subscale 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 2.0 (-2.3, 6.3). Reported control: -0.6 (-7.0, 5.8). Baseline scores: exercise group 69 (18); control group 71 (21)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: ADL KOOS subscale at 6 months; Group 1: mean 0.9 (SD 12.59); n=30, Group 2: mean -1.9 (SD 15.81); n=31; KOOS ADL subscale 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 0.9 (-3.8, 5.6). Reported control: -1.9 (-7.7, 3.9). Baseline scores: exercise group 69 (18); control group 71 (21)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 5: Pain at </=3 months
- Actual outcome: Pain KOOS subscale at 6 weeks; Group 1: mean 1.8 (SD 13.39); n=30, Group 2: mean -0.3 (SD 16.22); n=31; KOOS Pain subscale 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 1.8 (-3.2, 6.8). Reported control: -0.3 (-6.2, 5.7). Baseline scores: exercise group 60 (18); control group 64 (19)
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 6: Pain at > 3 months - Actual outcome: Pain KOOS subscale at 6 months; Group 1: mean 3.1 (SD 13.52); n=30, Group 2: mean -1.1 (SD 14.99); n=31; KOOS Pain subscale 0-100 Top=High is good outcome; Comments: Reported as change score (95% CI). Converted to mean (SD). Reported exercise: 3.1 (-1.9, 8.2). Reported control: -1.1 (-6.6, 4.4). Baseline scores: exercise group 60 (18); control group 64 (19) Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 3	
Protocol outcomes not reported by the study	Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Vaghela 2020[458]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=83)
Countries and setting	Conducted in India; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 month (end of intervention)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Bilateral osteoarthritis of the knee based on the clinical American College of Rheumatology criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with bilateral osteoarthritis of the knee according to the clinical American College of Rheumatology criteria; age 40-80 years; both genders; patients who are functionally ambulatory; patients who have not practiced any form of yoga or exercises in the past 2 months.
Exclusion criteria	Patients with unilateral knee osteoarthritis, symptoms of locking or instability of knee, buckling and shifting or "complain of giving way" in the past 3 months; patients treated with corticosteroid injections within the past 2 months; patients with total knee arthroplasty; inflammatory arthritis; any recent trauma of knee joint or lower limb; patients who are taking analgesics
Recruitment/selection of patients	People referred for physiotherapy at Shree Krishna Hospital, Karamsad and from the nearby old-age homes/hospitals.
Age, gender and ethnicity	Age - Mean (SD): 55.5 (9.4). Gender (M:F): 25:58. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated/unclear Duration of symptoms: Not stated/unclear. Clinical Trial Registry - India (CTRI Number CTRI/2019/02/017422).
Indirectness of population	No indirectness
Interventions	(n=43) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Yoga therapy. This included six asanas, that is, Tadasana, Uttitha Trikonasana, Virbhadrasana, Dandasana, Sputa Padangustasana, and Badhha Konasana. Each asana consisted of ten repetitions with short intervals of rest in between for a total of 30 minutes per session, three times per week for 4 weeks. In addition to conventional physiotherapy.. Duration 4 weeks. Concurrent medication/care: Conventional physiotherapy program included the following: Transelectrical nerve stimulation (10 minutes), isometric quadriceps exercise, straight leg-raising exercise in supine, terminal knee extension or vastus medialis oblique strengthening exercise in supine and high sitting; straight leg abduction exercise in side lying. Each exercise was performed for

	a total of three sets, with each set made up of ten repetitions for 20 minutes, three times a week for 4 weeks.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Yoga and strengthening). (n=40) Intervention 2: Exercise - Supervised strength exercise. Conventional physiotherapy only. Duration 4 weeks. Concurrent medication/care: Conventional physiotherapy program included the following: Transelectrical nerve stimulation (10 minutes), isometric quadriceps exercise, straight leg-raising exercise in supine, terminal knee extension or vastus medialis oblique strengthening exercise in supine and high sitting; straight leg abduction exercise in side lying. Each exercise was performed for a total of three sets, with each set made up of ten repetitions for 20 minutes, three times a week for 4 weeks.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Not applicable
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED STRENGTH EXERCISE

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF-36 physical functioning (domain A) at 4 weeks; Group 1: mean 7.6 (SD 1.41); n=43, Group 2: mean 7.55 (SD 1.1); n=40; SF-36 physical functioning 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 2.34 (0.48). Baseline strength: 3.6 (1.3).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 role physical (domain B) at 4 weeks; Group 1: mean 28.58 (SD 1.35); n=43, Group 2: mean 26.85 (SD 2.23); n=40; SF-36 role physical 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 16.39 (3.76). Baseline strength: 17.62 (3.86).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 role emotional (domain C) at 4 weeks; Group 1: mean 7.79 (SD 0.41); n=43, Group 2: mean 7.35 (SD 0.8); n=40; SF-36 role emotional 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 4.04 (0.21). Baseline strength: 4.47 (0.71).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 vitality (domain D) at 4 weeks; Group 1: mean 10.69 (SD 0.46); n=43, Group 2: mean 8.9 (SD 0.84); n=40; SF-36 vitality 0-100

Top=High is good outcome; Comments: Baseline mixed modality: 6.88 (0.76). Baseline strength: 6.42 (0.54).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 mental health (domain E) at 4 weeks; Group 1: mean 8.55 (SD 1.18); n=43, Group 2: mean 7.575 (SD 1.17); n=40; SF-36 mental health 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 2.46 (0.54). Baseline strength: 5.15 (1.16).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 social functioning (domain F) at 4 weeks; Group 1: mean 31.76 (SD 5.35); n=43, Group 2: mean 31.15 (SD 8.61); n=40; SF-36 social functioning 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 48.60 (2.68). Baseline strength: 42.32 (5.91).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 bodily pain (domain G) at 4 weeks; Group 1: mean 4.67 (SD 0.52); n=43, Group 2: mean 3.35 (SD 1.31); n=40; SF-36 bodily pain 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 2.18 (0.54). Baseline strength: 2.37 (0.54).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up
- Actual outcome: SF-36 general health (domain H) at 4 weeks; Group 1: mean 18.67 (SD 0.91); n=43, Group 2: mean 17.42 (SD 1.17); n=40; SF-36 general health 0-100 Top=High is good outcome; Comments: Baseline mixed modality: 13.48 (1.07). Baseline strength: 13.45 (1.31).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: WOMAC function (domain C) at 4 weeks; Group 1: mean 13.79 (SD 5.52); n=43, Group 2: mean 19.17 (SD 7.02); n=40; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline mixed modality: 42.32 (10.58). Baseline strength: 48.15 (9.96).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up

Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC pain (domain A) at 4 weeks; Group 1: mean 4.32 (SD 1.93); n=43, Group 2: mean 7.3 (SD 2.13); n=40; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline mixed modality: 13.16 (3.68). Baseline strength: 13.6 (2.46). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender and baseline values of outcomes. Reported that generally there was an absence of any baseline data.; Group 1 Number missing: 2, Reason: 2 lost to follow-up; Group 2 Number missing: 5, Reason: 5 lost to follow-up	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Van baar 2001[459]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=201)
Countries and setting	Conducted in Netherlands; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks + 24 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Osteoarthritis of the hip or knee according to the clinical criteria of the American College of Rheumatology with radiographic confirmation
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	OA of the hip or knee according to the clinical criteria of the American College of Rheumatology
Exclusion criteria	Another disease which might explain the complaints; complaints in fewer than 10 of 30 days; treatment for these complaints with exercise in the preceding six months; age under 40 or over 85; indication for hip or knee replacement; contraindication for exercise treatment; contraindications for analgesics or non-steroidal anti-inflammatory drugs (NSAIDs); and inability to understand the Dutch language
Recruitment/selection of patients	Selected by GPs
Age, gender and ethnicity	Age - Mean (SD): 68.0 (8.80. Gender (M:F): 44/157. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Median 1-no more than 6 years
Indirectness of population	No indirectness
Interventions	(n=99) Intervention 1: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Exercise treatment included exercises for muscle functions (strength and length), mobility, and coordination, and exercises for elementary movement abilities and locomotion abilities. Also, instructions for the adaptation of activities of daily living and home exercises were given. Content, intensity, and frequency of treatment were tailored to the patient's needs. Each session lasted approximately 30 minutes. Duration 12 weeks. Concurrent

	medication/care: GP provided patient education (including a brochure) and drug treatment, if necessary. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, coordination). (n=102) Intervention 2: No treatment. GP provided patient education (including a brochure) and drug treatment, if necessary, as in the experimental group. . Duration 12 weeks. Concurrent medication/care: GP provided patient education (including a brochure) and drug treatment, if necessary. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The Dutch Fund of Investigative Medicine of the Dutch HealthInsurance Council)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT Protocol outcome 1: Pain at </=3 months - Actual outcome: Pain past week at 12 weeks (end of intervention); MD; -17.0 (95%CI -23.6 to -10.4) VAS 0-100 Top=High is poor outcome, Comments: n (exercise) = 93. n (control) = 98. Baseline exercise: 34.0 (27.2). Baseline control: 28.7 (26.0).; Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 6; Group 2 Number missing: 4 Protocol outcome 2: Pain at > 3 months - Actual outcome: Pain past week at 36 weeks; MD; -6.6 (95%CI -14.7 to 1.6) VAS 0-100 Top=High is poor outcome, Comments: n (exercise) = 90. n (control) = 92. Baseline exercise: 34.0 (27.2). Baseline control: 28.7 (26.0).; Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 9; Group 2 Number missing: 10	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Wang 2007[478]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in USA; Setting: Not stated
Line of therapy	Unclear
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with osteoarthritis of the hip or knee
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed with osteoarthritis of the hip or knee; aged 25 years or older; able to speak and read English; able to obtain medical clearance through a primary healthcare provider to participate in the study
Exclusion criteria	Intra-articular corticosteroid injections in the past 30 days; had undergone joint replacement surgery in the past 6 months or were scheduled for joint replacement within 3 months of the start of the study; were currently exercising >20 minutes per week for the past 2 months; were currently using a wheelchair for mobility
Recruitment/selection of patients	Convenience sample of participants recruited from community sources such as flyers in local community centers, physicians' offices, YMCA offices and Parks and Recreation Departments. Also, invitations to Arthritis Foundation members and University campus advertisements
Age, gender and ethnicity	Age - Mean (SD): 66.2 (12.6). Gender (M:F): 5:32. Ethnicity: White: 33; Other: 5; Not reported: 4
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip or knee).
Extra comments	Severity: Number of tender joints = 6.8 (4.8) Duration of symptoms (mean [SD]): 13.5 (11.8)
Indirectness of population	No indirectness
Interventions	(n=21) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic exercise, which consisted of warm up, flexibility and strength training, and cool down. Exercises were divided into six sections (warm up, flexibility, endurance, lower body, upper body, cool down), focusing on joints in the trunk,

	shoulders, elbows, wrists, fingers, hip, knees, ankles and toes, and emphasizes muscle groups of upper and lower limbs. Duration 12 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (n=21) Intervention 2: No treatment. Participants were asked to continue their physical activity as usual and offered an opportunity to participate in the aquatic programme at the end of the trial.. Duration 12 weeks . Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Behavioural Nursing Research Training Grant, The Women's Health Nursing Research Training Grant, The Hester McLaw Nursing Scholarship, and the deTornyay Centre for Health Aging Scholarship)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT Protocol outcome 1: Physical function at </=3 months - Actual outcome: MDHAQ Physical functioning at 12 weeks (end of intervention); Group 1: mean 0.9 (SD 0.4); n=20, Group 2: mean 1 (SD 0.5); n=18; multidimensional Health Assessment Questionnaire 0-3 Top=High is poor outcome; Comments: Baseline scores: exercise group 0.9 (0.4); control group 0.95 (0.5) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment status (75% retired vs 44%); Group 1 Number missing: 1; Group 2 Number missing: 3 Protocol outcome 2: Pain at </=3 months - Actual outcome: VAS Bodily pain at 12 weeks (end of intervention); Group 1: mean 43.5 (SD 18.6); n=20, Group 2: mean 54.9 (SD 25.2); n=18; Visual analogue scale 0-100 Top=High is poor outcome; Comments: Baseline scores: exercise group 52.2 (23.8); control group 55.3 (24.6) Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment status (75% retired vs 44%); Group 1 Number missing: 1; Group 2 Number missing: 3	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at

	</=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Wang 2011[479]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=78)
Countries and setting	Conducted in Taiwan; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed by physician assessment based on symptoms and X-ray
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	(1) age over 55 years, (2) diagnosed with knee OA by physician assessment based on symptoms and X-ray and (3) consented to participate
Exclusion criteria	(1) having a medical condition precluding exercise (i.e. uncontrolled arrhythmias, third-degree heart block, myocardial infarction within six months, unstable angina, acute congestive heart failure and uncontrolled epilepsy), (2) having intra-articular corticosteroid injections in the past 30 days, (3) received a joint replacement previously, or (4) currently exercising more than 60 minutes per week for the past two months
Recruitment/selection of patients	Flyers and posters were distributed in local community centres and sport centres, and a recruitment social event was held
Age, gender and ethnicity	Age - Mean (SD): 67.7 (5.9). Gender (M:F): 11:67. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Low morbidity score (Mean number of comorbid conditions: 1.0 (1.0)). 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 6.8 (6.4).
Indirectness of population	No indirectness
Interventions	(n=28) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Aquatic exercise was based on the Arthritis Foundation Aquatics Program instructors manual. The main components of the programme include a 60-minute flexibility and aerobic training class, three times a week for 12 weeks. The exercise training focuses on joint in the trunk, shoulders, arms and legs and emphasises the muscle groups of the upper and lower limbs as well as balance and

	coordination. The mechanisms for fitness training involve changes in speed, surface area, direction of movement and turbulence in water to increase the exercise resistance and to create intensity variation. A trained exercise instructor taught the group classes at the public swimming pools. Duration 12 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Hydrotherapy (Aquatic exercise). (n=28) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). A land based exercise protocol was based on the People with Arthritis Can Exercise programme instructor's manual. The main components of the programme include a 60-minute flexibility and aerobic training class, three times a week for 12 weeks. The exercise training focuses on joints in the trunk, shoulders, arms and legs and emphasises the muscle groups of the upper and lower limbs as well as balance and coordination. To assure safe performance of the exercise, the classes include instruction about basic principles of arthritis exercise, correct body mechanics and joint protection. Movement against gravity and variations in speed, level of leg or arm raising, or moving both extremities simultaneously were used to create different levels of training intensity. The average number of repetitions for each exercise begins with 10 and gradually increases to 15. Classes were taught to a group of participants by the trained instructor at an indoor basketball court. Duration 12 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Other (Upper body, lower body training, flexibility, aerobic). (n=26) Intervention 3: No treatment. Control group, no further details reported. Duration 12 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (The National Science Council of Republic of China)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED)	

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS Quality of life at 12 weeks (end of intervention); Group 1: mean 73 (SD 12); n=26, Group 2: mean 74 (SD 11); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 2

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS ADL at 12 weeks (end of intervention); Group 1: mean 76 (SD 16); n=26, Group 2: mean 82 (SD 14); n=26; KOOS ADL 0-100 Top=High is good outcome; Comments: Baseline hydro: 73 (20). Baseline mixed: 75 (16).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 2

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS Pain at 12 weeks (end of intervention); Group 1: mean 72 (SD 18); n=26, Group 2: mean 76 (SD 15); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 2

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS Quality of life at 12 weeks (end of intervention); Group 1: mean 73 (SD 12); n=26, Group 2: mean 67 (SD 13); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS ADL at 12 weeks (end of intervention); Group 1: mean 76 (SD 16); n=26, Group 2: mean 69 (SD 18); n=26; KOOS ADL 0-100 Top=High is good outcome; Comments: Baseline hydro: 73 (20). Baseline control: 70 (19).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS Pain at 12 weeks (end of intervention); Group 1: mean 72 (SD 18); n=26, Group 2: mean 68 (SD 18); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus NO TREATMENT

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: KOOS Quality of life at 12 weeks (end of intervention); Group 1: mean 74 (SD 11); n=26, Group 2: mean 67 (SD 13); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

Protocol outcome 2: Physical function at </=3 months
- Actual outcome: KOOS ADL at 12 weeks (end of intervention); Group 1: mean 82 (SD 14); n=26, Group 2: mean 69 (SD 18); n=26; KOOS ADL 0-100 Top=High is good outcome; Comments: Baseline mixed: 75 (16). Baseline control: 70 (19).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

Protocol outcome 3: Pain at </=3 months
- Actual outcome: KOOS Pain at 12 weeks (end of intervention); Group 1: mean 76 (SD 15); n=26, Group 2: mean 68 (SD 18); n=26; Knee Injury and Osteoarthritis Outcome Score 0-100 Top=High is good outcome
Risk of bias: All domain – Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Williamson 2007[485]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=181)
Countries and setting	Conducted in United Kingdom; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks + 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People listed for knee arthroplasty with osteoarthritis of the knee
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients listed for knee arthroplasty due to OA; patients with unilateral or bilateral knee pain; pain lasting more than 3 months
Exclusion criteria	Taking anticoagulants; within 2 months after receiving anintra-articular steroid injection; experiencing back pain associated with referred leg pain; suffering from ipsilateral OA of the hip; suffering psoriasisor other skin disease in the region of the knee; suffering from rheumatoid arthritis; and if they had received acupuncture or physiotherapy treatment in the last year
Recruitment/selection of patients	Participants were on the waiting list for knee replacement surgery
Age, gender and ethnicity	Age - Mean (SD): 70.6 (9.0). Gender (M:F): 84/97. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: Exercise - Supervised strength exercise. Exercise was done in groups of 6–10 patients, hourly, once a week for 6 weeks. They carried out an exercise circuit devised and supervised by the same physiotherapist who provided the acupuncture. The exercises were: static quadriceps contractions; inner range quadriceps contractions; straight leg raises; sit to stands, stair climbing; calf stretches; theraband resisted knee extensions; wobble board balance training; knee flexion/extension sitting on gym ball and freestanding peddle revolutions.. Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness

	Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Not applicable (n=61) Intervention 2: No treatment. An exercise and advice leaflet, which had been designed by consensus between the physiotherapy, rheumatology and orthopaedic departments. Patients were told that they were in the 'home exercise group' . Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable (n=60) Intervention 3: Other. Acupuncture treatment. Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable Comments: This group was excluded as they did not fulfil the inclusion criteria
Funding	Academic or government funding (Research and Development Grant, The Great Western Hospital, Swindon)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Pain at </=3 months
- Actual outcome: Pain at 12 weeks (from baseline); Group 1: mean 6.36 (SD 2.6); n=60, Group 2: mean 7.24 (SD 2.07); n=61; VAS 0-10 Top=High is poor outcome; Comments: Baseline exercise: 6.8 (2.64). Baseline no treatment: 6.89 (2.29).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 19, Reason: Reasons not given; Group 2 Number missing: 26, Reason: Reasons not given

Protocol outcome 2: Psychological distress at </=3 months
- Actual outcome: HADs anxiety at 12 weeks (from baseline); Group 1: mean 7.08 (SD 5.16); n=60, Group 2: mean 6.54 (SD 3.93); n=61; HADs - anxiety 0-21 Top=High is poor outcome; Comments: Baseline exercise: 7.45 (4.94). Baseline no treatment: 6.69 (3.63).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 19, Reason: Reasons not given; Group 2 Number missing: 26, Reason: Reasons not given
- Actual outcome: HADs depression at 12 weeks (from baseline); Group 1: mean 6.75 (SD 3.84); n=60, Group 2: mean 7.13 (SD 3.54); n=61; HADS - depression 0-21 Top=High is poor outcome; Comments: Baseline exercise: 7.1 (3.88). Baseline no treatment: 7.43 (3.40).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 19, Reason: Reasons not given;

Group 2 Number missing: 26, Reason: Reasons not given	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Wortley 2013[489]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=39)
Countries and setting	Conducted in USA; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 10 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: The Classification Criteria for Knee OA of the American College of Rheumatology and bilateral knee x-rays
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Between the ages of 60 and 85 years, and have knee OA
Exclusion criteria	Received arthroscopic surgery or an intra-articular injection within the past 3 months, neurological disorders, or had participated in a resistance training or Tai Ji in the past 6 months
Recruitment/selection of patients	Senior centres, local newspaper advertisements, and a local newsletter for seniors
Age, gender and ethnicity	Age - Mean (SD): 69.2 (6.0). Gender (M:F): 9:22. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 1-3, median grade 2-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Exercise - Supervised strength exercise. A resistance training programme consisting of two 1-hour sessions per week. The program included the following knee and hip exercises performed with ankle cuff weights: seated leg extension, standing hamstring curl, straight leg raise, standing hip abduction, standing hip adduction,standing hip flexion, standing calf raise. Participants started with either a 5lb or 10 lb ankle weight and progressed from two sets of eight repetitions to three sets of 12 repetitions during the first 6 weeks, and were allowed to increase the weight as needed during the final 4 weeks. Duration 10 weeks. Concurrent medication/care: Participants were asked not to alter their regular physical activity or pain medications. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not stated

	/ Unclear 3. Type of exercise: Proprioception (n=15) Intervention 2: Exercise - Other supervised exercise (including flexibility, proprioception). A 1-h group training session twice per week which involved a program of 12 basic movements adapted from the Yang style Tai Ji. The program began by learning the first two movements during the first session, and then adding a new movement during each session for the first 5 weeks. In each training session of the first weeks, sufficient time was provided for practicing the new and previously learned movements. During the last 5 weeks, participants also practiced the movements in the opposite direction to the original direction in order to similarly “load” both lower limbs. Duration 10 weeks. Concurrent medication/care: Participants were asked not to alter their regular physical activity or pain medications. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (n=9) Intervention 3: No treatment. Participants asked not to alter their usual physical activity or medication during the 10 weeks of the intervention, and were contacted once by telephone during the intervention. Duration 10 weeks. Concurrent medication/care: Participants were asked not to alter their regular physical activity or pain medications. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	Academic or government funding (Supported in part by funds from UTK Office of Research,College of Education, Health and Human Sciences, and University of TennesseeMedical Center, The University of Tennessee)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION)

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC - physical function at 10 weeks (end of intervention); Group 1: mean 240 (SD 249); n=13, Group 2: mean 552 (SD 392); n=12; WOMAC 0-1800 Top=High is poor outcome; Comments: Baseline Tai Ji: 694 (361). Baseline strength: 494 (265).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudo-randomised"; Indirectness of outcome: No indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 2: Pain at </=3 months

- Actual outcome: WOMAC - pain at 10 weeks (end of intervention); Group 1: mean 71 (SD 100); n=13, Group 2: mean 141 (SD 107); n=12; WOMAC 0-500 Top=High is poor outcome; Comments: Baseline Tai Ji: 169 (135). Baseline strength: 155 (110).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudo-randomised"; Indirectness of outcome: No indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 2; Group 2 Number missing: 3

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SUPERVISED STRENGTH EXERCISE versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC - physical function at 10 weeks (end of intervention); Group 1: mean 240 (SD 249); n=13, Group 2: mean 475 (SD 282); n=6; WOMAC 0-1800 Top=High is poor outcome; Comments: Baseline strength: 494 (265). Baseline no treatment: 547 (369).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudo-randomised"; Indirectness of outcome: No indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 2; Group 2 Number missing: 3

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC - pain at 10 weeks (end of intervention); Group 1: mean 71 (SD 100); n=13, Group 2: mean 157 (SD 96); n=6; WOMAC 0-500 Top=High is poor outcome; Comments: Baseline strength: 155 (110). Baseline no treatment: 170 (86).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudo-randomised"; Indirectness of outcome: No indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 2; Group 2 Number missing: 3

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC - physical function at 10 weeks (end of intervention); Group 1: mean 552 (SD 392); n=12, Group 2: mean 475 (SD 282); n=6; WOMAC 0-1800 Top=High is poor outcome; Comments: Baseline Tai Ji: 694 (361). Baseline no treatment: 547 (369).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudorandomised"; Indirectness of outcome: No indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 3; Group 2 Number missing: 3

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC - pain at 10 weeks (end of intervention); Group 1: mean 141 (SD 107); n=12, Group 2: mean 157 (SD 96); n=6; WOMAC 0-500 Top=High is poor outcome; Comments: Baseline Tai Ji: 169 (135). Baseline no treatment: 170 (86).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Comments - States that the randomisation method was "pseudorandomised"; Indirectness of outcome: No

indirectness ; Baseline details: Difference in PASE baseline score; Group 1 Number missing: 3; Group 2 Number missing: 3	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	**Xiao 2020[491] (Xiao 2021[492])**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=98)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 months (the companion study reports an additional 3 months. However, this appears to re-randomise the original cohort so this will not be included in the data extraction as this makes the interpretation difficult to make).
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with knee osteoarthritis diagnosed by senior physicians based on standard clinical, endoscopic, radiologic and histological criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Subjects who had a clinical diagnosis of knee osteoarthritis, with or without radiographic changes
Exclusion criteria	Contraindication for exercise, NSAIDs or X-rays; had leg surgery/trauma within the last 6 months
Recruitment/selection of patients	Recruited from the Department of Orthopaedics, Haidian Hospital (Beijing, China)
Age, gender and ethnicity	Age - Mean (SD): 70.4 (9.72). Gender (M:F): 37:61. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging (Mixed - could have been with or without imaging). 3. Multimorbidity: Not applicable 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: X-ray classification grade I-II, median grade II Duration of symptoms (SD): 12.44 (4.17) months
Indirectness of population	No indirectness

Interventions	(n=49) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Wu Qin Xi Qigong exercise program. Three parts: warming up (10-15 minutes of aerobic activity at "rather strenuous level"), Wu Qin Xi Qigong exercise 40-45 minutes and cool-down 5 minutes. The exercise consists of 10 movement routines: 1) raising the tiger's paws, 2) seizing the prey, 3) colliding with the antlers, 4) running like a deer, 5) rotating the waist like a bear, 6) swaying like a bear, 7) lifting the monkey's paws, 8) picking fruit, 9) stretching upward, 10) flying like a bird. The whole protocol usually takes 12-15 minutes to complete at the usual pace. Each participant performed three repetitions, with a 2-minute rest period between the sets. Training took place in groups four times a week (each session 60 minutes) for 24 weeks. The exercise is led by an experienced physical therapist. Each participant can make the appropriate adjustments to the difficulty of movement when implemented in accordance with the physical condition of each person, which was reassessed every 4 weeks.. Duration 24 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Qigong). (n=49) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). Conventional physical therapy consisting of muscle-strength training of the lower extremity and aerobic training. The exercise program was conducted 4 days a week for 24 weeks, with a gradual increase in training intensity, knee load and exercise difficulty during the program. The resistance training was performed using three sets per exercise at intensities between 6 and 12 maximum repetitions. The aerobic training lasted 30 minutes and was performed at 75%-85% of heart rate. Training intensity and amount of attention from an experienced physical therapist were intended to be similar in both groups.. Duration 24 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength and aerobic).
Funding	Academic or government funding (This work was supported in part by Beijing Municipal Education Commission of Science and Technology for the general project (KM201910015001).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED)

Protocol outcome 1: Physical function at > 3 months
- Actual outcome: WOMAC function at 24 weeks; Group 1: mean 20.7 (SD 8.7); n=45, Group 2: mean 18.8 (SD 7.4); n=40; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline other supervised: 28.9 (11.7). Baseline mixed modality: 27.4 (10.9).
Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, x-ray classification, duration and baseline values of symptoms; Group 1 Number missing: 4, Reason: 4 failed to attend training at the scheduled time in the study; Group 2 Number missing: 9, Reason: 9 failed to attend training at the scheduled time in the study

Protocol outcome 2: Pain at > 3 months

- Actual outcome: WOMAC pain at 24 weeks; Group 1: mean 5 (SD 3.4); n=45, Group 2: mean 5.4 (SD 3.5); n=40; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline other supervised: 7.8 (3.8). Baseline mixed modality: 7.2 (3.9). Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, x-ray classification, duration and baseline values of symptoms; Group 1 Number missing: 4, Reason: 4 failed to attend training at the scheduled time in the study; Group 2 Number missing: 9, Reason: 9 failed to attend training at the scheduled time in the study Protocol outcome 3: Serious adverse events at > 3 months - Actual outcome: Serious adverse events at 24 weeks; Group 1: 0/49, Group 2: 0/49 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, weight, height, BMI, x-ray classification, duration and baseline values of symptoms; Group 1 Number missing: 4, Reason: 4 failed to attend training at the scheduled time in the study; Group 2 Number missing: 9, Reason: 9 failed to attend training at the scheduled time in the study	
Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at </=3 months; Pain at </=3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months

Study	**Ye 2019[496]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=50)
Countries and setting	Conducted in Malaysia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed with knee osteoarthritis according to the criteria of the American College of Rheumatology with radiographic grading of the severity between 2 and 3
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosed with knee osteoarthritis according to the criteria of the American College of Rheumatology, with radiographic grading of the severity between 2 and 3 and knee pain of <5 on the 10-point visual analogue scale; were aged between 50 and 80; were able to independently ambulate without language problem in order to perform movements.

Exclusion criteria	Suffered major diseases (Cardiovascular, respiratory or other musculoskeletal diseases) that required hospitalisation; had an implanted cardiac pacemaker; were on medication affecting the musculoskeletal system, or proprioception and postural stability (e.g. anti-depressants, dopaminergic ents, and hypnotic); partook in regular exercise of more than three times per week; fractured a bone within the past 12 months.
Recruitment/selection of patients	Recruited though advertisements and referral from their doctors of the Rehabilitation Hospital between January and December 2016.
Age, gender and ethnicity	Age - Mean (SD): 63.8 (6.2). Gender (M:F): 20:30. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Radiographic grade 2-3. Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=25) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). Banduajin Qigong training. 3 sessions per week with each session lasting 40 minutes (10 minutes for a warm-up and cool-down and 30 minutes for movements). The Banduajin Qigong training regime was in line with Health Qigong-Baduanjin published by the Health-Qigong Management Center of the General Administration of Sport of China in 2003. More specifically, this intervention involved two phases. Knee osteoarthritis patients were asked to attend group-based training for first 4 weeks at the Rehabilitation Hospital, administered by a certified instructor with at least of 5 years of teaching experience. In phase 2 (week 4-12), knee osteoarthritis patients were asked to practice at home. To maximize adherence to the training program, they were required to record themselves during practice. In addition, a reminder phone call was made every 2 weeks to increase exercise adherence. In the meanwhile, knee osteoarthritis patients were required to return to the Rehabilitation Hospital once per month and to attend the group-based training in which knee osteoarthritis patients were given the opportunities to ask questions about the exercise.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Qigong). (n=25) Intervention 2: No treatment. Patients in the control group were informed to maintain their unaltered lifestyle while refraining from other supervised exercise training program.. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable
Funding	-- (This work was supported by Natural Science Foundation of Fujian Province of China (Grant No. 2014J01347), National Natural Science Foundation of China (Grant No. 81173316), and Central Guide to Local Science and Technology Development (Grant No. 2018L3009). Equipment was supported by Fujian Provincial Rehabilitation Industrial Institution, Fujian Key Laboratory of Rehabilitation Technology, and Fujian Key Laboratory of Integrative Medicine on Geriatrics.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 12 weeks; Group 1: mean 15.64 (SD 8.87); n=25, Group 2: mean 17 (SD 8.71); n=25; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline other supervised exercise: 20.84 (11.04). Baseline no treatment: 19.08 (8.41).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 3.68 (SD 6.07); n=25, Group 2: mean 4.92 (SD 1.41); n=25; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline other supervised exercise: 6.64 (2.74). Baseline no treatment: 7.68 (9.34).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, height, weight, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	**Ye 2020[497]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=56)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis diagnosed according to the criteria of the American College of Rheumatology, with a radiographic grading of the severity between 2 and 3
Stratum	Overall
Subgroup analysis within study	Not applicable

Inclusion criteria	Males and females aged over 60 years; diagnosed with knee osteoarthritis according to criteria of the American College of Rheumatology, with a radiographic grading of the severity between 2 and 3; had no training experience in any kinds of mind-body exercise (Tai Chi, Qigong and/or Yoga) prior to six months before enrollment; able to ambulate without any device assistance; enrolled as an inpatient.
Exclusion criteria	Participants with severe cardiovascular, respiratory or other musculoskeletal diseases; participants on medications that could affect the musculoskeletal system or postural stability (e.g. antidepressants, dopaminergic agents, and hypnotics); participants who had a bone fracture within one year.
Recruitment/selection of patients	People recruited through referral from their doctors of the Rehabilitation Hospital between January 2015 and January 2016.
Age, gender and ethnicity	Age - Mean (SD): 64.4 (5.1). Gender (M:F): 19:37. Ethnicity: Not stated/unclear
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Radiographic grade between 2 and 3 Duration of symptoms: Not stated/unclear
Indirectness of population	No indirectness
Interventions	(n=28) Intervention 1: Exercise - Other supervised exercise (including flexibility, proprioception). 12 weeks Banduanjin program. People were asked to perform three banduanjin sessions per week, with each session lasting for 40 minutes. This training scheme was consistent with a guideline recommended by the Chinese Health-Qigong Association. The content of the exercise includes eight sections, namely, section 1: elevate both hands to the sky, section 2: draw a bow on both sides, section 3: raise single arm each time, section 4: look back, section 5: sway the head and shake the tail, section 5: sway the head and shake the tail, section 6: touch toes by hands with flexion of hip and extension of knee joint, section 7: clench fists, section 8: bounce on the toes. People in the intervention group took part in group-based training sessions in the hospital for the first four weeks under the supervision of a qualified instructor. After the initial in-hospital training, participants were instructed to continue to practice at home for the remaining time (till week 12). In order to maximise adherence, all people were asked t keep a daily log and the research team checked in via telephone.. Duration 12 weeks. Concurrent medication/care: All people received conventional therapies (acupuncture, massage and moxibustion), one hour each day, five days a week for the first four weeks.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Group session 3. Type of exercise: Mind-body (e.g. Tai Chi, Yoga, Qiqong) (Banduajin Qigong). (n=28) Intervention 2: No treatment. Usual care. Participants received the exercise intervention after 12 weeks.. Duration 12 weeks. Concurrent medication/care: All people received conventional therapies (acupuncture, massage and moxibustion), one hour each day, five days a week for the first four weeks.. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Not applicable 3. Type of exercise: Not applicable

Funding	Academic or government funding (This work was supported by the Central Guide to Local Science and Technology Development (grant no. 2018L3009). The Fujian Provincial Rehabilitation Industrial Institution and the Fujian Key Laboratory of Rehabilitation Technology provided equipment.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OTHER SUPERVISED EXERCISE (INCLUDING FLEXIBILITY, PROPRIOCEPTION) versus NO TREATMENT

Protocol outcome 1: Physical function at </=3 months
- Actual outcome: WOMAC physical function at 12 weeks; Group 1: mean 9.96 (SD 5.95); n=28, Group 2: mean 16 (SD 6.54); n=28; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline exercise: 20.36 (10.52). Baseline no treatment: 18.25 (8.04).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, height, weight, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Pain at </=3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 3.79 (SD 5.83); n=28, Group 2: mean 4.64 (SD 1.91); n=28; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline exercise: 6.21 (2.67). Baseline no treatment: 7.36 (8.89).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported gender, age, height, weight, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcomes not reported by the study	Health related quality of life at </=3 months; Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Study	Yilmaz 2019[498]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in Turkey; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 6 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis with radiographic grade 2-3 changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Primer knee OA diagnosis, grade II–III knee osteoarthritis according to Kellgren Lawrence stage, no intra-articular steroid injection or surgery in any joint in the last six months, and no participation in a physical therapy and rehabilitation program in the last six months
Exclusion criteria	Exclusion criteria were as follows: any orthopedic injuries and neurological problems affecting the balance and function of the person
Recruitment/selection of patients	Not reported
Age, gender and ethnicity	Age - Mean (SD): 60.27 (9.77). Gender (M:F): 20/60. Ethnicity: Not reported
Further population details	1. Age: under or aged 75 years 2. Diagnosis with or without imaging: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee osteoarthritis
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Median 6-12 months
Indirectness of population	No indirectness
Interventions	(n=39) Intervention 1: Exercise - Unsupervised mixed modality exercise (e.g. aerobic and strength exercise combined). A home exercise programme designed to improve strength, flexibility, functional ability and quality of life. Strengthening, stretching and range of motion exercises were performed daily. A home exercise brochure was given by the physiotherapist without any training. The purpose and methods were not explained. Exercises were given everyday twice a day, with 15 repetitions . Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, stretching, range of motion).

	(n=41) Intervention 2: Exercise - Supervised mixed modality exercise (e.g. aerobic and strength exercise combined). The same home exercise programme was instructed by a physiotherapist, and home exercises were taught at the hospital. Patients were instructed about knee joint protection and were interviewed once a week by telephone about their clinical status. Duration 6 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Class of medicine: Not applicable 2. Group or individual : Individual session 3. Type of exercise: Other (Strength, stretching, range of motion).
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: UNSUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED) versus SUPERVISED MIXED MODALITY EXERCISE (E.G. AEROBIC AND STRENGTH EXERCISE COMBINED)

Protocol outcome 1: Health related quality of life at </=3 months
- Actual outcome: SF36 - physical function at 6 weeks (end of intervention); Group 1: mean 60 (SD 15.36); n=39, Group 2: mean 64.05 (SD 12.9); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 51.84 (17.88). Baseline supervised: 44.76 (14.75).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00
- Actual outcome: SF36 - role physical at 6 weeks (end of intervention); Group 1: mean 61.84 (SD 32.66); n=39, Group 2: mean 77.38 (SD 28.4); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 25 (37.26). Baseline supervised: 34.52 (33.98).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00
- Actual outcome: SF36 - bodily pain at 6 weeks (end of intervention); Group 1: mean 64.11 (SD 21.56); n=39, Group 2: mean 74.1 (SD 12.63); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 34.95 (18.84). Baseline supervised:34.62 (17.05).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00
- Actual outcome: SF36 - general health at 6 weeks (end of intervention); Group 1: mean 57.89 (SD 16); n=39, Group 2: mean 67.62 (SD 15.4); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 53.89 (13.57). Baseline supervised: 56.62 (15).
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00
- Actual outcome: SF36 - vitality at 6 weeks (end of intervention); Group 1: mean 50 (SD 22.28); n=39, Group 2: mean 51.67 (SD 23.41); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 43.16 (22.80). Baseline supervised: 44.01 (24).

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00 - Actual outcome: SF36 - social function at 6 weeks (end of intervention); Group 1: mean 65.53 (SD 17.61); n=39, Group 2: mean 5.95 (SD 16.31); n=41; SF36 0-100 Top=High is good outcome; Comments: The supervised group score seems likely to be a typo. Baseline unsupervised: 56.37 (25.44). Baseline supervised: 57.43 (23.94). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00 - Actual outcome: SF36 - role emotional at 6 weeks (end of intervention); Group 1: mean 61.21 (SD 38.96); n=39, Group 2: mean 87.19 (SD 24.83); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 45.47 (43.32). Baseline supervised: 42.62 (38.13). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00 - Actual outcome: SF36 - mental health at 6 weeks (end of intervention); Group 1: mean 75.62 (SD 17.1); n=39, Group 2: mean 75.62 (SD 17.1); n=41; SF36 0-100 Top=High is good outcome; Comments: Baseline unsupervised: 63.21 (18.37). Baseline supervised: 66.67 (16.79). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00 Protocol outcome 2: Physical function at </=3 months - Actual outcome: WOMAC - function at 6 weeks (end of intervention); Group 1: mean 18.89 (SD 8.29); n=39, Group 2: mean 13.71 (SD 9.01); n=41; WOMAC 0-68 Top=High is poor outcome; Comments: Baseline unsupervised: 26.79 (9.03). Baseline supervised: 25.81 (10.24). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00 Protocol outcome 3: Pain at </=3 months - Actual outcome: WOMAC - pain at 6 weeks (end of intervention); Group 1: mean 6.74 (SD 4.64); n=39, Group 2: mean 5.95 (SD 3.2); n=41; WOMAC 0-20 Top=High is poor outcome; Comments: Baseline unsupervised: 9.74 (3.26). Baseline supervised: 11.52 (3.25). Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in education level, and baseline Kellgren Lawrence level; Group 1 Number missing: 00; Group 2 Number missing: 00	
Protocol outcomes not reported by the study	Health related quality of life at > 3 months; Physical function at > 3 months; Pain at > 3 months; Osteoarthritis flares at </=3 months; Osteoarthritis flares at > 3 months; Psychological distress at </=3 months; Psychological distress at > 3 months; Serious adverse events at </=3 months; Serious adverse events at > 3 months

Appendix E – Forest plots

E.1 Supervised strength exercise compared to unsupervised strength exercise

Figure 2: Pain (VAS, 0-100, high is poor, final values) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Kuru colak 2017	39.58	27.9	39	50.09	44.4	39	2.4%	-10.51 [-26.97, 5.95]
Nambi 2020	18	4	19	38	4	18	97.6%	-20.00 [-22.58, -17.42]
Total (95% CI)			**58**			**57**	**100.0%**	**-19.77 [-22.32, -17.23]**

Heterogeneity: Chi² = 1.25, df = 1 (P = 0.26); I² = 20%
Test for overall effect: Z = 15.21 (P < 0.00001)

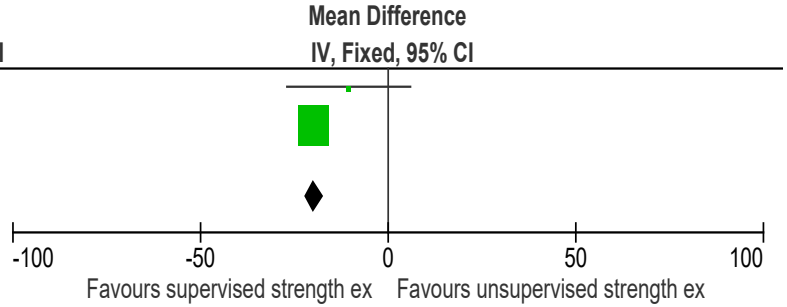
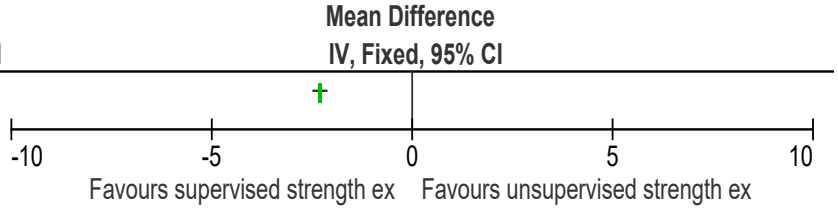


Figure 3: Pain (VAS, 0-10, high is poor, final value) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Nambi 2020	0.8	0.3	18	3.1	0.2	18	-2.30 [-2.47, -2.13]



E.2 Supervised strength exercise compared to supervised aerobic exercise

Figure 4: Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Supervised aerobic ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Bokaeian 2021	60.3	26.9	18	44.4	24.6	19	55.1%	0.60 [-0.06, 1.27]
Samut 2015	4	3	15	3.29	2.4	14	44.9%	0.25 [-0.48, 0.98]
Total (95% CI)			**33**			**33**	**100.0%**	**0.45 [-0.04, 0.94]**

Heterogeneity: Chi² = 0.49, df = 1 (P = 0.48); I² = 0%

Test for overall effect: Z = 1.78 (P = 0.07)

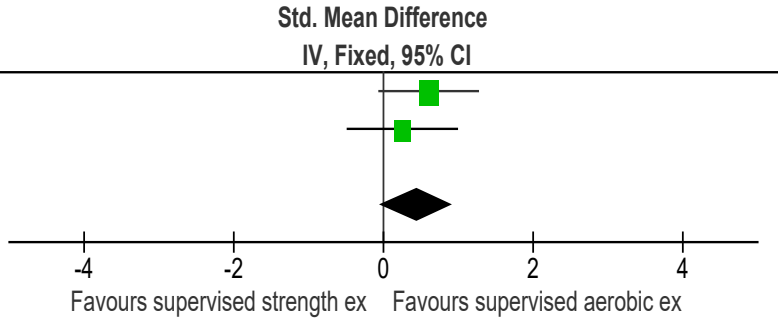


Figure 5: Pain (Arthritis Self-Efficacy pain subscale, 0-100, high is poor, change score) at >3 months

Study or Subgroup	Mean Difference	SE	Supervised strength ex Total	Supervised aerobic ex Total	Mean Difference IV, Fixed, 95% CI
Bieler 2017	11.1	5.6123	50	50	11.10 [0.10, 22.10]

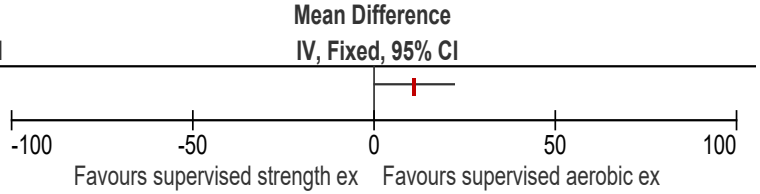


Figure 6: Pain (Arthritis Self-Efficacy pain subscale, 0-100, high is poor, change score) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Supervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Beckwee 2017	7.93	9.77	14	10.1	7.27	16	-2.17 [-8.41, 4.07]

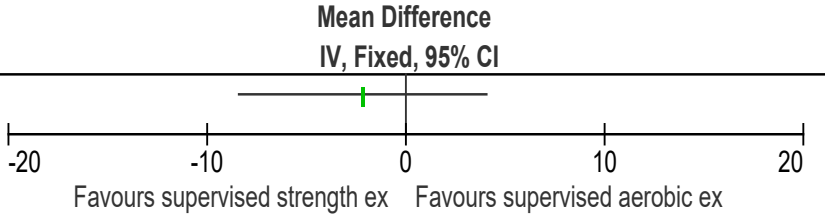


Figure 7: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Supervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Samut 2015	16.08	11.27	15	14.57	11.74	14	1.51 [-6.88, 9.90]

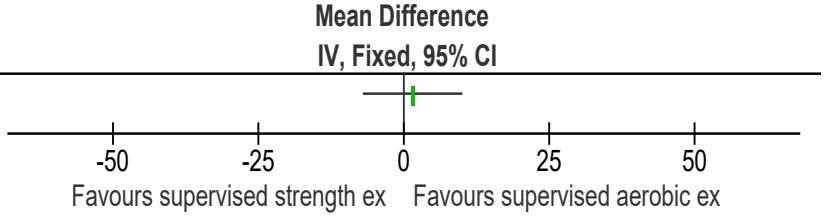
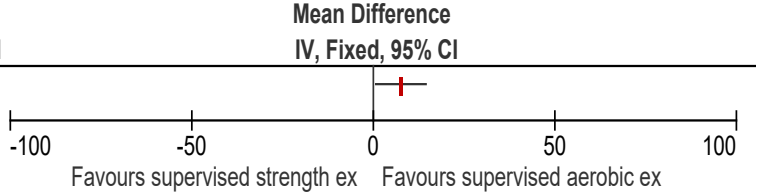


Figure 8: Physical function (Arthritis Self-Efficacy function subscale, 0-100, high is poor, change score) at >3 months

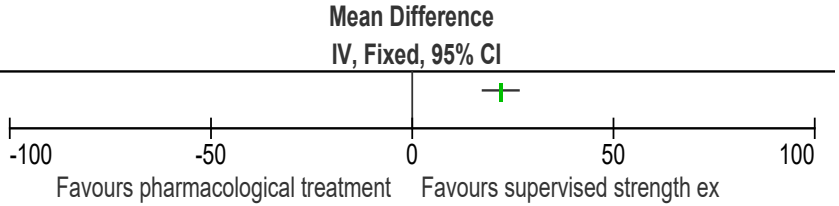
Study or Subgroup	Mean Difference	SE	Supervised strength ex Total	Supervised aerobic ex Total	Mean Difference IV, Fixed, 95% CI
Bieler 2017	7.6	3.5205	50	50	7.60 [0.70, 14.50]



E.3 Supervised strength exercise compared to pharmacological treatment

Figure 9: Quality of life (SF-36 total, scale range unclear, high is good, final values) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Chao 2020	105.4	21.5	92	83.4	4.2	74	22.00 [17.50, 26.50]



E.4 Supervised strength exercise compared to no treatment

Figure 10: Quality of life (KOOS, 0-100, high is good, change scores) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Henriksen 2014	5.8	14.2	25	-0.3	14.3	23	52.8%	6.10 [-1.97, 14.17]
Park 2021	21.53	30.66	25	-5.4	9.06	25	47.2%	26.93 [14.40, 39.46]
Total (95% CI)			**50**			**48**	**100.0%**	**15.94 [-4.44, 36.32]**

Heterogeneity: Tau² = 188.02; Chi² = 7.50, df = 1 (P = 0.006); I² = 87%
Test for overall effect: Z = 1.53 (P = 0.13)

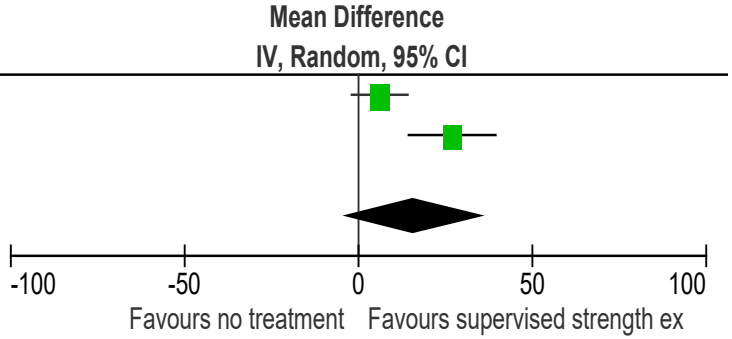


Figure 11: Quality of life (EQ-5D, KOOS, HOOS, Assessment of Quality of Life Scale [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Bautch 1997	-23.37	9.605	15	-17.88	7.165	15	14.3%	-0.63 [-1.37, 0.11]
Hermann 2016	38.8	17.2	40	31.2	13.9	40	19.3%	0.48 [0.04, 0.93]
Kigozi 2018	0.708	0.188	176	0.686	0.201	175	22.9%	0.11 [-0.10, 0.32]
Nahayatbin 2018	57.31	19.39	16	40	15.24	16	14.3%	0.97 [0.23, 1.71]
Nejati 2015	39.4	3.26	28	35.74	3.26	28	17.2%	1.11 [0.54, 1.67]
Pazit 2018	0.72	0.19	10	0.63	0.12	10	12.0%	0.54 [-0.35, 1.44]
Total (95% CI)			**285**			**284**	**100.0%**	**0.42 [-0.01, 0.86]**

Heterogeneity: Tau² = 0.21; Chi² = 20.99, df = 5 (P = 0.0008); I² = 76%

Test for overall effect: Z = 1.89 (P = 0.06)

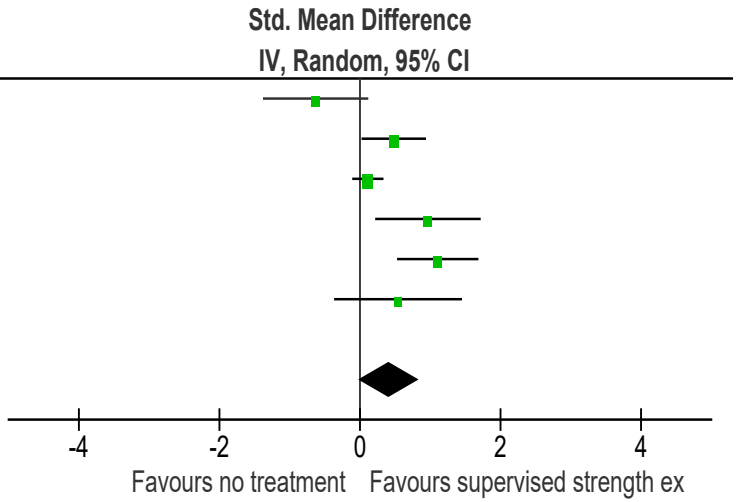


Figure 12: Quality of life (SF-36 physical component summary, SF-12 physical score, 0-100, high is good, change score and final values) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Bruce-brand 2012	53.2	25.09	10	67.83	21.71	6	16.5%	-14.63 [-37.94, 8.68]
Foley 2003	31.4	12.7	35	28.8	11	35	30.8%	2.60 [-2.97, 8.17]
Salli 2010	59.2	16.3	47	38.4	9.5	24	30.5%	20.80 [14.79, 26.81]
Thorstensson 2005	3	25.84	30	0.3	37.9	31	22.1%	2.70 [-13.53, 18.93]
Total (95% CI)			**122**			**96**	**100.0%**	**5.33 [-8.19, 18.85]**

Heterogeneity: Tau² = 146.46; Chi² = 24.23, df = 3 (P < 0.0001); I² = 88%
Test for overall effect: Z = 0.77 (P = 0.44)

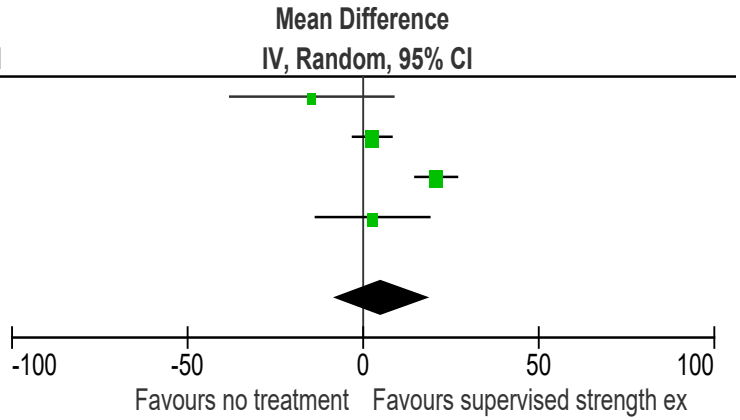


Figure 13: Quality of life (SF-36 mental component summary, SF-12 mental score, 0-100, high is good, change score and final values) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Bruce-brand 2012	65.3	24.91	10	70.5	22.4	6	10.5%	-5.20 [-28.86, 18.46]
Foley 2003	57.9	19.5	35	50.5	14	35	34.7%	7.40 [-0.55, 15.35]
Salli 2010	68.3	12.4	47	50.9	12.5	24	39.4%	17.40 [11.27, 23.53]
Thorstensson 2005	1.6	34.28	30	-2.1	38.71	31	15.3%	3.70 [-14.63, 22.03]
Total (95% CI)			**122**			**96**	**100.0%**	**9.45 [0.79, 18.11]**

Heterogeneity: Tau² = 39.74; Chi² = 7.06, df = 3 (P = 0.07); I² = 58%
Test for overall effect: Z = 2.14 (P = 0.03)

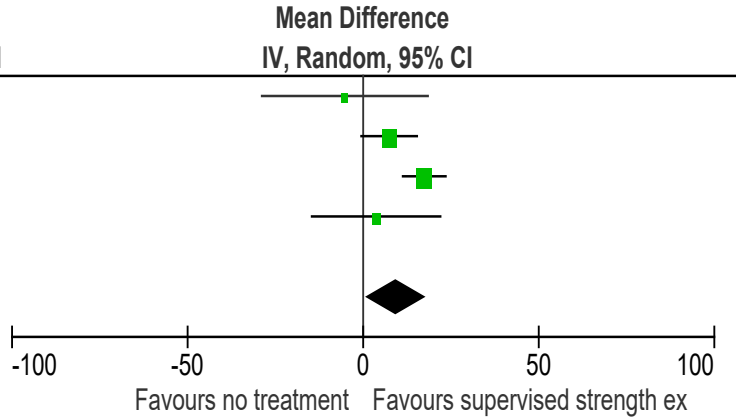


Figure 14: Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Imoto 2012	20.28	27.42	50	6.96	26.79	50	46.6%	13.32 [2.69, 23.95]
Jorge 2015	49.8	21.9	29	30.8	16.8	31	53.4%	19.00 [9.07, 28.93]
Total (95% CI)			**79**			**81**	**100.0%**	**16.35 [9.10, 23.61]**

Heterogeneity: Chi² = 0.59, df = 1 (P = 0.44); I² = 0%
Test for overall effect: Z = 4.42 (P < 0.00001)

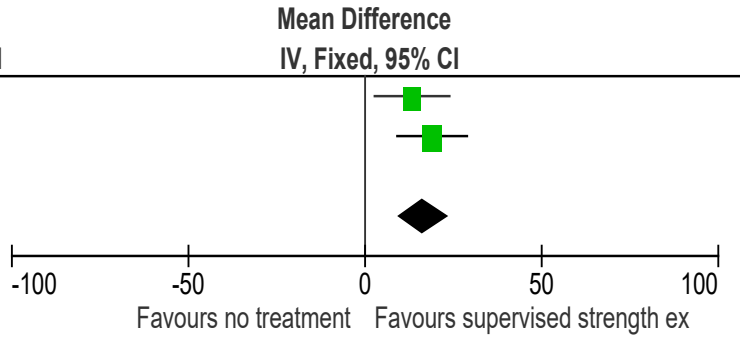


Figure 15: Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Imoto 2012	16.4	36.85	50	6.14	41.16	50	36.6%	10.26 [-5.05, 25.57]
Jorge 2015	58.6	25	29	41.7	20.6	31	63.4%	16.90 [5.26, 28.54]
Total (95% CI)			**79**			**81**	**100.0%**	**14.47 [5.21, 23.73]**

Heterogeneity: Chi² = 0.46, df = 1 (P = 0.50); I² = 0%
Test for overall effect: Z = 3.06 (P = 0.002)

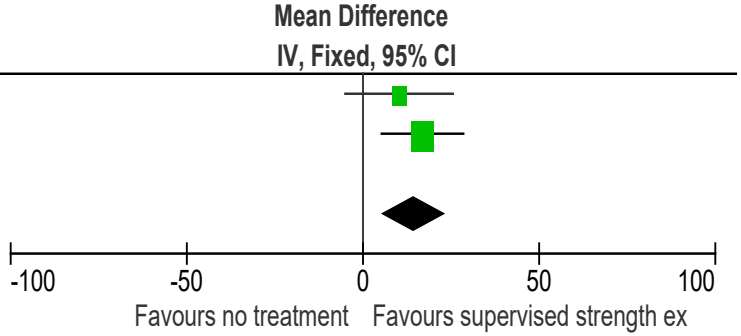


Figure 16: Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Imoto 2012	27.85	67.5	50	13.39	64.9	50	30.8%	14.46 [-11.49, 40.41]
Jorge 2015	48.3	41.7	29	16.9	23.6	31	69.2%	31.40 [14.10, 48.70]
Total (95% CI)			**79**			**81**	**100.0%**	**26.19 [11.79, 40.58]**

Heterogeneity: Chi² = 1.13, df = 1 (P = 0.29); I² = 12%
Test for overall effect: Z = 3.57 (P = 0.0004)

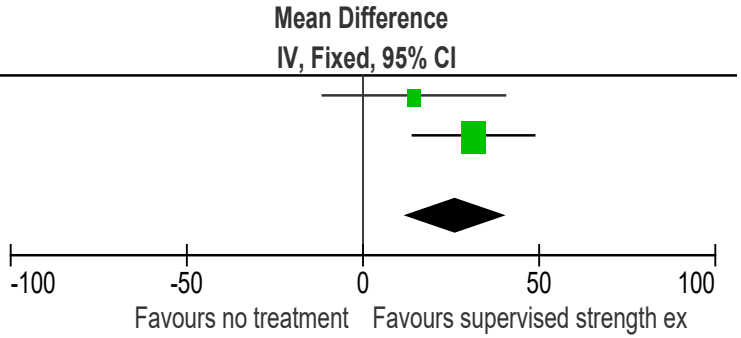


Figure 17: Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Imoto 2012	10	37.9	50	3.17	40.62	50	37.2%	6.83 [-8.57, 22.23]
Jorge 2015	64	25.2	29	52.4	21.3	31	62.8%	11.60 [-0.25, 23.45]
Total (95% CI)			**79**			**81**	**100.0%**	**9.83 [0.44, 19.22]**

Heterogeneity: Chi² = 0.23, df = 1 (P = 0.63); I² = 0%
Test for overall effect: Z = 2.05 (P = 0.04)

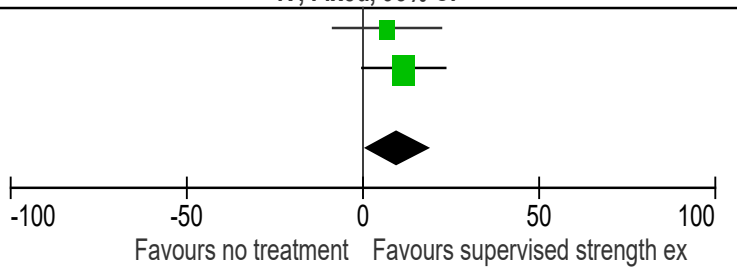


Figure 18: Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Imoto 2012	8.05	24.68	50	5.89	26.1	50	52.3%	2.16 [-7.80, 12.12]
Jorge 2015	66.1	21.8	29	52.6	21.8	31	47.7%	13.50 [2.46, 24.54]
Total (95% CI)			**79**			**81**	**100.0%**	**7.57 [-3.53, 18.67]**

Heterogeneity: Tau² = 35.54; Chi² = 2.24, df = 1 (P = 0.13); I² = 55%
Test for overall effect: Z = 1.34 (P = 0.18)

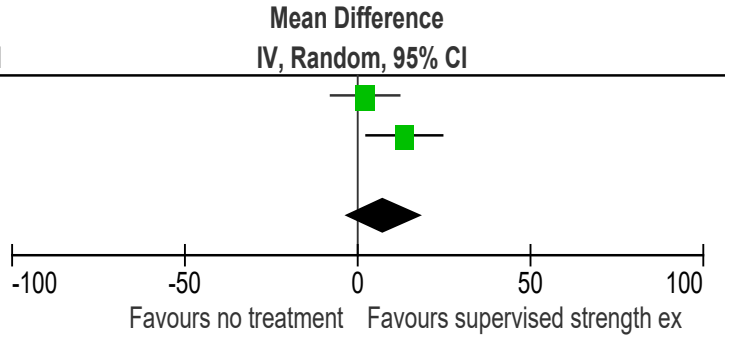


Figure 19: Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Imoto 2012	3.77	29.15	50	1.28	32.99	50	47.1%	2.49 [-9.71, 14.69]
Jorge 2015	76.4	18.7	29	59.5	21.2	31	52.9%	16.90 [6.80, 27.00]
Total (95% CI)			**79**			**81**	**100.0%**	**10.12 [-3.98, 24.22]**

Heterogeneity: Tau² = 71.17; Chi² = 3.18, df = 1 (P = 0.07); I² = 69%
Test for overall effect: Z = 1.41 (P = 0.16)

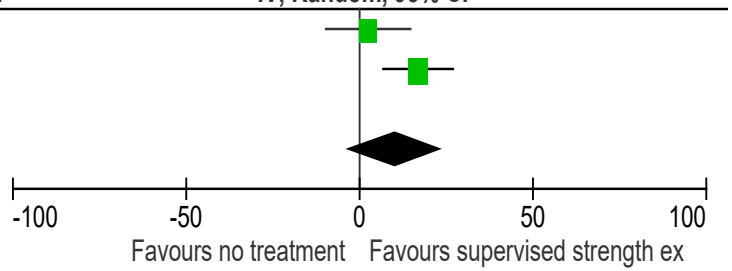


Figure 20: Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Imoto 2012	16.22	68.22	50	13.21	86.46	50	30.2%	3.01 [-27.52, 33.54]
Jorge 2015	72.4	39.9	29	49.5	39.3	31	69.8%	22.90 [2.84, 42.96]
Total (95% CI)			**79**			**81**	**100.0%**	**16.90 [0.14, 33.67]**

Heterogeneity: Chi² = 1.14, df = 1 (P = 0.29); I² = 12%

Test for overall effect: Z = 1.98 (P = 0.05)

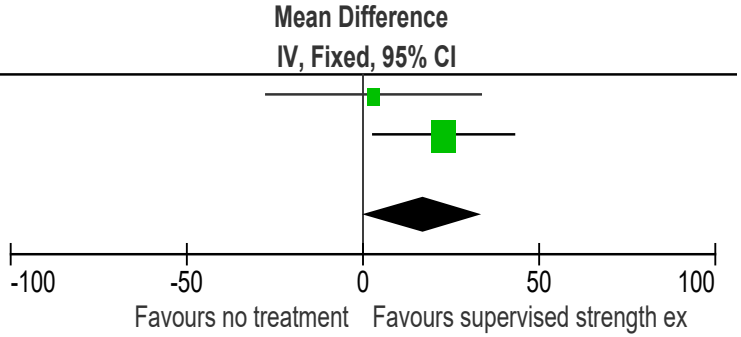


Figure 21: Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at ≤3 months

	Supervised strength ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Imoto 2012	9.57	37.27	50	0.35	51.7	50	39.9%	9.22 [-8.45, 26.89]
Jorge 2015	77.2	28.9	29	57.7	27.9	31	60.1%	19.50 [5.11, 33.89]
Total (95% CI)			**79**			**81**	**100.0%**	**15.40 [4.24, 26.56]**

Heterogeneity: Chi² = 0.78, df = 1 (P = 0.38); I² = 0%

Test for overall effect: Z = 2.71 (P = 0.007)

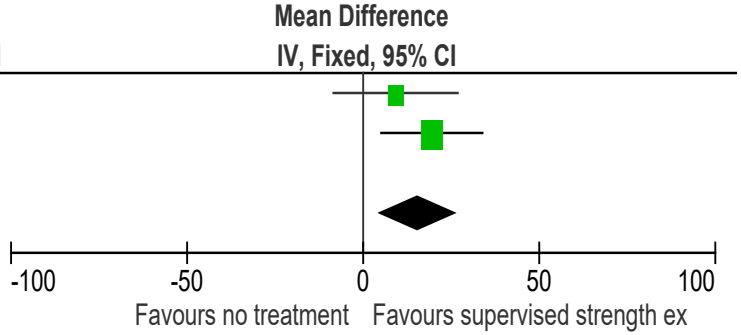


Figure 22: Quality of life (EQ-5D, KOOS [different scale ranges], high is good, final values) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Kigozi 2018	0.7	0.206	176	0.7	0.219	175	86.5%	0.00 [-0.21, 0.21]
Nejati 2015	30.26	18.7	28	38.21	18.7	28	13.5%	-0.42 [-0.95, 0.11]
Total (95% CI)			**204**			**203**	**100.0%**	**-0.06 [-0.25, 0.14]**

Heterogeneity: Chi² = 2.08, df = 1 (P = 0.15); I² = 52%
Test for overall effect: Z = 0.57 (P = 0.57)

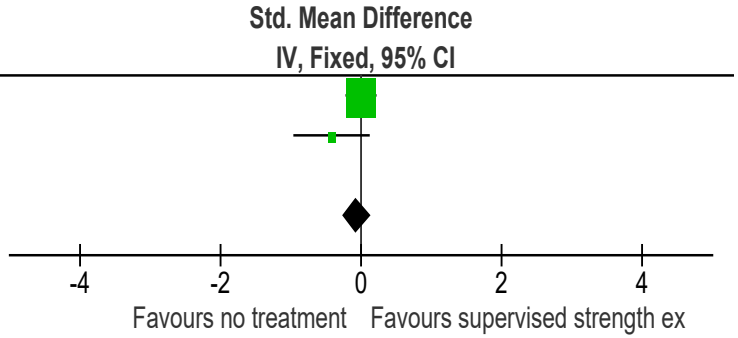


Figure 23: Quality of life (SF-36 physical component summary, 0-100, high is good, change scores) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Thorstensson 2005	3	29.73	30	-0.7	33.53	31	3.70 [-12.19, 19.59]

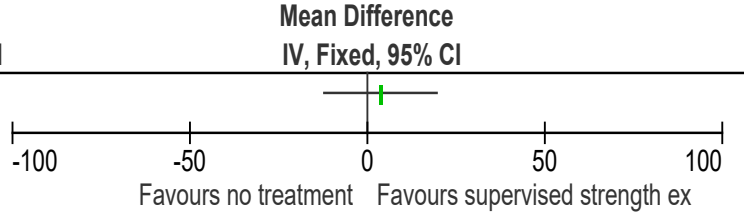


Figure 24: Quality of life (SF-36 mental component summary, 0-100, high is good, change scores) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Thorstensson 2005	0.7	41.91	30	-0.7	40.35	31	1.40 [-19.26, 22.06]

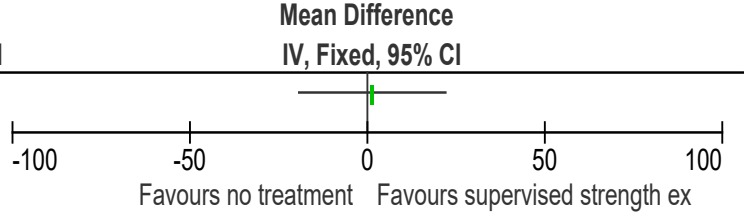


Figure 25: Pain (KOOS, WOMAC, NRS, VAS [different scale ranges], high is poor, change scores) at ≤3 months

	Supervised strength ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Anwer 2014	-4.81	0.1	21	-1.71	0.23	21	0.2%	-17.15 [-21.05, -13.25]
Henriksen 2014	-6.1	9.4	25	0.7	9.5	23	9.5%	-0.71 [-1.29, -0.12]
Huang 2005A	-1.2	1.6	35	-0.5	1.7	35	14.5%	-0.42 [-0.89, 0.05]
Huang 2005B	-1.2	1.4	30	-0.4	1.6	30	12.3%	-0.53 [-1.04, -0.01]
Imoto 2012	-3.17	3.84	50	-0.88	3.73	50	20.2%	-0.60 [-1.00, -0.20]
Oliveira 2012	-3.87	4.15	50	-1.05	4.65	50	20.1%	-0.63 [-1.04, -0.23]
Park 2021	-3.57	7	25	0.63	9.78	25	10.3%	-0.49 [-1.05, 0.08]
Thorstensson 2005	-1.8	13.39	30	0.3	16.22	31	12.9%	-0.14 [-0.64, 0.36]
Total (95% CI)			**266**			**265**	**100.0%**	**-0.55 [-0.73, -0.37]**

Heterogeneity: Chi² = 73.03, df = 7 (P < 0.00001); I² = 90%

Test for overall effect: Z = 5.93 (P < 0.00001)

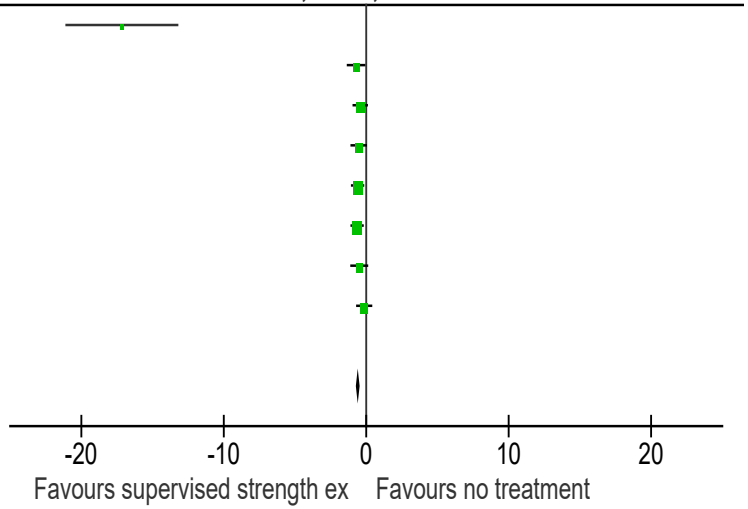


Figure 26: Pain (KOOS, HOOS, AUSCAN, WOMAC, NRS, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Anon 2016	3.2	1.9	19	7.2	5.1	16	4.1%	-1.05 [-1.77, -0.34]
Bautch 1997	2.19	1.6654	15	2.08	2.0914	15	4.1%	0.06 [-0.66, 0.77]
Borjesson 1996	3	1.5	34	3.3	1.5	34	5.1%	-0.20 [-0.67, 0.28]
Bruce-brand 2012	9.6	4.14	10	8.33	4.08	6	3.0%	0.29 [-0.73, 1.31]
Foley 2003	8	5	35	10	4	35	5.1%	-0.44 [-0.91, 0.04]
Hermann 2016	-55.4	16.9	40	-45.9	14.1	40	5.2%	-0.60 [-1.05, -0.16]
Huang 2003	3.1	1	99	4.2	0.4	33	5.3%	-1.23 [-1.65, -0.81]
Jan 2008	4.8	3.1	68	7.1	3.4	30	5.2%	-0.71 [-1.16, -0.27]
Jorge 2015	4.9	3.2	29	9.5	3.2	31	4.7%	-1.42 [-1.99, -0.85]
Kang 2019	42.07	5.26	15	56.5	6.19	14	3.0%	-2.45 [-3.44, -1.45]
Kuptniratsaikul 2002	4.14	2.28	193	5.15	2.26	199	6.0%	-0.44 [-0.64, -0.24]
Lin 2009	4.2	3	36	7.3	3.4	36	5.0%	-0.96 [-1.45, -0.47]
Nahayatbin 2018	-58.44	9.51	16	-50.31	10.77	16	4.0%	-0.78 [-1.50, -0.06]
Nejati 2015	-63.39	19.3	29	-46.65	19.3	28	4.8%	-0.86 [-1.40, -0.31]
Nery 2021	4.97	4.07	30	8.23	4.42	30	4.8%	-0.76 [-1.28, -0.23]
Pazit 2018	117	132.6	10	249.7	309.3	10	3.4%	-0.53 [-1.43, 0.36]
Rezasoltani 2020	-80.8	1.8	15	-55.3	3.5	15	0.8%	-8.92 [-11.44, -6.39]
Rosedale 2014	-56	17	120	-46	16	60	5.7%	-0.60 [-0.91, -0.28]
Salli 2010	3.4	1.9	47	6.5	1.8	24	4.7%	-1.64 [-2.21, -1.08]
Samut 2015	4	3	15	7.31	2.84	13	3.7%	-1.10 [-1.90, -0.29]
Sayers 2012	9.8	3.1	22	10.2	2.5	11	4.0%	-0.13 [-0.86, 0.59]
Williamson 2007	6.36	2.6	60	7.24	2.07	61	5.5%	-0.37 [-0.73, -0.01]
Wortley 2013	71	100	13	157	96	6	3.0%	-0.83 [-1.84, 0.18]
Total (95% CI)			**970**			**763**	**100.0%**	**-0.81 [-1.06, -0.57]**

Heterogeneity: Tau² = 0.24; Chi² = 104.35, df = 22 (P < 0.00001); I² = 79%

Test for overall effect: Z = 6.58 (P < 0.00001)

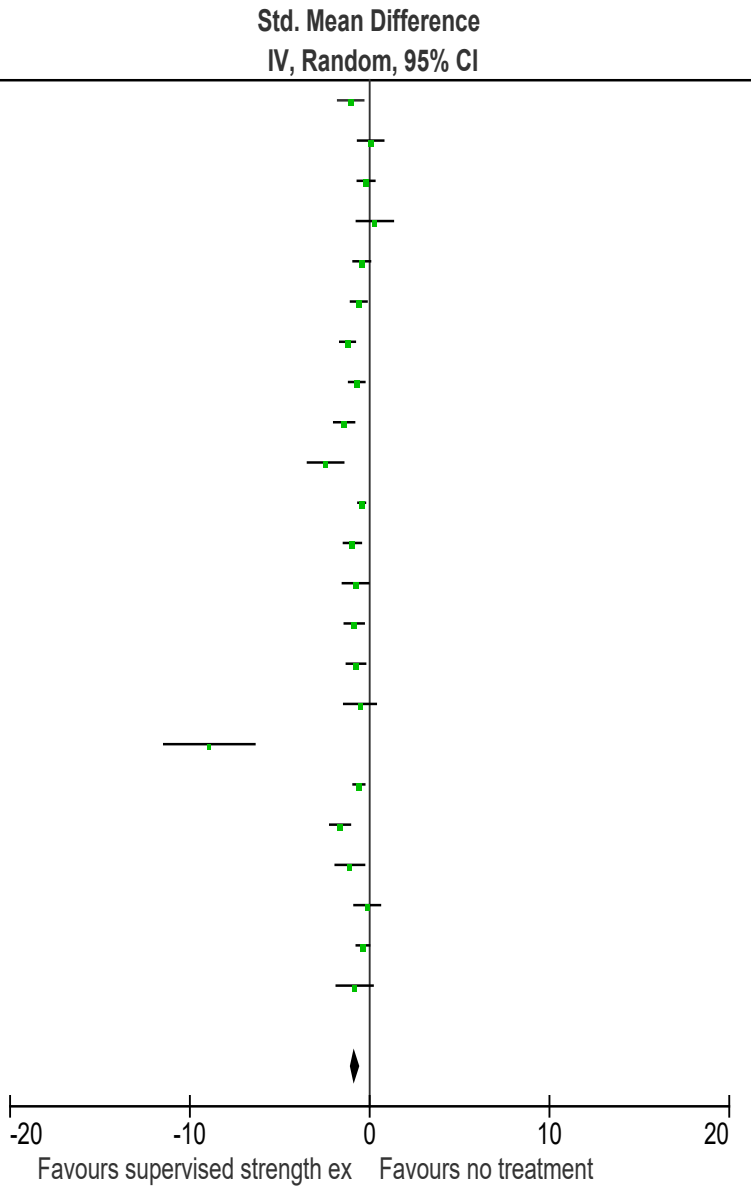


Figure 27: Pain (KOOS, VAS [different scale ranges], high is poor, change score and final values) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Huang 2003	2.6	1.7	99	6.1	1.3	33	14.4%	-2.16 [-2.63, -1.69]
Huang 2005A	3.9	1.4	35	6.6	1.5	35	14.1%	-1.84 [-2.40, -1.28]
Huang 2005B	3.5	1.7	30	6	1.3	30	14.0%	-1.63 [-2.22, -1.04]
Kuptniratsaikul 2002	4.25	2.7	193	4.57	2.69	199	15.1%	-0.12 [-0.32, 0.08]
Nejati 2015	-48.07	1.73	28	-49.03	1.73	28	14.2%	0.55 [0.01, 1.08]
Salli 2010	3.5	1.9	47	6.3	1.5	24	14.1%	-1.56 [-2.12, -1.00]
Thorstensson 2005	-3.1	13.52	30	1.1	14.99	31	14.3%	-0.29 [-0.80, 0.21]
Total (95% CI)			**462**			**380**	**100.0%**	**-1.00 [-1.76, -0.23]**

Heterogeneity: Tau² = 1.00; Chi² = 125.00, df = 6 (P < 0.00001); I² = 95%

Test for overall effect: Z = 2.56 (P = 0.01)

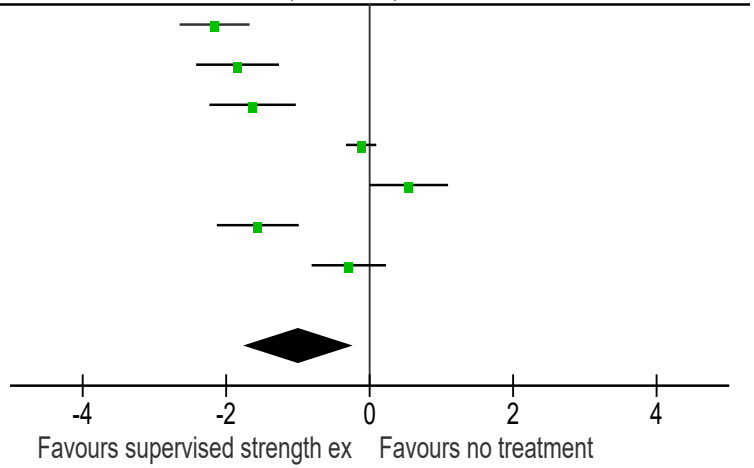


Figure 28: Physical function (KOOS, WOMAC [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Anwer 2014	-16.66	1.09	21	-6.47	0.13	21	0.9%	-12.88 [-15.84, -9.92]
Henriksen 2014	-4.2	11	25	-1.4	11	23	25.1%	-0.25 [-0.82, 0.32]
Oliveira 2012	-10.95	14.05	50	-1.97	16.58	50	50.7%	-0.58 [-0.98, -0.18]
Park 2021	-4.21	5.04	25	1.52	6.12	25	23.3%	-1.01 [-1.60, -0.41]
Total (95% CI)			**121**			**119**	**100.0%**	**-0.71 [-1.00, -0.43]**

Heterogeneity: Chi² = 68.99, df = 3 (P < 0.00001); I² = 96%

Test for overall effect: Z = 4.89 (P < 0.00001)

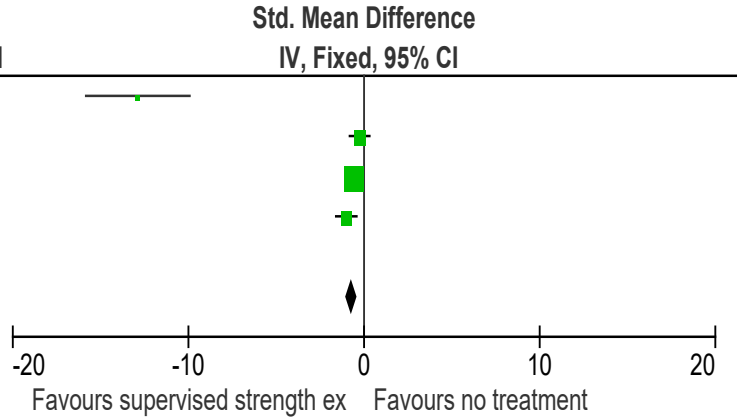


Figure 29: Physical function (KOOS, HOOS, AUSCAN, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at ≤3 months

	Supervised strength ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Anon 2016	6.7	3.7	19	19.8	13.8	16	5.2%	-1.32 [-2.06, -0.58]
Bruce-brand 2012	31.5	14.4	10	21.67	18.9	6	4.3%	0.58 [-0.46, 1.61]
Foley 2003	27	12	35	37	13	35	5.9%	-0.79 [-1.28, -0.30]
Hermann 2016	-59.9	17.1	40	-48.7	13.9	40	6.0%	-0.71 [-1.16, -0.26]
Jan 2008	14.8	8.9	68	22.5	10.9	30	6.0%	-0.80 [-1.24, -0.36]
Jorge 2015	17.3	12.4	29	26.7	10.2	31	5.8%	-0.82 [-1.35, -0.29]
Kang 2019	50.93	7.01	15	56.64	5.26	14	5.1%	-0.89 [-1.66, -0.12]
Kuptniratsaikul 2002	6.08	3.14	193	6.38	3.58	199	6.5%	-0.09 [-0.29, 0.11]
Lin 2009	10.1	8.3	36	24.9	11.8	36	5.8%	-1.44 [-1.96, -0.91]
Nahayatbin 2018	-59	10.25	16	-61.31	10.39	16	5.4%	0.22 [-0.48, 0.91]
Nejati 2015	-64.99	3.37	28	-50.81	3.37	28	4.6%	-4.15 [-5.10, -3.20]
Nery 2021	8.77	7.4	30	13.8	7.42	30	5.8%	-0.67 [-1.19, -0.15]
Pazit 2018	277.8	237	10	565.7	282.5	10	4.6%	-1.06 [-2.01, -0.11]
Rezasoltani 2020	-80.2	1.8	15	-58.4	3.5	15	2.0%	-7.62 [-9.81, -5.43]
Rosedale 2014	-61	17	120	-52	16	60	6.3%	-0.54 [-0.85, -0.22]
Salli 2010	20.8	10.2	47	32.6	11.6	24	5.8%	-1.09 [-1.62, -0.57]
Samut 2015	16.08	11.27	15	29.92	11.25	13	5.0%	-1.19 [-2.01, -0.38]
Sayers 2012	29.7	10.2	22	34.8	13.9	11	5.2%	-0.43 [-1.16, 0.30]
Wortley 2013	240	249	13	475	282	6	4.4%	-0.87 [-1.88, 0.15]
Total (95% CI)			**761**			**620**	**100.0%**	**-1.00 [-1.37, -0.63]**

Heterogeneity: Tau² = 0.53; Chi² = 151.81, df = 18 (P < 0.00001); I² = 88%

Test for overall effect: Z = 5.34 (P < 0.00001)

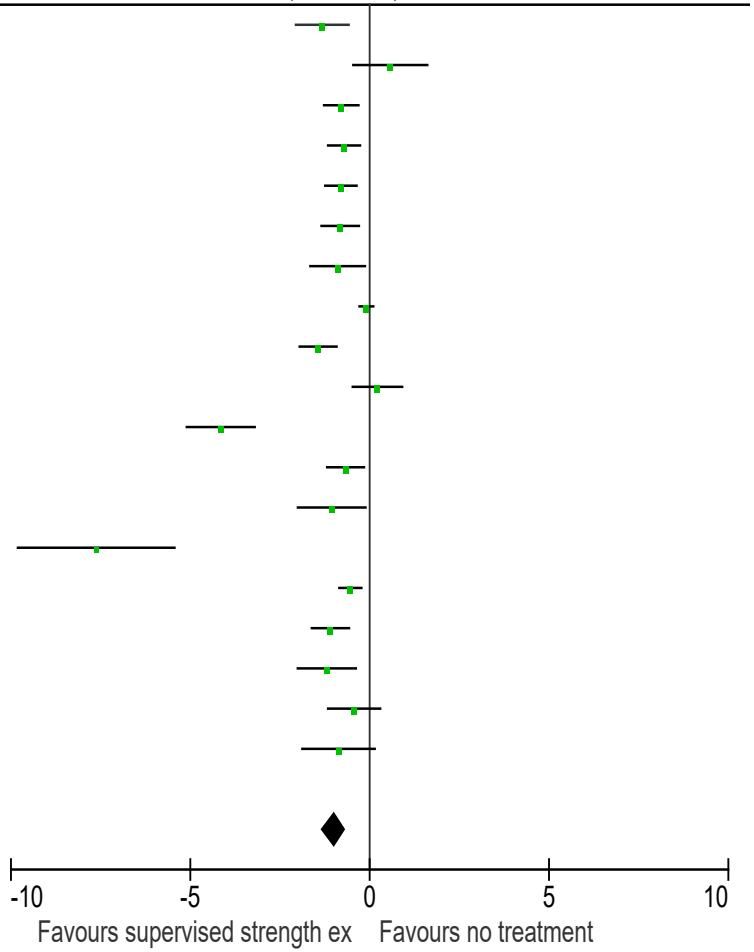


Figure 30: Physical function (KOOS ADL, 0-100, high is good, change score) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Thorstensson 2005	0.9	12.59	30	-1.9	15.81	31	2.80 [-4.36, 9.96]

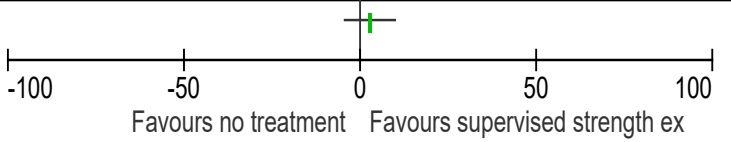


Figure 31: Physical function (KOOS, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at >3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Kuptniratsaikul 2002	5.28	3.46	193	6.32	3.63	199	36.4%	-0.29 [-0.49, -0.09]
Nejati 2015	-46.98	20.3	28	-58.88	20.3	28	31.8%	0.58 [0.04, 1.11]
Salli 2010	19.6	10.5	47	32.7	11.3	24	31.8%	-1.20 [-1.73, -0.67]
Total (95% CI)			**268**			**251**	**100.0%**	**-0.31 [-1.09, 0.48]**

Heterogeneity: Tau² = 0.43; Chi² = 21.39, df = 2 (P < 0.0001); I² = 91%
Test for overall effect: Z = 0.76 (P = 0.45)

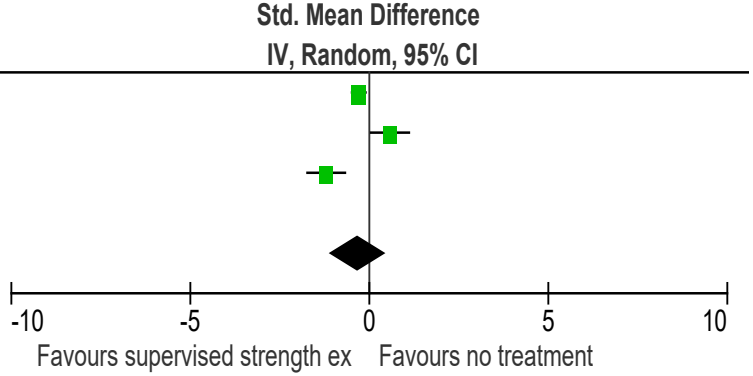


Figure 32: Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Williamson 2007	7.08	5.16	60	6.54	3.93	61	0.54 [-1.10, 2.18]

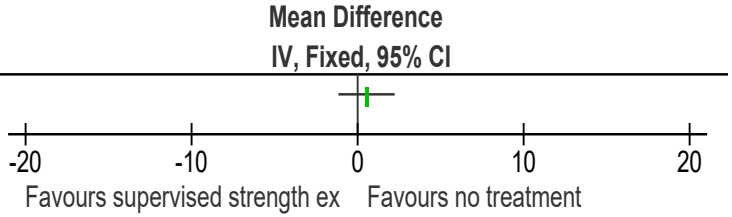


Figure 33: Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Williamson 2007	6.75	3.84	60	7.13	3.54	61	-0.38 [-1.70, 0.94]

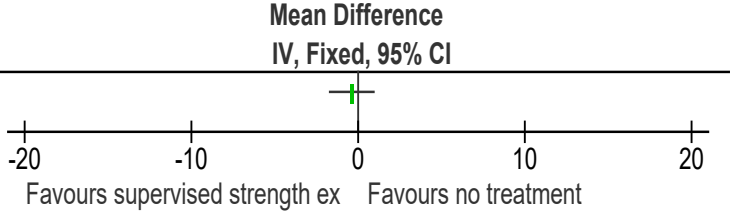
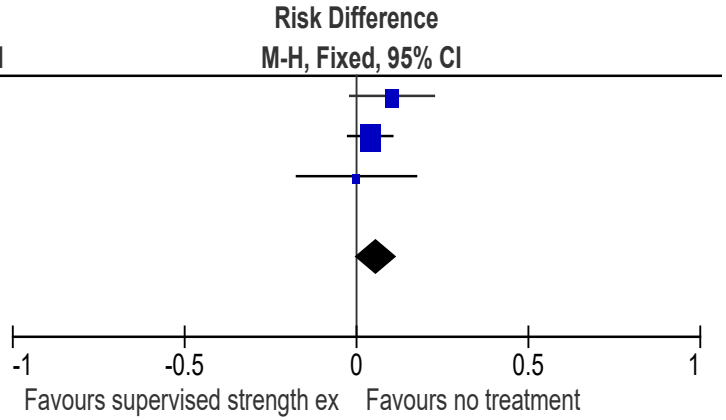


Figure 34: Serious adverse events at ≤3 months

Study or Subgroup	Supervised strength ex Events	Total	No treatment Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Jorge 2015	3	29	0	31	33.3%	0.10 [-0.02, 0.23]
Oliveira 2012	2	50	0	50	55.6%	0.04 [-0.03, 0.11]
Pazit 2018	0	10	0	10	11.1%	0.00 [-0.17, 0.17]
Total (95% CI)		**89**		**91**	**100.0%**	**0.06 [-0.00, 0.12]**
Total events	5		0			

Heterogeneity: Chi² = 1.22, df = 2 (P = 0.54); I² = 0%

Test for overall effect: Z = 1.85 (P = 0.06)



E.5 Unsupervised strength exercise compared to unsupervised aerobic exercise

Figure 35: Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	Unsupervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	29.5	4.8	27	8.6	4	28	20.90 [18.56, 23.24]

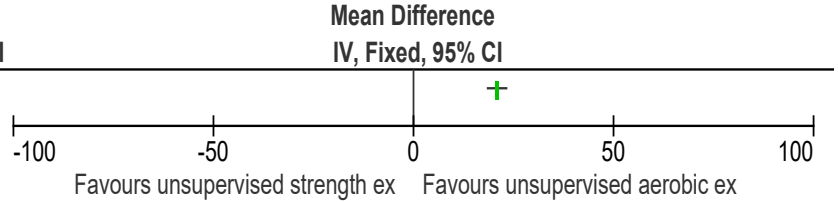


Figure 36: Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	Unsupervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	9.8	3.1	27	9	3.3	28	0.80 [-0.89, 2.49]

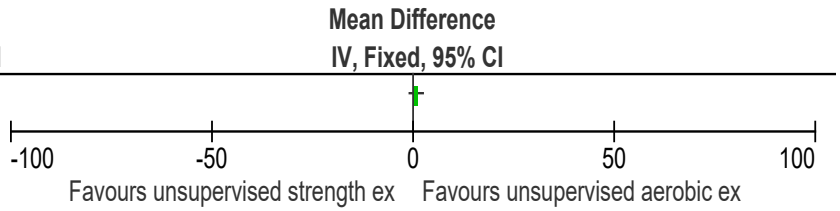


Figure 37: Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	Unsupervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	33.4	2.1	27	14.6	1.3	28	18.80 [17.87, 19.73]

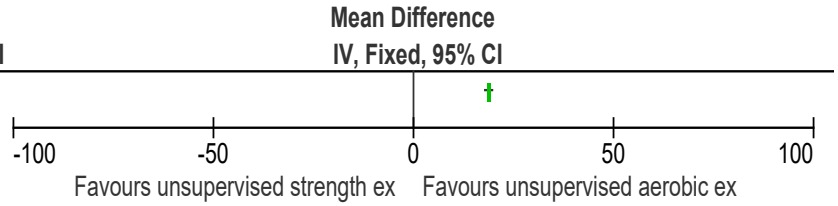


Figure 38: Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	Unsupervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	31.9	4.9	27	19.6	4	28	12.30 [9.93, 14.67]

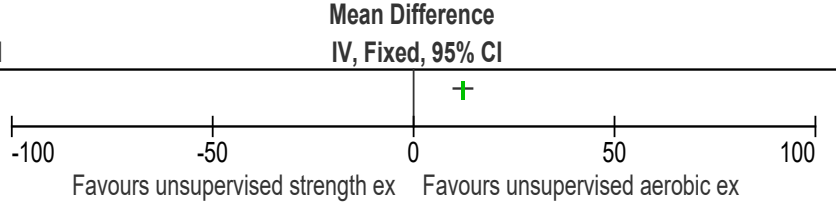


Figure 39: Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months

	Unsupervised strength ex			Unsupervised aerobic ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Evcik 2002	19.1	2.2	27	6.9	3	28	12.20 [10.81, 13.59]	

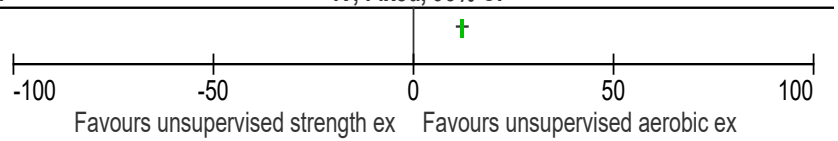


Figure 40: Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months

	Unsupervised strength ex			Unsupervised aerobic ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Evcik 2002	17.1	4.1	27	17.3	3.9	28	-0.20 [-2.32, 1.92]	

Figure 41: Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months

	Unsupervised strength ex			Unsupervised aerobic ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Evcik 2002	3	1.7	27	3.4	1.3	28	-0.40 [-1.20, 0.40]	

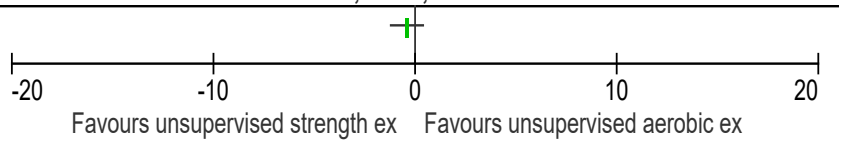
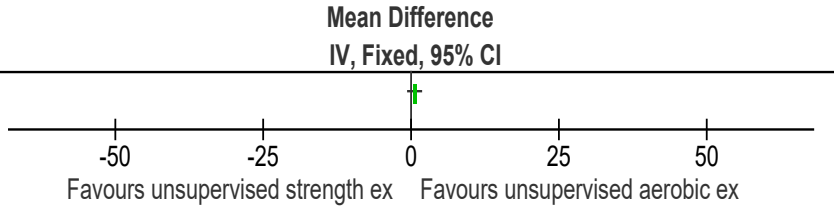


Figure 42: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	Unsupervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	10.8	1.8	27	10.2	2.4	28	0.60 [-0.52, 1.72]



E.6 Unsupervised strength exercise compared to no treatment

Figure 43: Quality of life (EQ-5D, Arthritis Impact Measurement Scale 2 - Short form [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chen 2019	82	9.96	71	77.9	9.52	70	50.5%	0.42 [0.08, 0.75]
Dziedzic 2015	0.66	0.22	65	0.665	0.24	65	49.5%	-0.02 [-0.37, 0.32]
Total (95% CI)			**136**			**135**	**100.0%**	**0.20 [-0.23, 0.63]**

Heterogeneity: Tau² = 0.07; Chi² = 3.24, df = 1 (P = 0.07); I² = 69%

Test for overall effect: Z = 0.91 (P = 0.36)

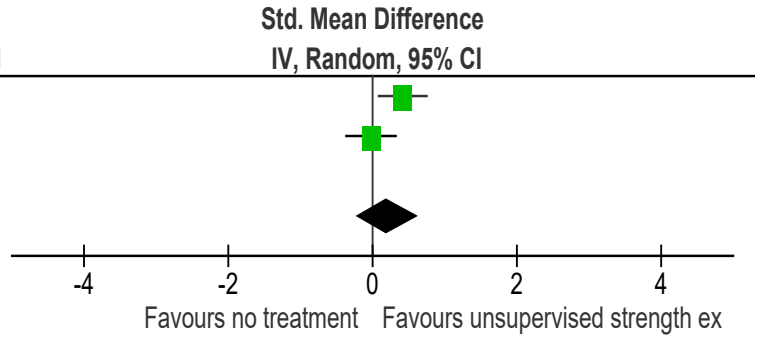


Figure 44: Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	29.5	4.8	27	36.6	6.1	26	-7.10 [-10.06, -4.14]

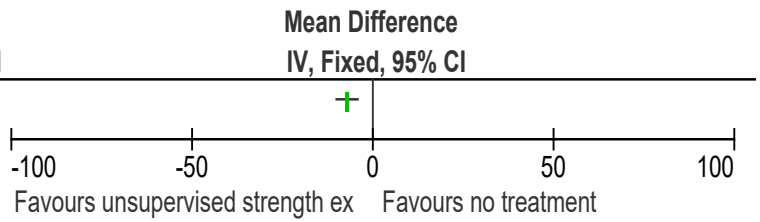


Figure 45: Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	9.8	3.1	27	20.4	3.2	26	-10.60 [-12.30, -8.90]

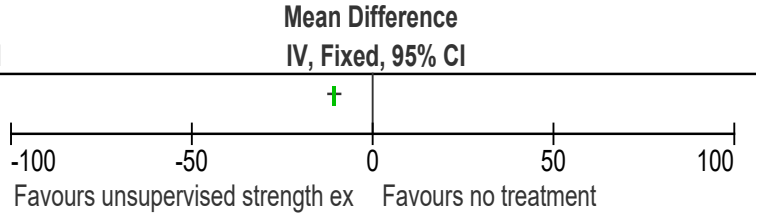


Figure 46: Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	33.4	2.1	27	49.3	1.7	26	-15.90 [-16.93, -14.87]

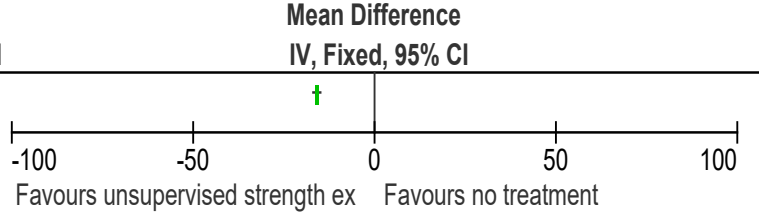


Figure 47: Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months

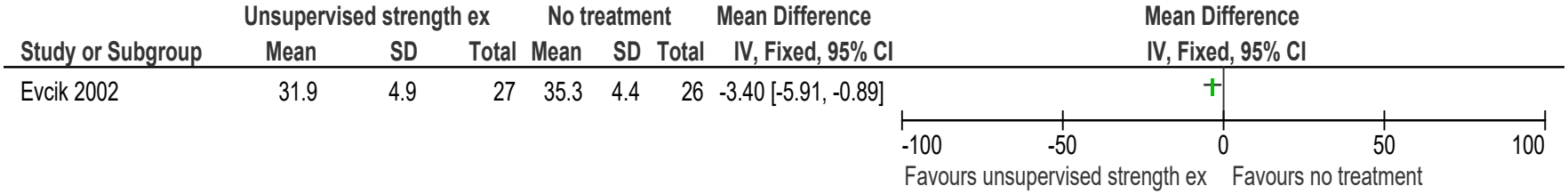


Figure 48: Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months

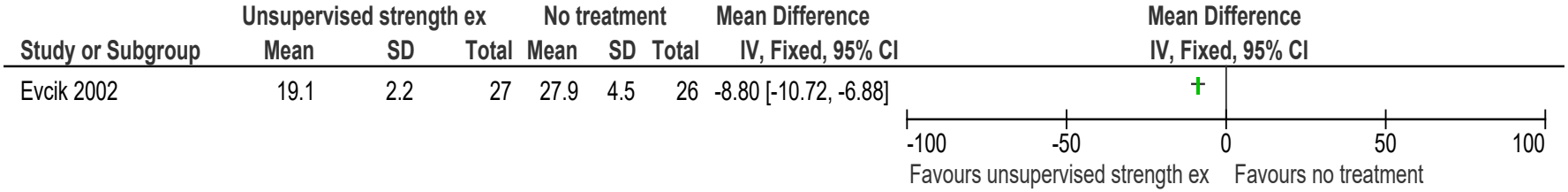


Figure 49: Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months

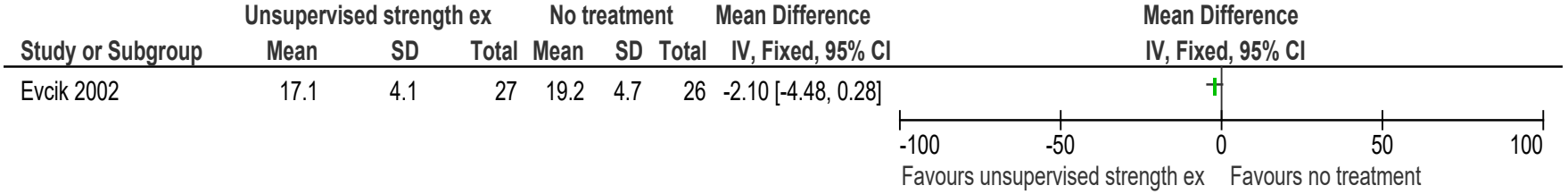


Figure 50: Quality of life (EQ-5D, 0-1, high is good, final value) at >3 months

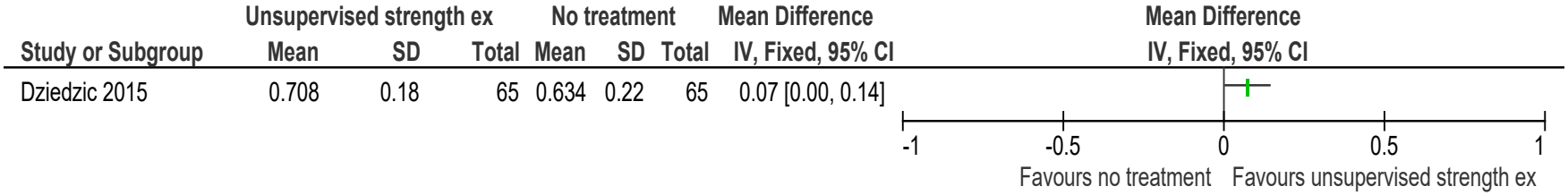


Figure 51: Quality of life (SF-36 physical functioning, 0-100, high is good, change score) at >3 months

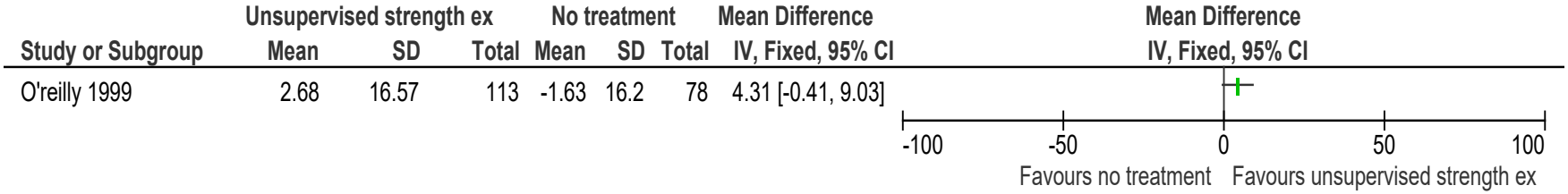


Figure 52: Quality of life (SF-36 bodily pain, 0-100, high is good, change score) at >3 months

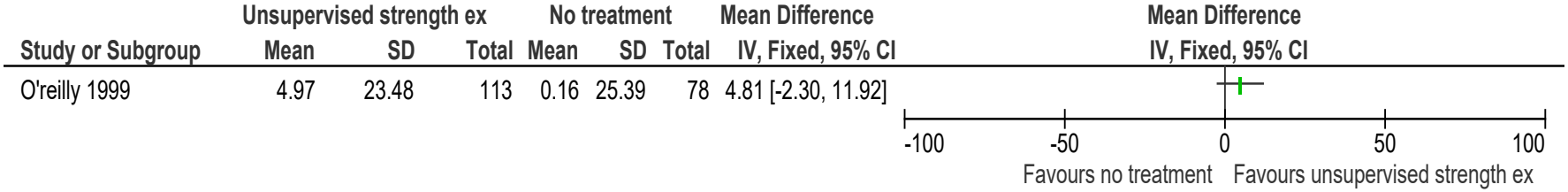


Figure 53: Quality of life (SF-36 role physical, 0-100, high is good, change score) at >3 months

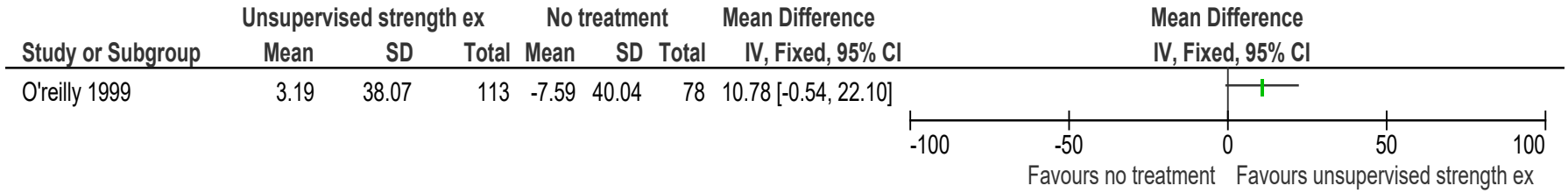


Figure 54: Quality of life (SF-36 vitality, 0-100, high is good, change score) at >3 months

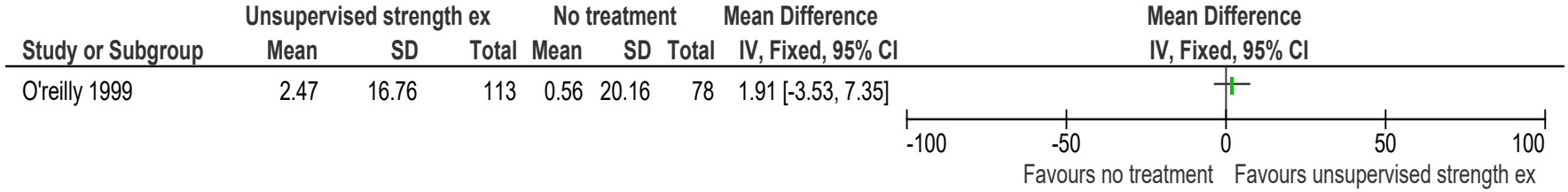


Figure 55: Quality of life (SF-36 general health, 0-100, high is good, change score) at >3 months

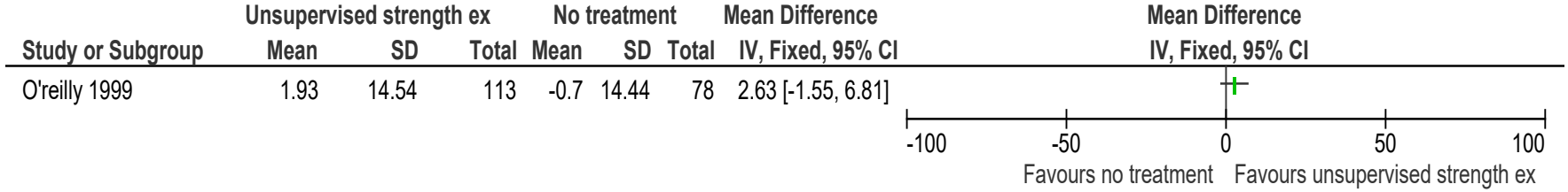


Figure 56: Quality of life (SF-36 mental health, 0-100, high is good, change score) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
O'reilly 1999	-0.21	13.86	113	-2.91	16.69	78	2.70 [-1.80, 7.20]

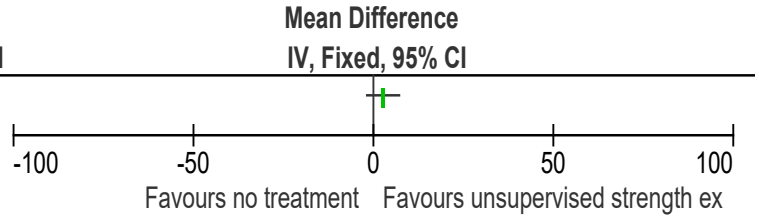


Figure 57: Quality of life (SF-36 role emotional, 0-100, high is good, change score) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
O'reilly 1999	1.85	46.15	113	0.48	62.34	78	1.37 [-14.87, 17.61]

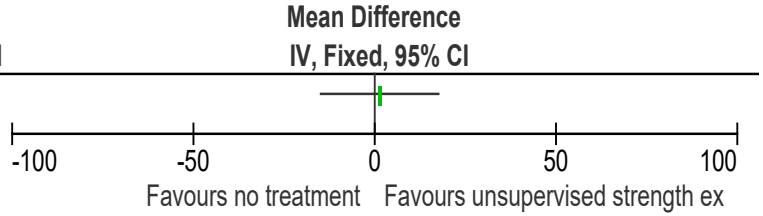


Figure 58: Quality of life (SF-36 social functioning, 0-100, high is good, change score) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
O'reilly 1999	1.89	25.79	113	1.9	41.12	78	-0.01 [-10.30, 10.28]

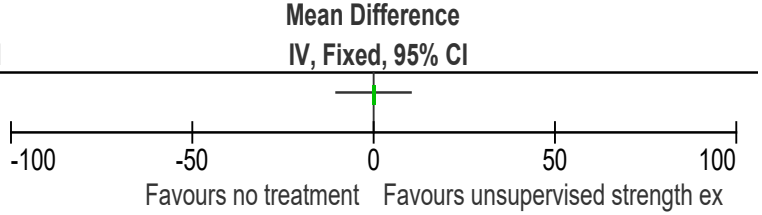


Figure 59: Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Bennell 2010	-2.6	2.6	45	-0.48	2.7	44	25.8%	-0.79 [-1.23, -0.36]
Chang 2012	-2.3	1.3	24	-0.9	1.5	17	11.0%	-0.99 [-1.65, -0.33]
Hennig 2015	-1.1	2.6	40	0.3	1.6	40	23.8%	-0.64 [-1.09, -0.19]
Karadag 2019	-6.2	1.12	30	-3.28	2.71	32	15.5%	-1.37 [-1.93, -0.82]
Lim 2008	-8.9	4.8	53	-1.9	2.9	54	23.9%	-1.76 [-2.20, -1.31]
Total (95% CI)			**192**			**187**	**100.0%**	**-1.10 [-1.32, -0.88]**

Heterogeneity: Chi² = 15.18, df = 4 (P = 0.004); I² = 74%
Test for overall effect: Z = 9.82 (P < 0.00001)

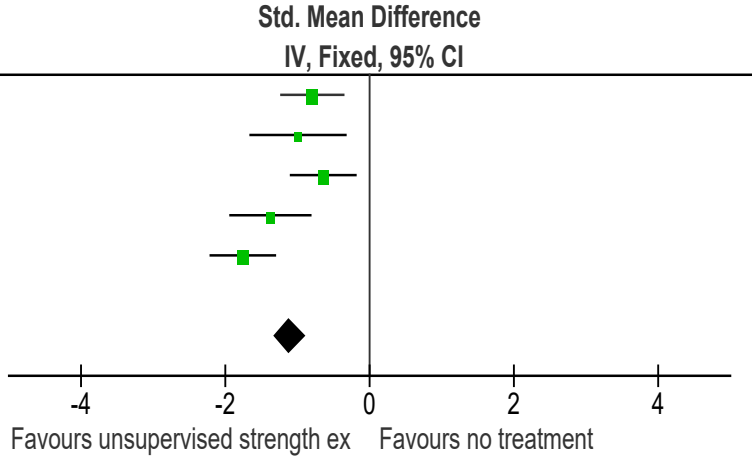


Figure 60: Pain (WOMAC, NRS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chen 2019	4.28	3.3	71	5.73	3.54	70	26.8%	-0.42 [-0.76, -0.09]
Evcik 2002	3	1.7	27	6	3.3	26	20.5%	-1.13 [-1.72, -0.55]
Juhakoski 2011	27.6	16.3	60	24.3	16.8	58	26.1%	0.20 [-0.16, 0.56]
Osteras 2014	3.7	2.1	65	4.3	2.1	65	26.5%	-0.28 [-0.63, 0.06]
Total (95% CI)			**223**			**219**	**100.0%**	**-0.37 [-0.81, 0.08]**

Heterogeneity: Tau² = 0.16; Chi² = 15.60, df = 3 (P = 0.001); I² = 81%
Test for overall effect: Z = 1.62 (P = 0.10)

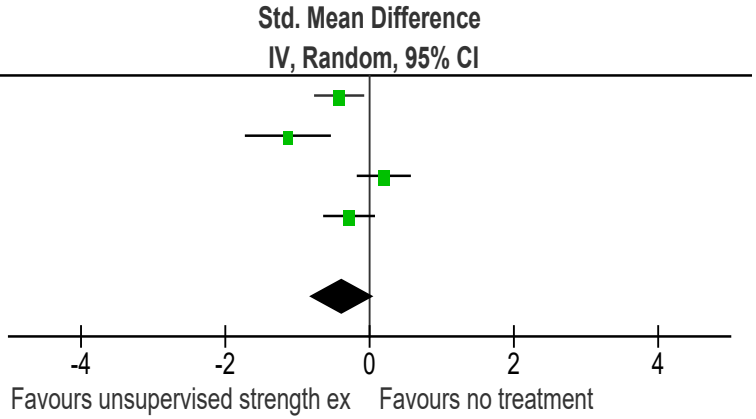


Figure 61: Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
O'reilly 1999	-1.45	3.2	113	0.42	3.02	78	10.6%	-0.60 [-0.89, -0.30]
Ravaud 2004	-19.7	28.7	735	-19.1	28.8	760	89.4%	-0.02 [-0.12, 0.08]
Total (95% CI)			**848**			**838**	**100.0%**	**-0.08 [-0.18, 0.01]**

Heterogeneity: Chi² = 13.05, df = 1 (P = 0.0003); I² = 92%
Test for overall effect: Z = 1.67 (P = 0.10)

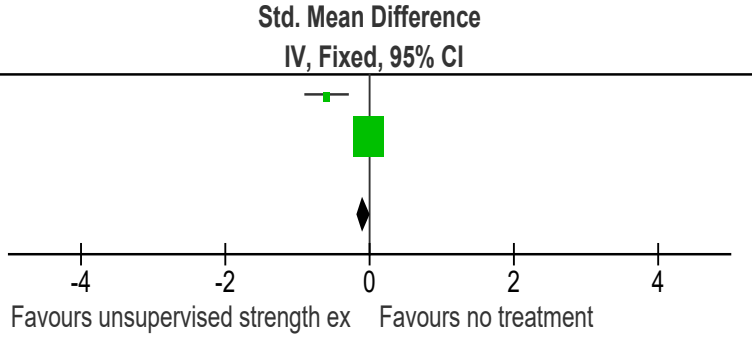


Figure 62: Pain (WOMAC, NRS, 0-100, high is poor, final values) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Juhakoski 2011	24.1	22.5	60	27.9	22.8	58	46.2%	-3.80 [-11.98, 4.38]
Osteras 2014	43	23	65	43	21	65	53.8%	0.00 [-7.57, 7.57]
Total (95% CI)			**125**			**123**	**100.0%**	**-1.75 [-7.31, 3.80]**

Heterogeneity: Chi² = 0.45, df = 1 (P = 0.50); I² = 0%
Test for overall effect: Z = 0.62 (P = 0.54)

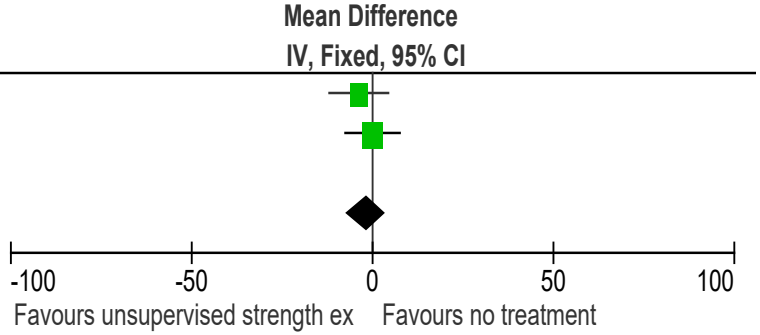


Figure 63: Physical function (WOMAC, FIHOA [different scale ranges], high is poor, change scores) at ≤3 months

	Unsupervised strength ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Bennell 2010	-8.07	7.7	45	-1.9	7.7	44	24.3%	-0.79 [-1.23, -0.36]
Chang 2012	-10.7	5.9	24	-4.5	4.4	17	10.0%	-1.14 [-1.81, -0.47]
Hennig 2015	-2.2	5.8	40	1.7	2.6	40	21.6%	-0.86 [-1.32, -0.40]
Karadag 2019	-16.97	4.69	30	-7.72	9.03	32	15.1%	-1.26 [-1.81, -0.71]
Lim 2008	-5.7	4.1	53	-2.9	2.2	54	29.0%	-0.85 [-1.24, -0.45]
Total (95% CI)			**192**			**187**	**100.0%**	**-0.93 [-1.14, -0.72]**

Heterogeneity: Chi² = 2.38, df = 4 (P = 0.67); I² = 0%
Test for overall effect: Z = 8.53 (P < 0.00001)

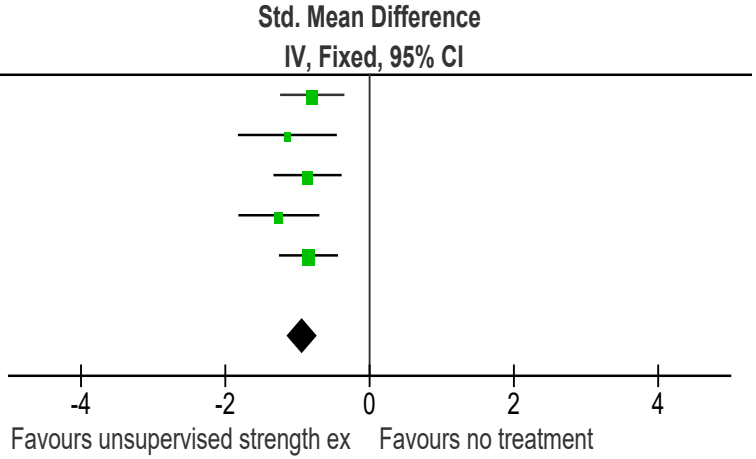


Figure 64: Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at ≤3 months

	Unsupervised strength ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Evcik 2002	10.8	1.8	27	20.7	4.4	26	31.2%	-2.92 [-3.71, -2.13]
Juhakoski 2011	27.4	13.9	60	25.9	14.5	58	34.4%	0.10 [-0.26, 0.47]
Osteras 2014	10.3	4.7	65	10	4.8	65	34.4%	0.06 [-0.28, 0.41]
Total (95% CI)			**152**			**149**	**100.0%**	**-0.85 [-2.15, 0.44]**

Heterogeneity: Tau² = 1.24; Chi² = 50.57, df = 2 (P < 0.00001); I² = 96%
Test for overall effect: Z = 1.29 (P = 0.20)

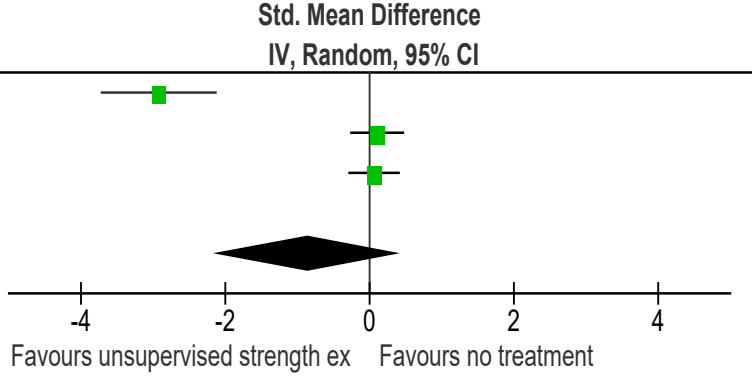


Figure 65: Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
O'reilly 1999	-3.55	9.74	113	-0.01	7.82	78	10.8%	-0.39 [-0.68, -0.10]
Ravaud 2004	-12.4	19.2	735	-11.1	20.2	760	89.2%	-0.07 [-0.17, 0.04]
Total (95% CI)			**848**			**838**	**100.0%**	**-0.10 [-0.20, -0.01]**

Heterogeneity: Chi² = 4.28, df = 1 (P = 0.04); I² = 77%
Test for overall effect: Z = 2.07 (P = 0.04)

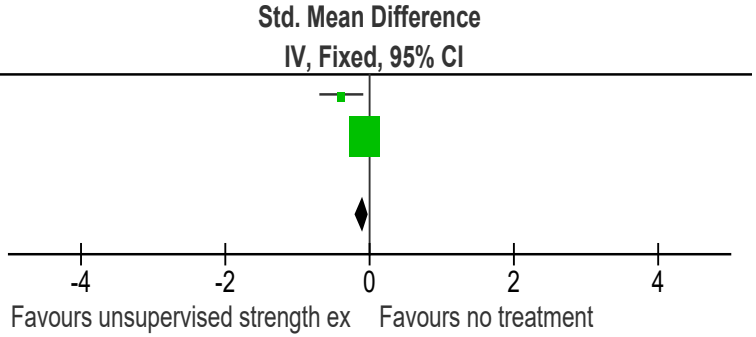


Figure 66: Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Juhakoski 2011	24.4	20.9	60	30	31.3	58	47.5%	-0.21 [-0.57, 0.15]
Osteras 2014	10.9	5.4	65	10.5	4.9	65	52.5%	0.08 [-0.27, 0.42]
Total (95% CI)			**125**			**123**	**100.0%**	**-0.06 [-0.31, 0.19]**

Heterogeneity: Chi² = 1.27, df = 1 (P = 0.26); I² = 21%
Test for overall effect: Z = 0.46 (P = 0.64)

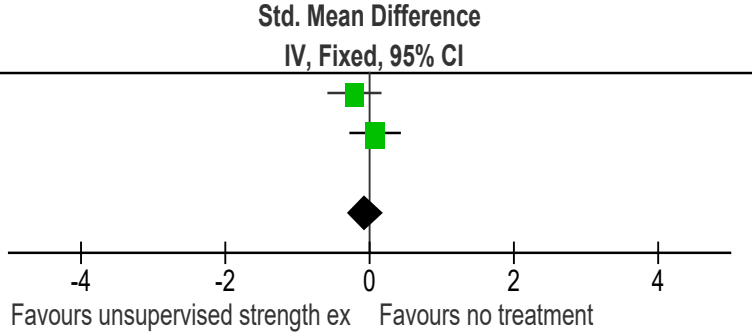


Figure 67: Psychological distress (HADS anxiety, 0-21, high is poor, change score) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
O'reilly 1999	-0.57	3.09	113	0.06	3.22	78	-0.63 [-1.54, 0.28]

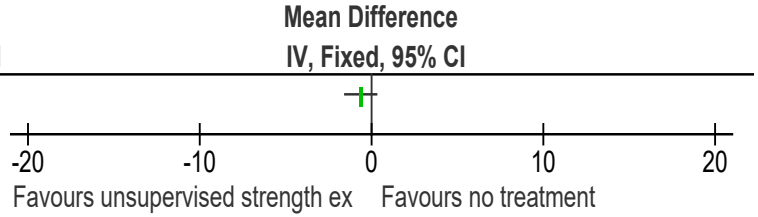


Figure 68: Psychological distress (HADS depression, 0-21, high is poor, change score) at >3 months

Study or Subgroup	Unsupervised strength ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
O'reilly 1999	-0.57	2.09	113	0.11	2.16	78	-0.68 [-1.30, -0.06]

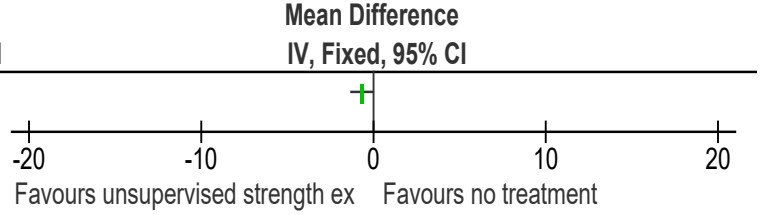


Figure 69: Serious adverse events at ≤3 months

Study or Subgroup	Unsupervised strength ex Events	Total	No treatment Events	Total	Peto Odds Ratio Peto, Fixed, 95% CI
Bennell 2010	5	45	0	44	7.94 [1.32, 47.77]

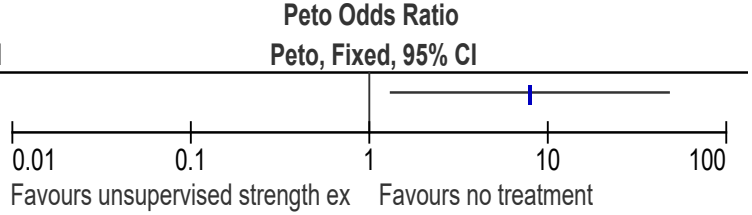
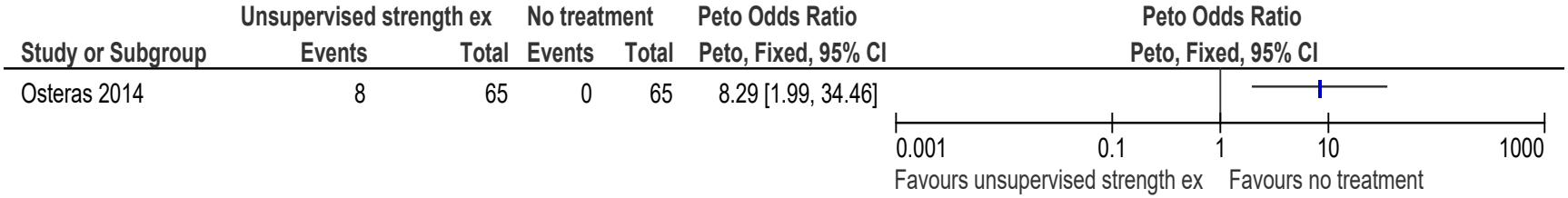


Figure 70: Serious adverse events at >3 months



E.7 Supervised aerobic exercise compared to no treatment

Figure 71: Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months

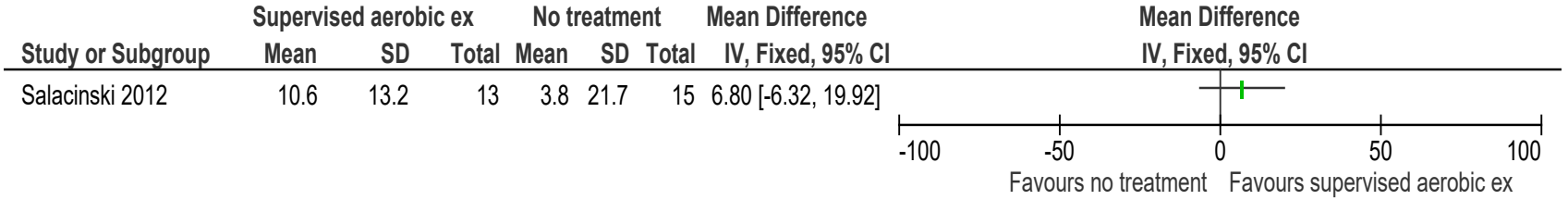


Figure 72: Quality of life (SF-36 physical component, 0-100, high is good, change score and final value) at >3 months

Study or Subgroup	Supervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Brosseau 2012	42.82	9.24	44	45.149	8.93	36	31.9%	-2.33 [-6.32, 1.67]
Christensen 2015	3.8	7.8	64	4.4	8	64	68.1%	-0.60 [-3.34, 2.14]
Total (95% CI)			**108**			**100**	**100.0%**	**-1.15 [-3.41, 1.11]**

Heterogeneity: Chi² = 0.49, df = 1 (P = 0.48); I² = 0%
Test for overall effect: Z = 1.00 (P = 0.32)

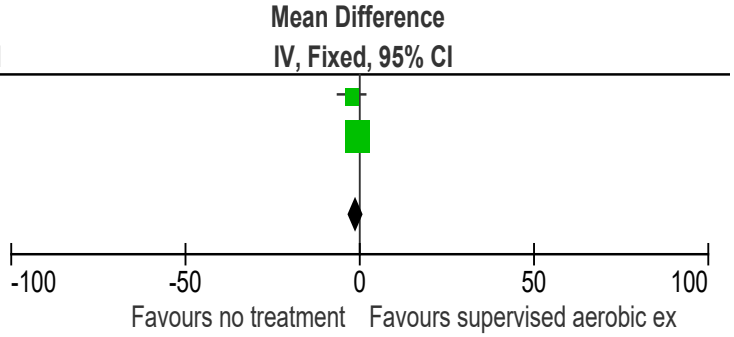


Figure 73: Quality of life (SF-36 mental component, 0-100, high is good, change score) at >3 months

Study or Subgroup	Supervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Brosseau 2012	51.993	11	44	53.101	9.914	36	24.8%	-1.11 [-5.70, 3.48]
Christensen 2015	0.1	7.6	64	1.3	7.6	64	75.2%	-1.20 [-3.83, 1.43]
Total (95% CI)			**108**			**100**	**100.0%**	**-1.18 [-3.46, 1.11]**

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.97); I² = 0%
Test for overall effect: Z = 1.01 (P = 0.31)

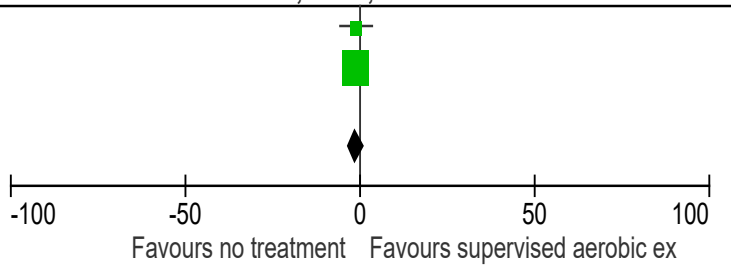


Figure 74: Pain (KOOS, 0-100, high is good, change score) at ≤3 months

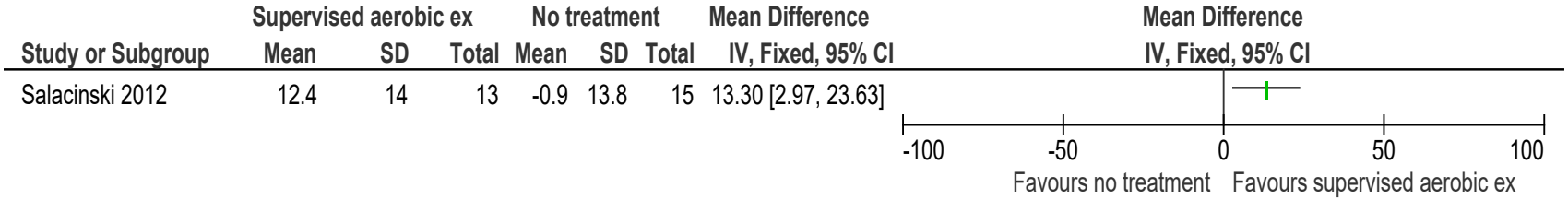


Figure 75: Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months

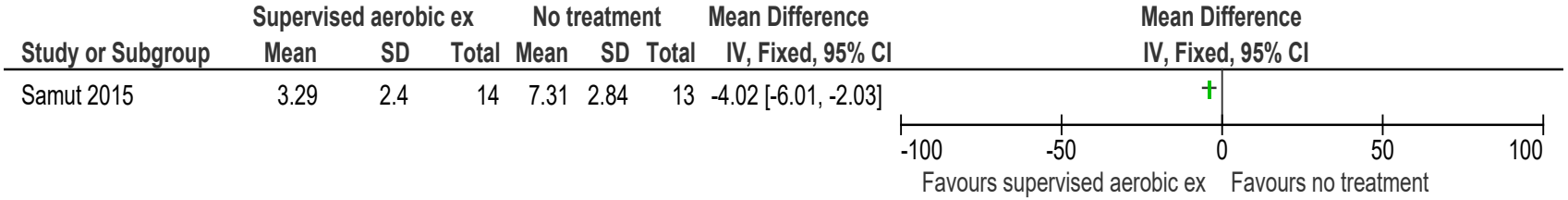


Figure 76: Pain (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months

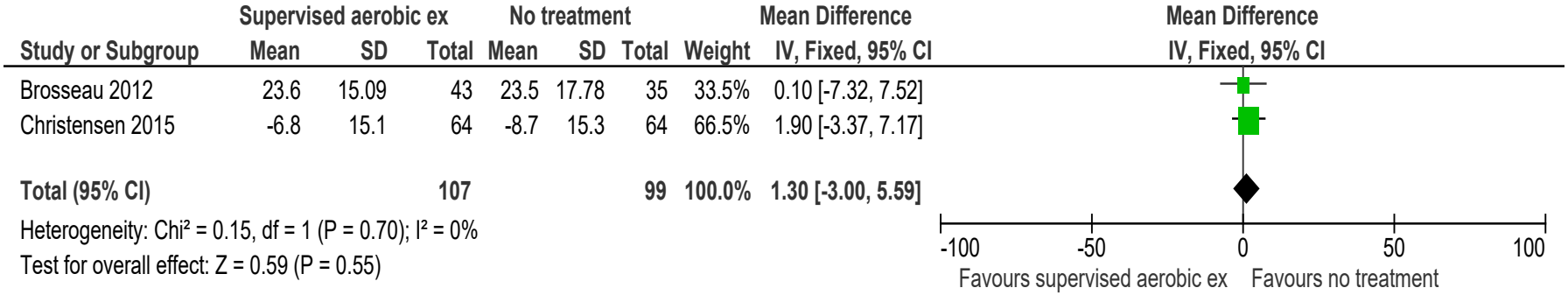


Figure 77: Physical function (KOOS, 0-100, high is good, change score) at ≤3 months

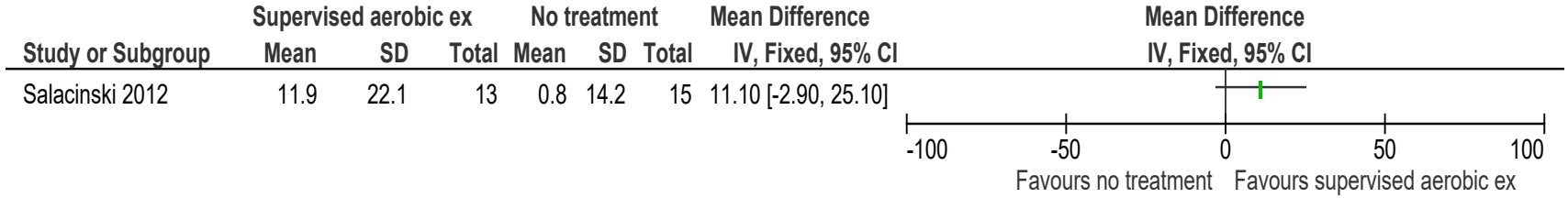


Figure 78: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

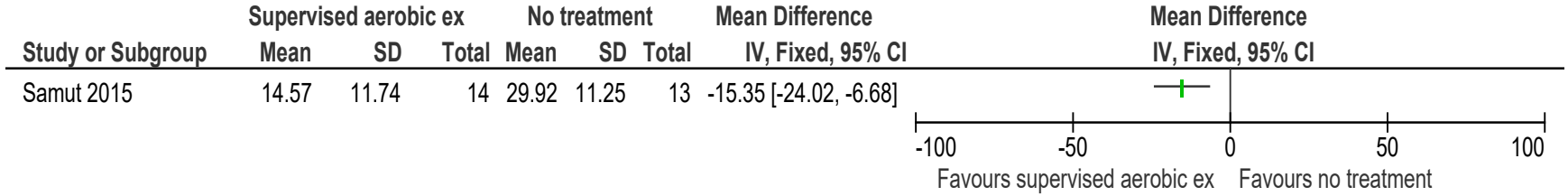


Figure 79: Physical function (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months

Study or Subgroup	Supervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Brosseau 2012	18.2	14.63	43	19.4	17.08	35	33.0%	-1.20 [-8.35, 5.95]
Christensen 2015	-8.4	14.5	64	-6.2	14.5	64	67.0%	-2.20 [-7.22, 2.82]
Total (95% CI)			**107**			**99**	**100.0%**	**-1.87 [-5.98, 2.24]**

Heterogeneity: Chi² = 0.05, df = 1 (P = 0.82); I² = 0%
Test for overall effect: Z = 0.89 (P = 0.37)

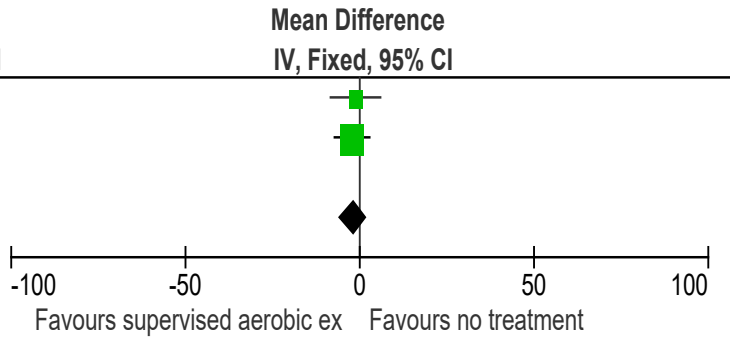
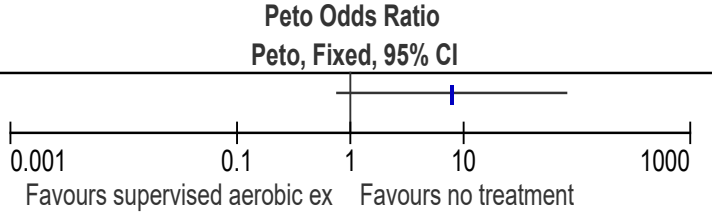


Figure 80: Serious adverse events at ≤3 months

Study or Subgroup	Supervised aerobic ex Events	Total	No treatment Events	Total	Peto Odds Ratio Peto, Fixed, 95% CI
Salacinski 2012	3	19	0	18	7.86 [0.77, 80.77]



E.8 Unsupervised aerobic exercise compared to no treatment

Figure 81: Quality of life (KOOS, 0-100, high is good, final value) at ≤3 months

	Unsupervised aerobic ex			No treatment			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Bossen 2013	49.4	36	85	47.3	35.8	80	2.10 [-8.86, 13.06]	

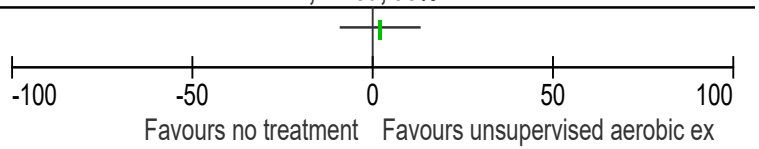


Figure 82: Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months

	Unsupervised aerobic ex			No treatment			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Evcik 2002	8.6	4	28	36.6	6.1	26	-28.00 [-30.77, -25.23]	

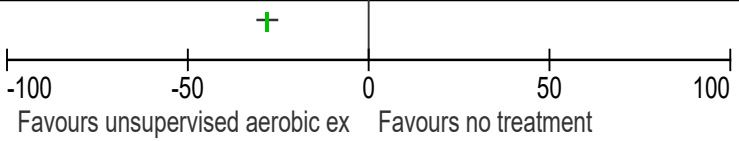


Figure 83: Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months

	Unsupervised aerobic ex			No treatment			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Evcik 2002	9	3.3	28	20.4	3.2	26	-11.40 [-13.13, -9.67]	

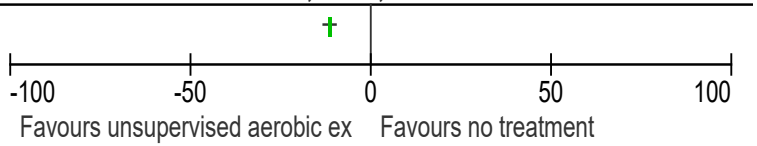


Figure 84: Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	14.6	1.3	28	49.3	1.7	26	-34.70 [-35.51, -33.89]

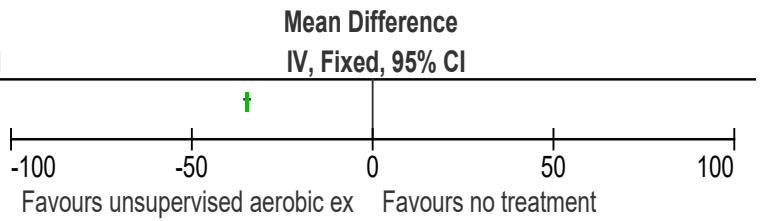


Figure 85: Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	19.6	4	28	35.3	4.4	26	-15.70 [-17.95, -13.45]

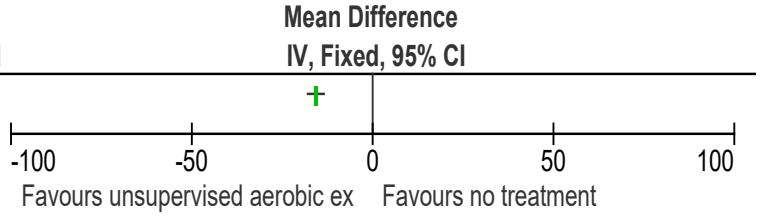


Figure 86: Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Evcik 2002	6.9	3	28	27.9	4.5	26	-21.00 [-23.06, -18.94]

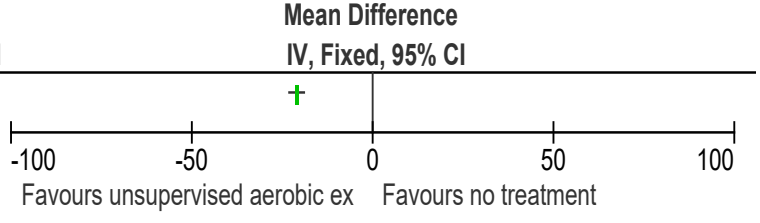


Figure 87: Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months

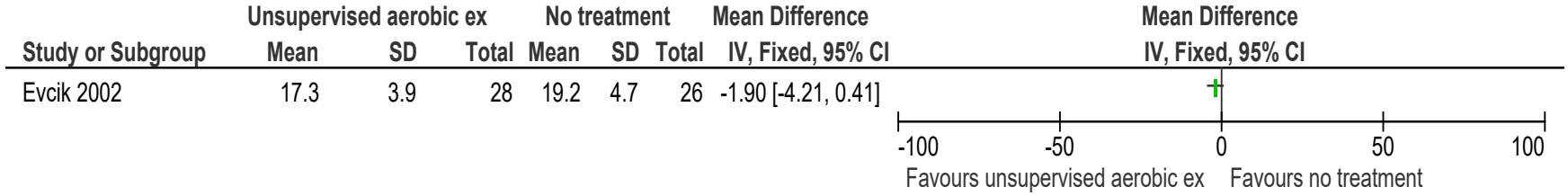


Figure 88: Quality of life (KOOS, 0-100, high is good, final value) at >3 months

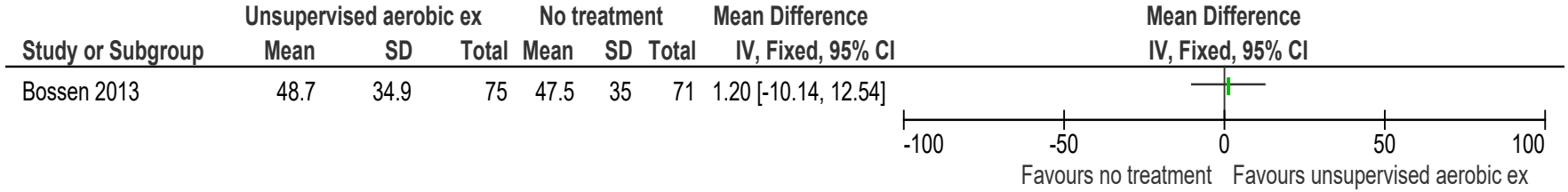


Figure 89: Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Bossen 2013	3.5	4.9	85	4.5	5.3	81	34.3%	-0.20 [-0.50, 0.11]
Evcik 2002	3.4	1.3	28	6	3.3	26	33.3%	-1.04 [-1.61, -0.47]
Shahine 2020	17.5	18.1	33	77	17.3	33	32.3%	-3.32 [-4.08, -2.56]
Total (95% CI)			**146**			**140**	**100.0%**	**-1.49 [-3.11, 0.14]**

Heterogeneity: Tau² = 1.98; Chi² = 57.81, df = 2 (P < 0.00001); I² = 97%

Test for overall effect: Z = 1.79 (P = 0.07)

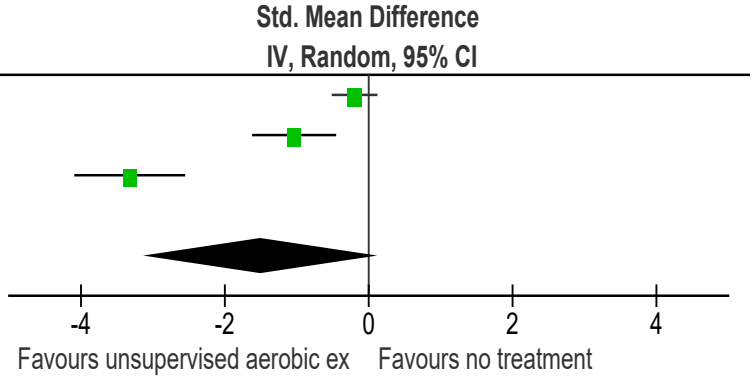


Figure 90: Pain (VAS, 0-10, high is poor, final value) at >3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	3.5	4.7	76	3.8	4.7	71	-0.30 [-1.82, 1.22]

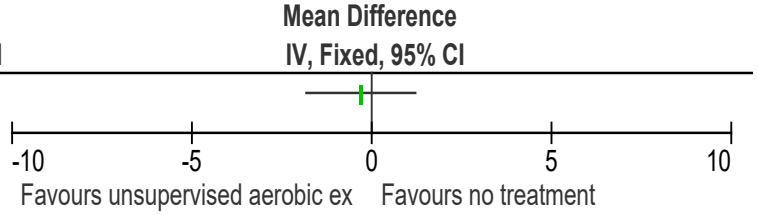


Figure 91: Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Bossen 2013	-67.8	40.2	84	-61.3	39.2	80	34.0%	-0.16 [-0.47, 0.14]
Evcik 2002	10.2	2.4	28	20.7	4.4	26	32.9%	-2.95 [-3.74, -2.16]
Shahine 2020	30	14.8	33	79.9	15.5	33	33.0%	-3.25 [-4.00, -2.51]
Total (95% CI)			**145**			**139**	**100.0%**	**-2.10 [-4.38, 0.18]**

Heterogeneity: Tau² = 3.94; Chi² = 86.23, df = 2 (P < 0.00001); I² = 98%

Test for overall effect: Z = 1.81 (P = 0.07)

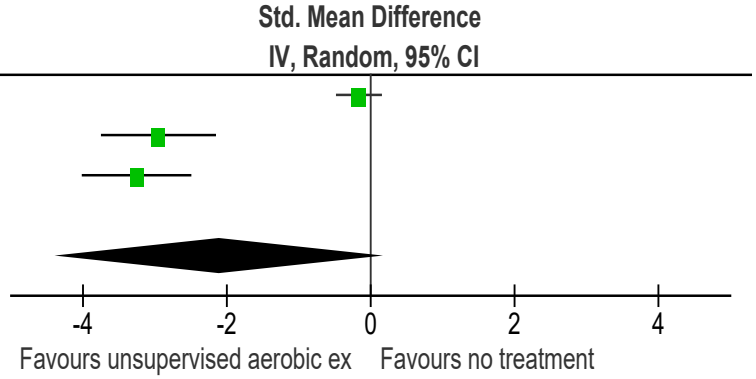


Figure 92: Physical function (WOMAC, 0-100, high is good, final value) at >3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	67.9	38.9	75	62.9	38.1	72	5.00 [-7.45, 17.45]

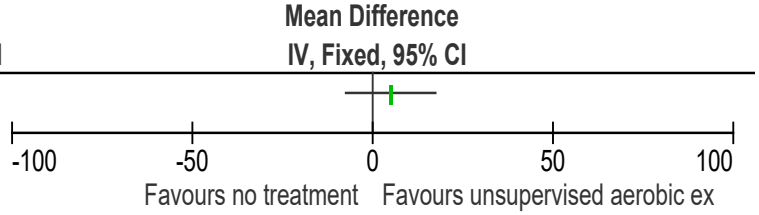


Figure 93: Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	3.5	4.7	85	4.2	4.8	79	-0.70 [-2.16, 0.76]

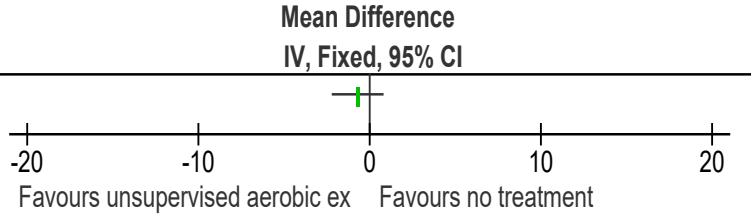


Figure 94: Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	2.6	5.2	85	3.2	5	79	-0.60 [-2.16, 0.96]

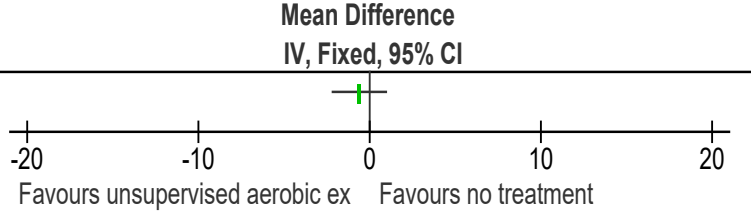


Figure 95: Psychological distress (HADS anxiety, 0-21, high is poor, final value) at >3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	3.1	5.1	75	4.1	5	72	-1.00 [-2.63, 0.63]

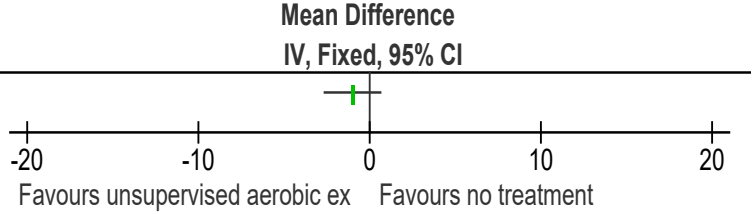
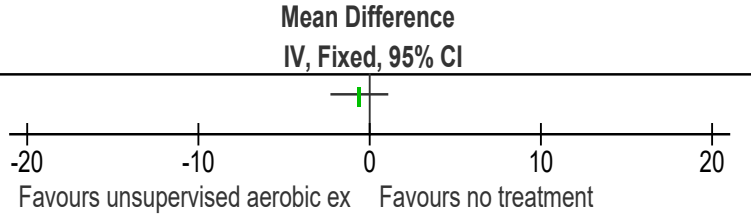


Figure 96: Psychological distress (HADS depression, 0-21, high is poor, final value) at >3 months

Study or Subgroup	Unsupervised aerobic ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bossen 2013	2.4	5.1	75	3	5	72	-0.60 [-2.23, 1.03]



E.9 Other supervised exercise compared to supervised strength exercise

Figure 97: Quality of life (KOOS, Assessment of Quality of Life Instrument version two, WHO Quality of Life total [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Bennell 2014	0.78	0.14	38	0.78	0.16	44	55.3%	0.00 [-0.43, 0.43]
Khurakhorn 2021	93.06	5.8	17	92.88	12.23	17	23.0%	0.02 [-0.65, 0.69]
Nahayatbin 2018	57.13	16.41	16	57.31	19.39	16	21.7%	-0.01 [-0.70, 0.68]
Total (95% CI)			**71**			**77**	**100.0%**	**0.00 [-0.32, 0.32]**

Heterogeneity: Chi² = 0.00, df = 2 (P = 1.00); I² = 0%
Test for overall effect: Z = 0.01 (P = 0.99)

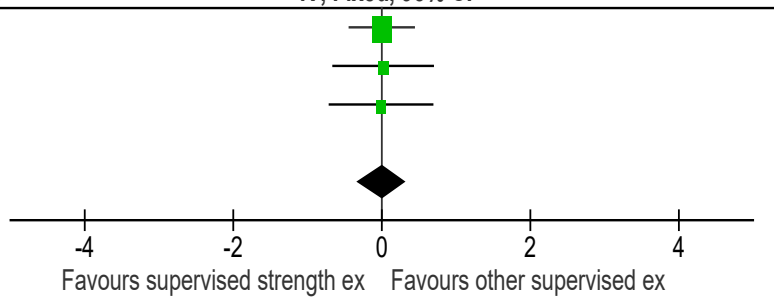


Figure 98: Quality of life (SF-12 physical score, 0-100, high is good, final value) at ≤3 months

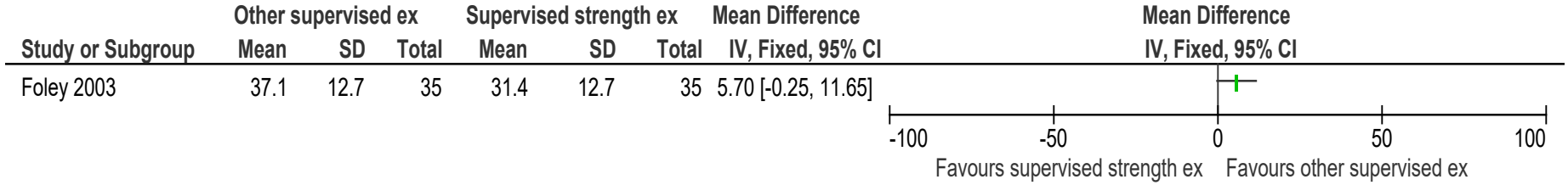


Figure 99: Quality of life (SF-36 mental component, SF-12 mental score, 0-100, high is good, final values) at ≤3 months

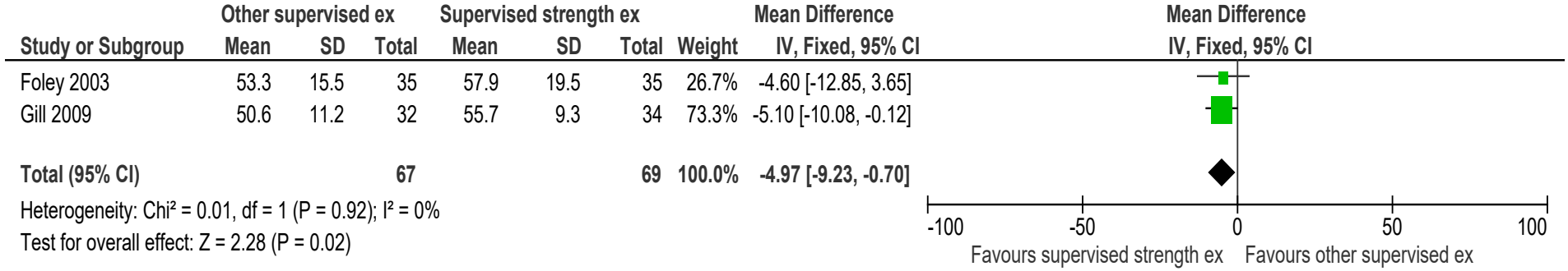


Figure 100: Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at ≤3 months

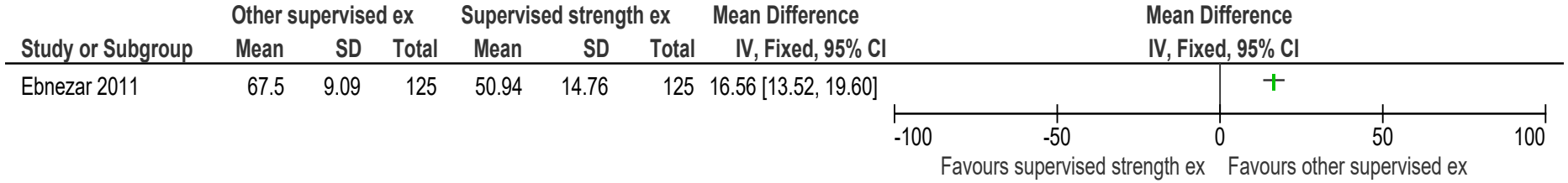


Figure 101: Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	73.77	12.67	125	46.93	11.22	125	26.84 [23.87, 29.81]

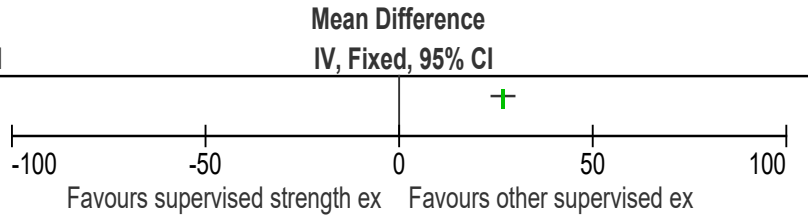


Figure 102: Quality of life (SF-36 role physical, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	86.44	16.55	125	58.33	44.52	125	28.11 [19.78, 36.44]

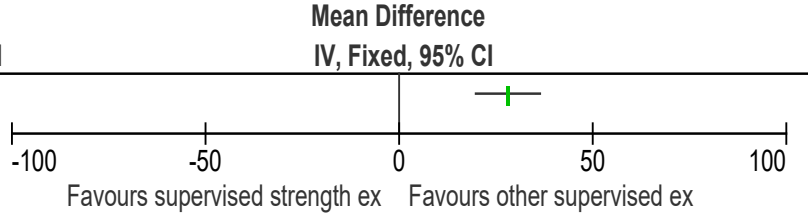


Figure 103: Quality of life (SF-36 vitality, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	36.35	6.08	125	53.2	6.86	125	-16.85 [-18.46, -15.24]

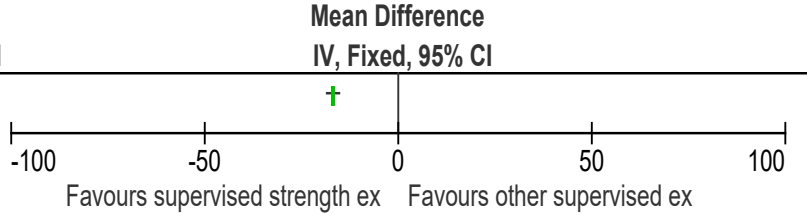


Figure 104: Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	77.47	20.91	125	60.12	12.57	125	17.35 [13.07, 21.63]

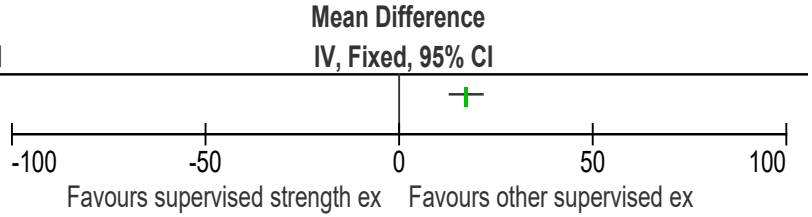


Figure 105: Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	34.33	5.46	125	52.27	5.91	125	-17.94 [-19.35, -16.53]

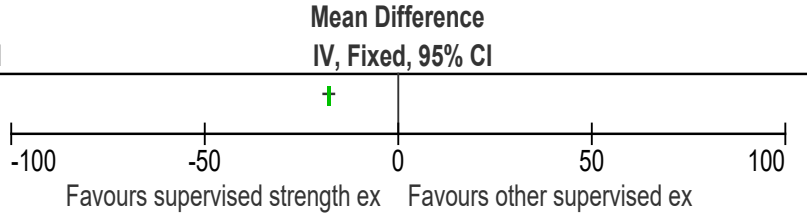


Figure 106: Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	86.41	17.59	125	58.75	38.94	125	27.66 [20.17, 35.15]

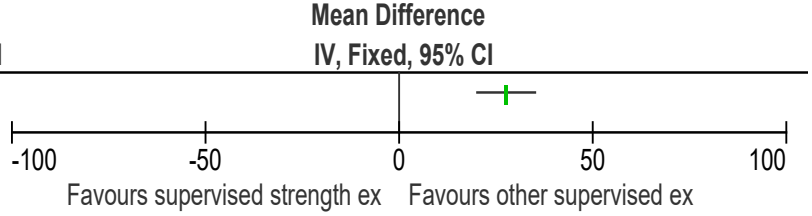


Figure 107: Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Ebnezar 2011	64.04	8.92	125	57.15	10.42	125	6.89 [4.49, 9.29]

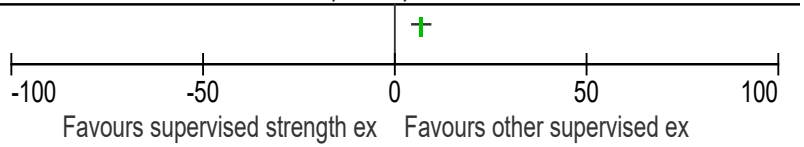


Figure 108: Quality of life (SF-36 mental component, 0-100, high is good, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Gill 2009	51.2	10.5	32	51.9	12.1	34	-0.70 [-6.16, 4.76]

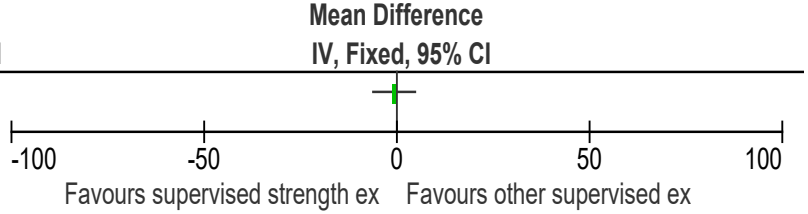


Figure 109: Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Gomiero 2018	57.5	43.3	32	50.8	38.2	32	6.70 [-13.31, 26.71]

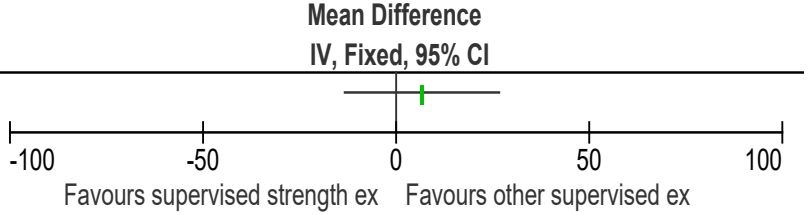


Figure 110: Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at >3 months

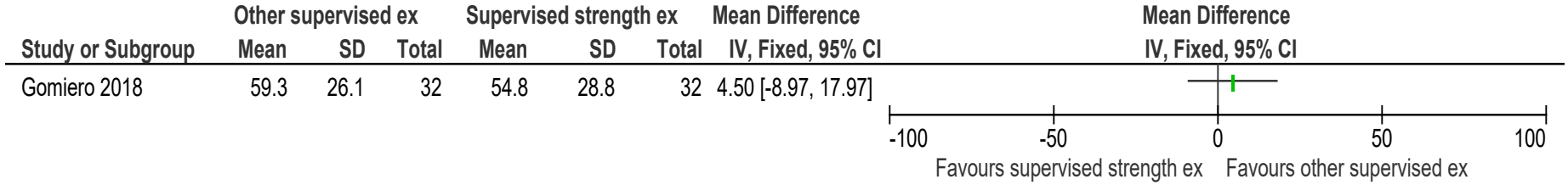


Figure 111: Quality of life (SF-36 role physical, 0-100, high is good, final value) at >3 months

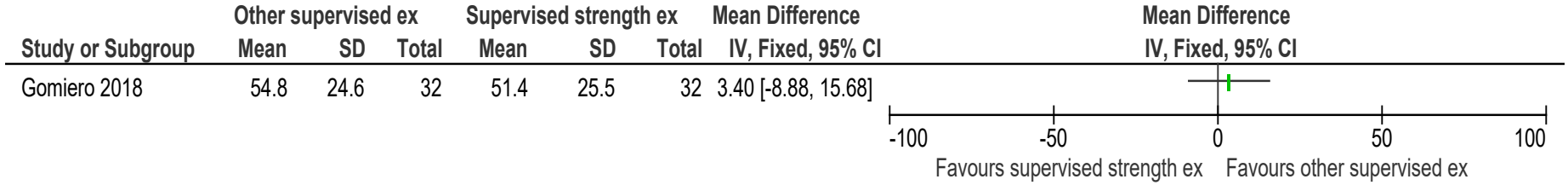


Figure 112: Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months

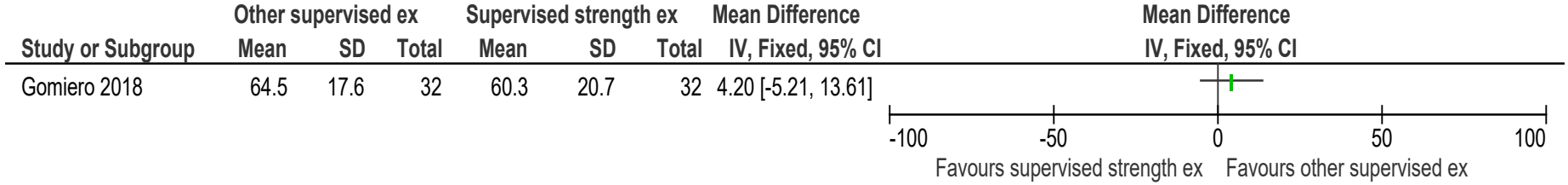


Figure 113: Quality of life (SF-36 general health, 0-100, high is good, final value) at >3 months

	Other supervised ex			Supervised strength ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Gomiero 2018	60.8	20	32	62	21.4	32	-1.20 [-11.35, 8.95]	

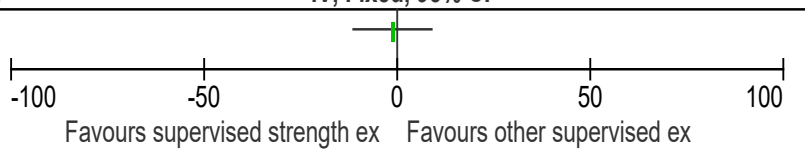


Figure 114: Quality of life (SF-36 mental health, 0-100, high is good, final value) at >3 months

	Other supervised ex			Supervised strength ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Gomiero 2018	74.1	17	32	65.6	19.8	32	8.50 [-0.54, 17.54]	

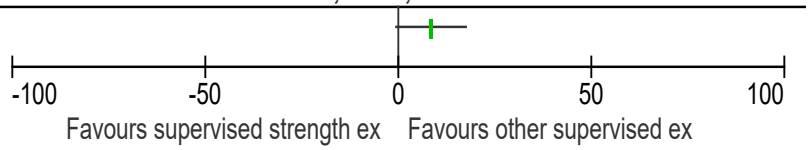


Figure 115: Quality of life (SF-36 role emotional, 0-100, high is good, final value) at >3 months

	Other supervised ex			Supervised strength ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Gomiero 2018	61.1	42.9	32	64.6	42.3	32	-3.50 [-24.37, 17.37]	

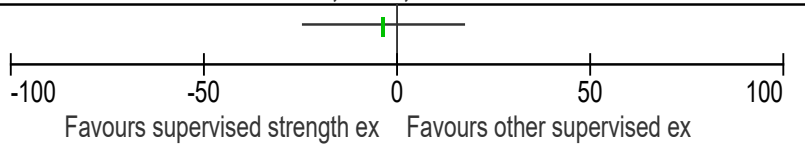


Figure 116: Quality of life (SF-36 social functioning, 0-100, high is good, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Gomiero 2018	74	23.7	32	67.3	27.2	32	6.70 [-5.80, 19.20]

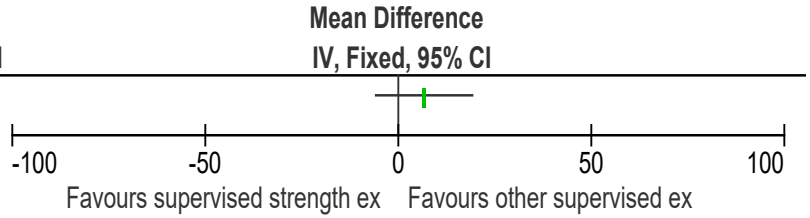


Figure 117: Quality of life (WHO Quality of Life Total, 0-100, high is good, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Khurakhorn 2021	98.18	3.43	17	96.24	8.05	17	1.94 [-2.22, 6.10]

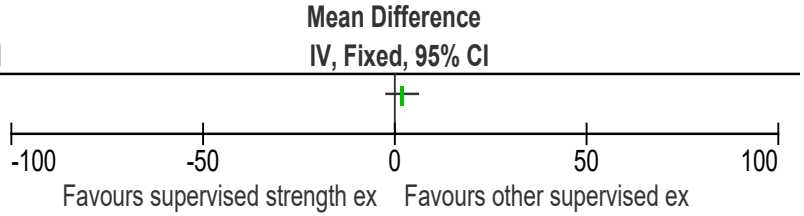


Figure 118: Pain (WOMAC, 0-20, high is poor, change score) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McCaffrey 2019	-4.4	2.1	9	-4.4	2.4	9	83.7%	0.00 [-2.08, 2.08]
Rogers 2011	-2.67	4.41	6	-4	4.79	9	16.3%	1.33 [-3.39, 6.05]
Total (95% CI)			**15**			**18**	**100.0%**	**0.22 [-1.69, 2.12]**

Heterogeneity: Chi² = 0.26, df = 1 (P = 0.61); I² = 0%

Test for overall effect: Z = 0.22 (P = 0.82)

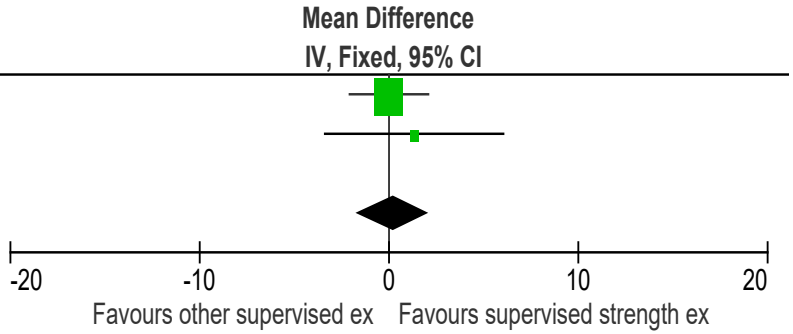


Figure 119: Pain (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Avelar 2011	189	29	11	165	32	10	7.2%	0.76 [-0.14, 1.65]
Bennell 2014	6.4	3.1	38	6.4	2.9	44	8.2%	0.00 [-0.43, 0.43]
Bokaeian 2021	39.8	36	22	44.4	24.6	19	7.9%	-0.14 [-0.76, 0.47]
Ebnezar 2011	1.94	1.11	125	4.17	1.51	125	8.4%	-1.68 [-1.97, -1.39]
Foley 2003	10	4	35	8	5	35	8.1%	0.44 [-0.04, 0.91]
Gill 2009	10.1	2.9	32	9.2	3.7	34	8.1%	0.27 [-0.22, 0.75]
Khurakhorn 2021	7.47	6.85	17	7.94	9.22	17	7.7%	-0.06 [-0.73, 0.62]
Lin 2009	4.3	2.3	36	4.2	3	36	8.1%	0.04 [-0.43, 0.50]
Nahayatbin 2018	-70.13	11.8	16	-58.44	9.51	16	7.6%	-1.06 [-1.81, -0.32]
Nambi 2020	3.6	0.3	18	1.8	0.4	19	6.0%	4.96 [3.60, 6.32]
Ojoawo 2016	3.71	3.4	23	6.5	3.83	22	7.9%	-0.76 [-1.36, -0.15]
Rogers 2012	4.87	3.6	11	4.25	3.45	11	7.4%	0.17 [-0.67, 1.01]
Wortley 2013	141	107	12	71	100	13	7.4%	0.65 [-0.15, 1.46]
Total (95% CI)			**396**			**401**	**100.0%**	**0.18 [-0.43, 0.79]**

Heterogeneity: Tau² = 1.13; Chi² = 181.49, df = 12 (P < 0.00001); I² = 93%

Test for overall effect: Z = 0.57 (P = 0.57)

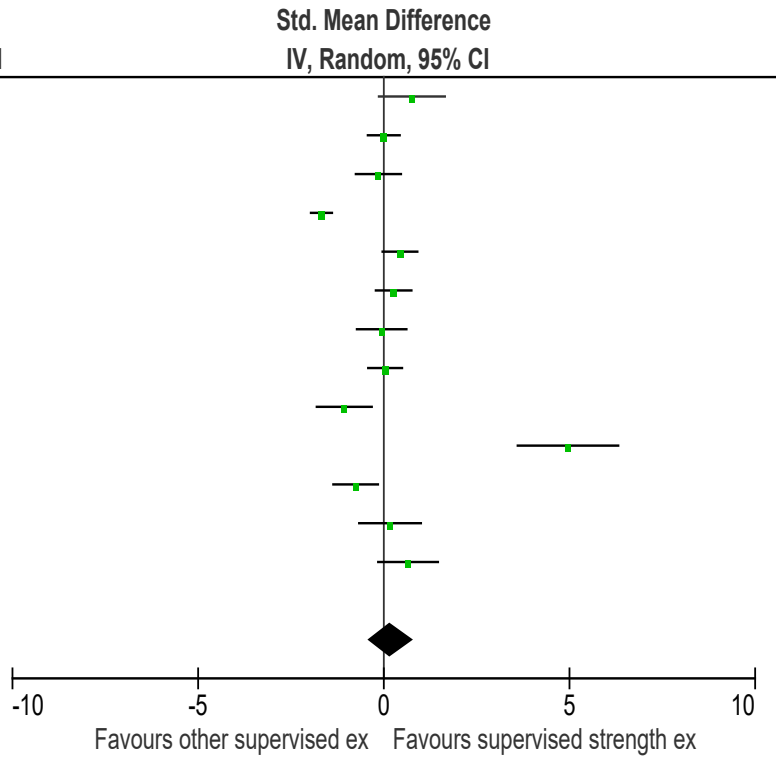


Figure 120: Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Gill 2009	10.3	3.4	32	10	2.3	34	49.4%	0.10 [-0.38, 0.59]
Gomiero 2018	4.6	2.2	32	4.1	2.7	32	47.7%	0.20 [-0.29, 0.69]
Nambi 2020	2.9	0.2	18	0.9	0.3	18	2.9%	7.67 [5.68, 9.66]
Total (95% CI)			**82**			**84**	**100.0%**	**0.37 [0.03, 0.71]**

Heterogeneity: Chi² = 53.44, df = 2 (P < 0.00001); I² = 96%
Test for overall effect: Z = 2.14 (P = 0.03)

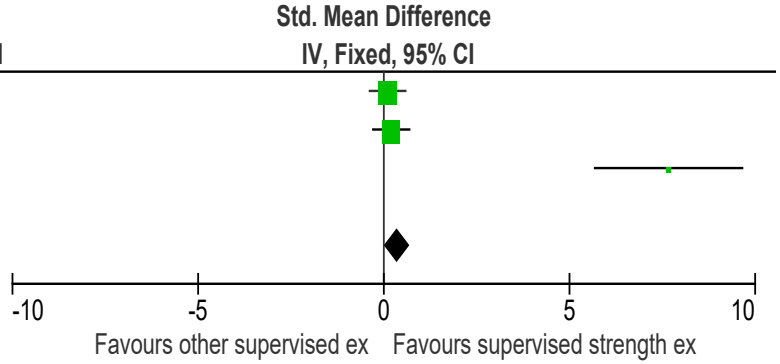


Figure 121: Physical function (WOMAC, 0-68, high is poor, change score) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
McCaffrey 2019	-17.4	14.4	9	-14.9	13.6	9	47.8%	-2.50 [-15.44, 10.44]
Rogers 2011	-16.5	5.69	4	-9.37	15.94	8	52.2%	-7.13 [-19.50, 5.24]
Total (95% CI)			**13**			**17**	**100.0%**	**-4.92 [-13.86, 4.02]**

Heterogeneity: Chi² = 0.26, df = 1 (P = 0.61); I² = 0%
Test for overall effect: Z = 1.08 (P = 0.28)

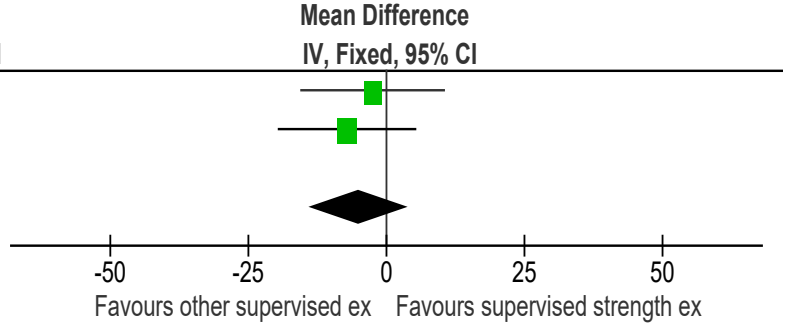


Figure 122: Physical function (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Avelar 2011	718	94	11	777	130	10	7.9%	-0.50 [-1.38, 0.37]
Bennell 2014	18.3	9.6	38	20.1	9.8	44	12.0%	-0.18 [-0.62, 0.25]
Foley 2003	33	17	35	27	12	35	11.6%	0.40 [-0.07, 0.88]
Gill 2009	32.3	10.4	32	29.2	12.7	34	11.5%	0.26 [-0.22, 0.75]
Khurakhorn 2021	20.24	18.81	17	24.35	28.61	17	9.7%	-0.17 [-0.84, 0.51]
Lin 2009	14.6	9.6	36	10.1	8.3	36	11.7%	0.50 [0.03, 0.97]
Nahayatbin 2018	-74.69	12.54	16	-59	10.25	16	8.7%	-1.34 [-2.11, -0.56]
Ojoawo 2016	10.14	11.48	23	17.67	8.66	22	10.4%	-0.73 [-1.33, -0.12]
Rogers 2012	20	9.2	11	16.25	12.53	11	8.2%	0.33 [-0.51, 1.17]
Wortley 2013	552	392	12	240	249	13	8.3%	0.93 [0.09, 1.76]
Total (95% CI)			**231**			**238**	**100.0%**	**-0.03 [-0.40, 0.33]**

Heterogeneity: Tau² = 0.24; Chi² = 32.56, df = 9 (P = 0.0002); I² = 72%
Test for overall effect: Z = 0.17 (P = 0.87)

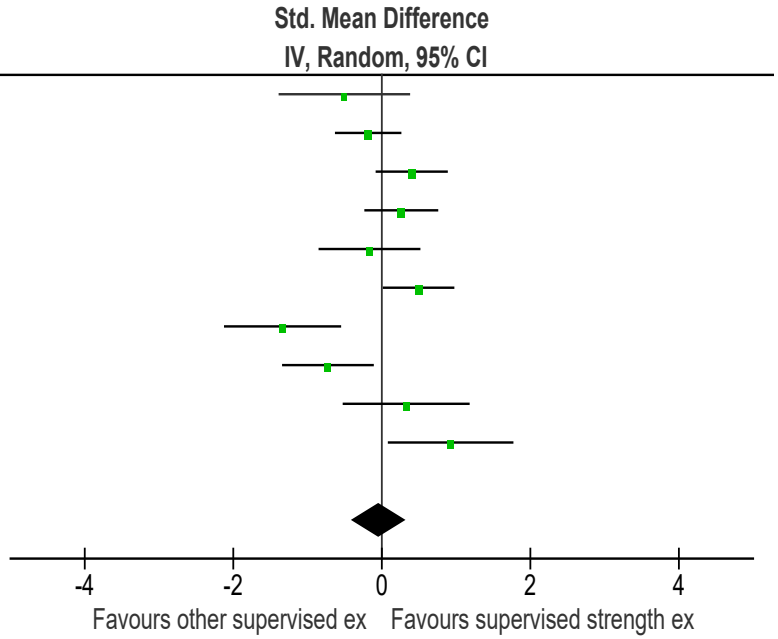


Figure 123: Physical function (WOMAC, 0-68, high is poor, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Gill 2009	32.6	10.7	32	32.2	12.4	34	0.40 [-5.18, 5.98]

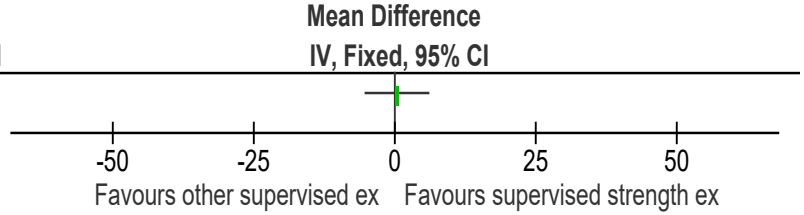


Figure 124: Serious adverse events at ≤3 months

Study or Subgroup	Other supervised ex Events	Total	Supervised strength ex Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Bennell 2014	13	46	10	44	83.3%	0.06 [-0.12, 0.23]
McCaffrey 2019	0	9	0	9	16.7%	0.00 [-0.19, 0.19]
Total (95% CI)		**55**		**53**	**100.0%**	**0.05 [-0.11, 0.20]**
Total events	13		10			

Heterogeneity: Chi² = 0.23, df = 1 (P = 0.63); I² = 0%

Test for overall effect: Z = 0.59 (P = 0.56)

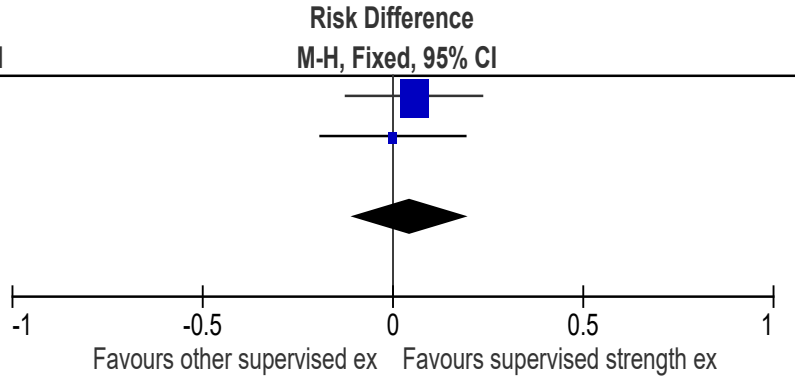
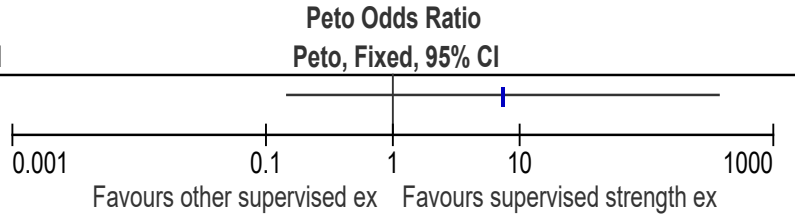


Figure 125: Serious adverse events at >3 months

Study or Subgroup	Other supervised ex Events	Total	Supervised strength ex Events	Total	Peto Odds Ratio Peto, Fixed, 95% CI
Gomiero 2018	1	32	0	32	7.39 [0.15, 372.38]



E.10 Other supervised exercise compared to unsupervised strength exercise

Figure 126: Quality of life (SF-36 physical component, 0-100, high is good, final value) at ≤3 months

	Other supervised ex			Unsupervised strength ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Lim 2010	38.8	7.7	24	36.9	9.6	20	1.90 [-3.31, 7.11]

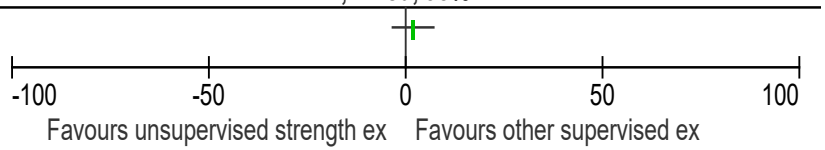


Figure 127: Quality of life (SF-36 mental component, 0-100, high is good, final value) at ≤3 months

	Other supervised ex			Unsupervised strength ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Lim 2010	54.8	8.8	24	48.4	14.3	20	6.40 [-0.79, 13.59]

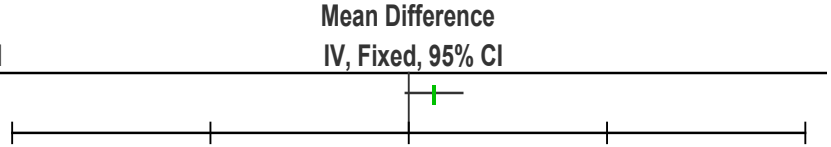


Figure 128: Pain (NRS, 0-10, high is poor, change score) at ≤3 months

	Other supervised ex			Unsupervised strength ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Kuptniratsaikul 2019	-2.3	1.9	40	-1.8	1.7	40	-0.50 [-1.29, 0.29]

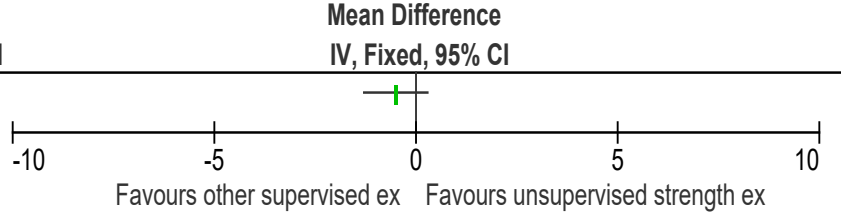


Figure 129: Pain (WOMAC, NRS [different scale ranges], high is poor, final scores) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Kars fertelli 2018	7	4.44	60	14.43	6.4	60	56.4%	-1.34 [-1.74, -0.94]
Lim 2010	3.27	1.67	24	4.55	1.88	20	23.6%	-0.71 [-1.32, -0.10]
Nambi 2020	3.6	0.3	18	3.8	0.4	18	20.0%	-0.55 [-1.22, 0.11]
Total (95% CI)			**102**			**98**	**100.0%**	**-1.03 [-1.33, -0.74]**

Heterogeneity: Chi² = 5.35, df = 2 (P = 0.07); I² = 63%

Test for overall effect: Z = 6.80 (P < 0.00001)

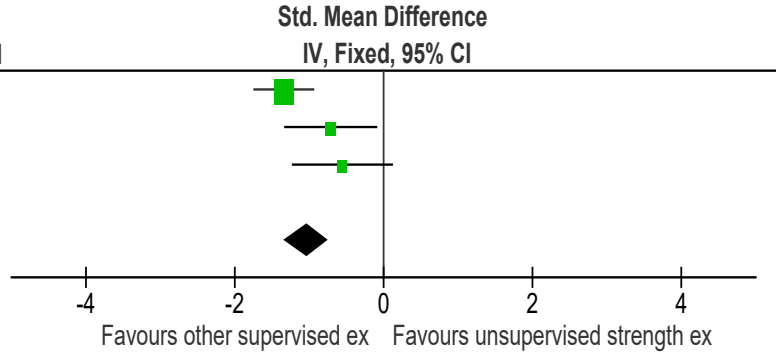


Figure 130: Pain (VAS, 0-10, high is poor, final value) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Nambi 2020	2.9	0.2	18	3.1	0.2	18	-0.20 [-0.33, -0.07]

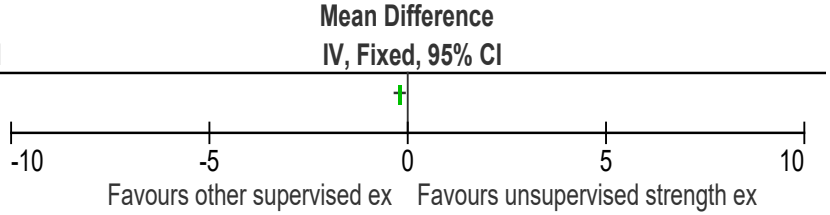


Figure 131: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Kars fertelli 2018	26.1	15.59	60	46.9	17.22	60	-20.80 [-26.68, -14.92]

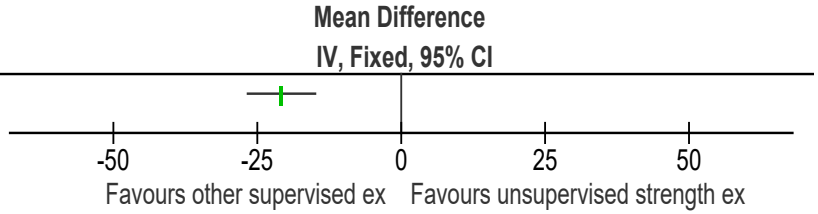
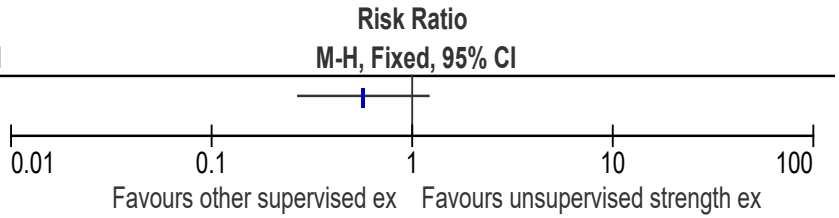


Figure 132: Serious adverse events at ≤3 months

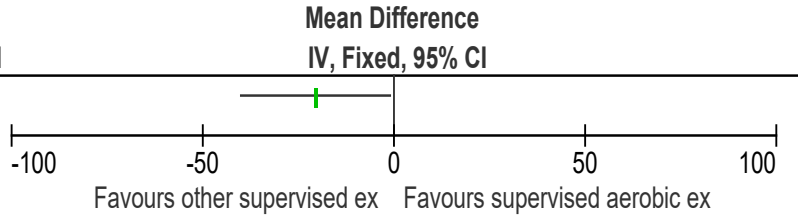
Study or Subgroup	Other supervised ex Events	Total	Unsupervised strength ex Events	Total	Risk Ratio M-H, Fixed, 95% CI
Kuptniratsaikul 2019	8	40	14	40	0.57 [0.27, 1.21]



E.11 Other supervised exercise compared to supervised aerobic exercise

Figure 133: Pain (VAS, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	Supervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bokaeian 2021	39.8	36	22	60.3	26.9	18	-20.50 [-40.01, -0.99]



E.12 Other supervised exercise compared to no treatment

Figure 134: Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Other supervised ex Total	No treatment Total	Weight	Std. Mean Difference IV, Random, 95% CI
Hinman 2007	-0.3462	0.2393	36	35	26.5%	-0.35 [-0.82, 0.12]
Nahayatbin 2018	1.0545	0.3805	16	16	21.0%	1.05 [0.31, 1.80]
Rewald 2020	0.7027	0.2049	55	47	27.7%	0.70 [0.30, 1.10]
Wang 2011	0.4724	0.2815	26	26	24.8%	0.47 [-0.08, 1.02]
Total (95% CI)			**133**	**124**	**100.0%**	**0.44 [-0.14, 1.02]**

Heterogeneity: Tau² = 0.28; Chi² = 14.93, df = 3 (P = 0.002); I² = 80%

Test for overall effect: Z = 1.49 (P = 0.14)

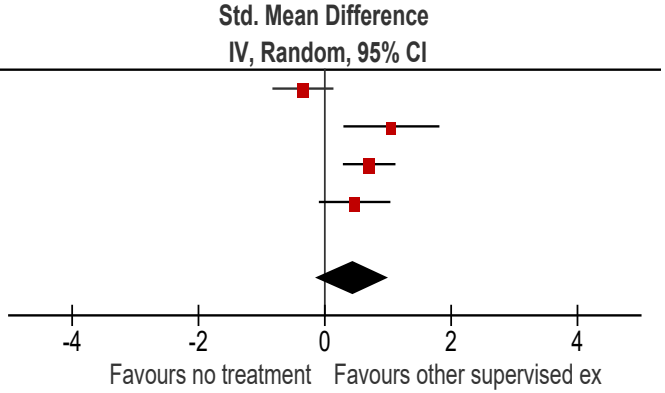


Figure 135: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change scores and final values) at ≤3 months

	Other supervised ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Cheung 2014	38	4.2	18	38.7	4.2	18	24.3%	-0.70 [-3.44, 2.04]
Cheung 2017	41.5	8.5	32	39	8.4	23	19.2%	2.50 [-2.02, 7.02]
Foley 2003	37.1	12.7	35	28.8	11	35	16.4%	8.30 [2.73, 13.87]
Fransen 2007	36.7	10.6	111	33.1	10.6	41	21.3%	3.60 [-0.20, 7.40]
Lee 2009	17.1	14.9	29	5.6	12.9	15	10.5%	11.50 [3.01, 19.99]
Mcilroy 2017	4.3	12.8	7	0.01	4.1	6	8.3%	4.29 [-5.74, 14.32]
Total (95% CI)			**232**			**138**	**100.0%**	**4.00 [0.56, 7.44]**

Heterogeneity: Tau² = 10.74; Chi² = 14.16, df = 5 (P = 0.01); I² = 65%

Test for overall effect: Z = 2.28 (P = 0.02)

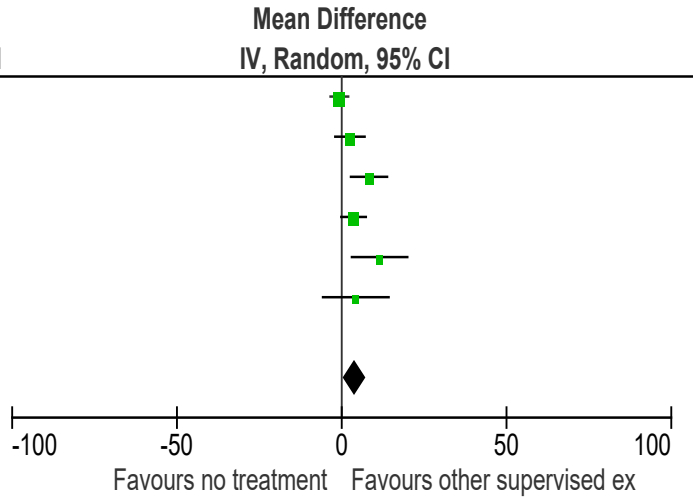


Figure 136: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change scores and final values) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Cheung 2014	49.7	5.1	18	51.7	5.1	18	23.0%	-2.00 [-5.33, 1.33]
Cheung 2017	55.2	8.7	32	52.8	8.8	23	19.0%	2.40 [-2.29, 7.09]
Foley 2003	53.3	15.5	35	50.5	14	35	13.5%	2.80 [-4.12, 9.72]
Fransen 2007	52.7	10	111	48	11.4	41	21.1%	4.70 [0.75, 8.65]
Lee 2009	19.2	15.9	29	9.1	10.3	15	11.8%	10.10 [2.31, 17.89]
Mcilroy 2017	8.2	9.2	7	1.2	5.2	6	11.5%	7.00 [-0.99, 14.99]
Total (95% CI)			**232**			**138**	**100.0%**	**3.37 [-0.11, 6.85]**

Heterogeneity: Tau² = 10.85; Chi² = 13.01, df = 5 (P = 0.02); I² = 62%

Test for overall effect: Z = 1.90 (P = 0.06)

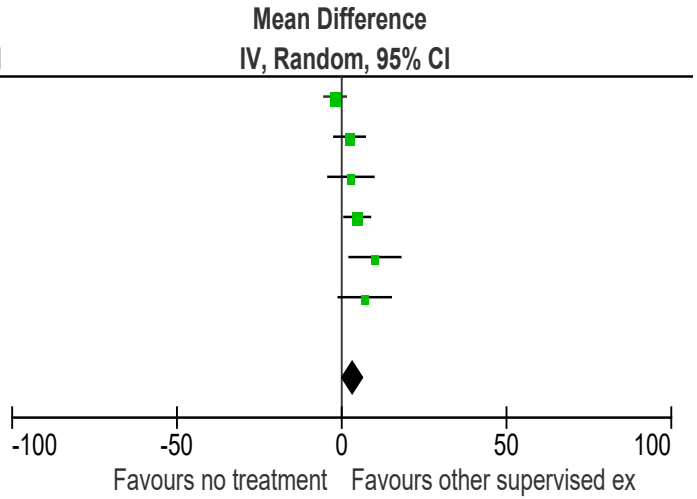


Figure 137: Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
An 2008	61.2	17.9	11	49.1	25.9	10	12.10 [-7.12, 31.32]

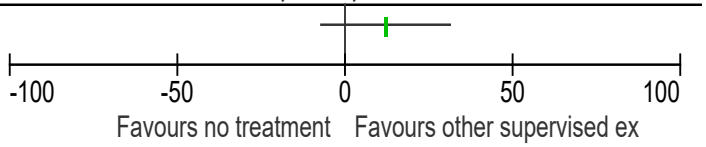


Figure 138: Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months

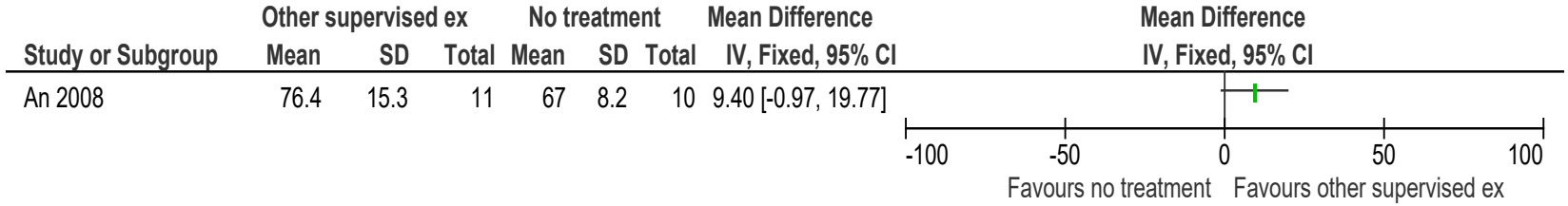


Figure 139: Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months

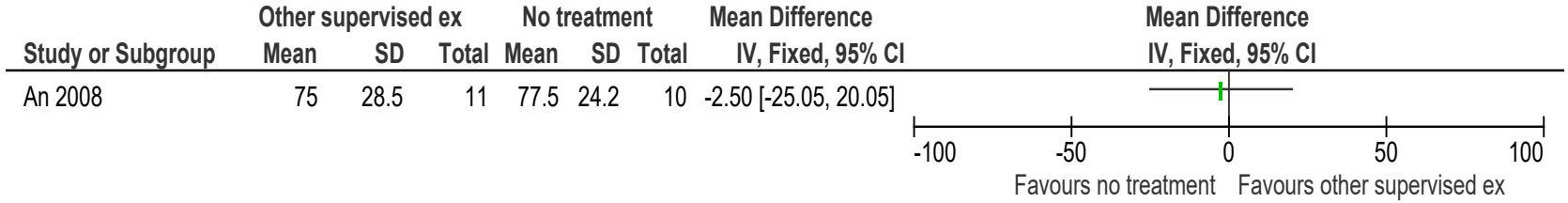


Figure 140: Quality of life (KOOS, 0-100, high is good, change score) at >3 months

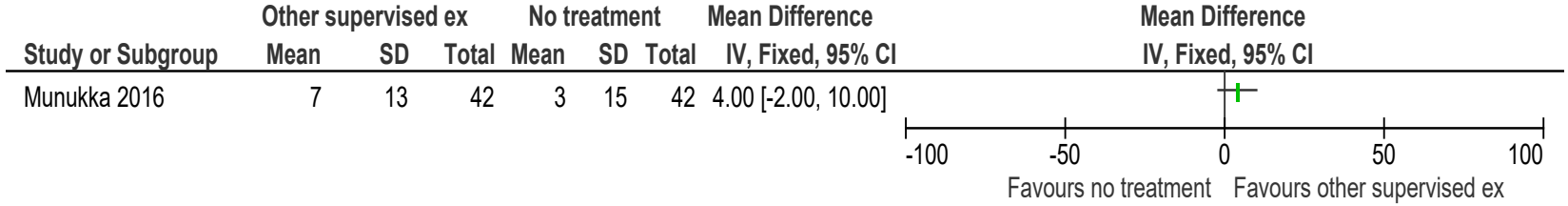


Figure 141: Quality of life (EQ-5D VAS, Quality of Well-being scale [different scale ranges], high is good, final values) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Cochrane 2005	62	19	150	60	19	157	58.2%	0.11 [-0.12, 0.33]
Patrick 2001	0.606	0.069	101	0.599	0.079	121	41.8%	0.09 [-0.17, 0.36]
Total (95% CI)			**251**			**278**	**100.0%**	**0.10 [-0.07, 0.27]**

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.95); I² = 0%
Test for overall effect: Z = 1.15 (P = 0.25)

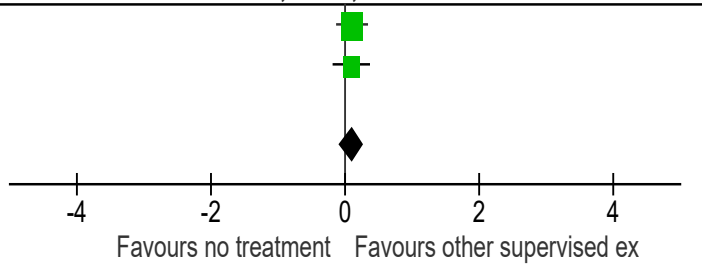


Figure 142: Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Lee 2009	-2.2	4.1	29	-0.2	1.8	15	15.5%	-0.56 [-1.20, 0.07]
Mcilroy 2017	-5.1	4.2	7	0.3	1.9	6	3.8%	-1.50 [-2.79, -0.21]
Robbins 2021	-3.62	2.73	86	-1.43	2.51	86	64.3%	-0.83 [-1.14, -0.52]
Song 2003	-2.45	3.9	22	0.61	5.1	21	16.5%	-0.66 [-1.28, -0.05]
Total (95% CI)			**144**			**128**	**100.0%**	**-0.79 [-1.04, -0.54]**

Heterogeneity: Chi² = 1.88, df = 3 (P = 0.60); I² = 0%
Test for overall effect: Z = 6.17 (P < 0.00001)

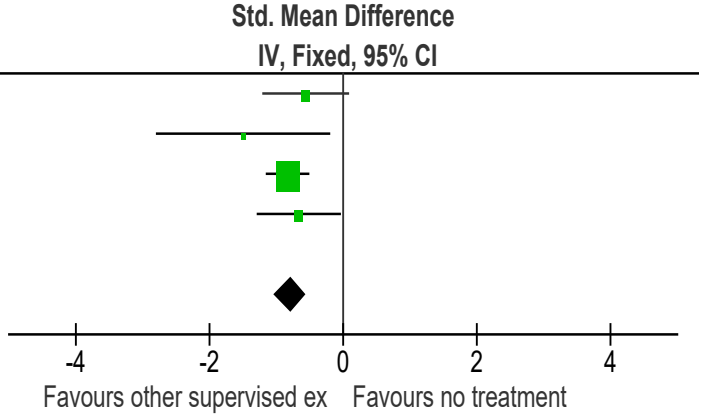


Figure 143: Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Other supervised ex Total	No treatment Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
An 2008	-0.5788	0.448	11	10	2.3%	-0.58 [-1.46, 0.30]
Cheung 2014	-0.873	0.3507	18	18	3.8%	-0.87 [-1.56, -0.19]
Cheung 2017	-0.5111	0.278	32	23	6.0%	-0.51 [-1.06, 0.03]
Duman 2012	-0.09	0.274	30	24	6.1%	-0.09 [-0.63, 0.45]
Foley 2003	0	0.239	35	35	8.1%	0.00 [-0.47, 0.47]
Fransen 2007	-0.6007	0.1861	111	41	13.3%	-0.60 [-0.97, -0.24]
Hinman 2007	-0.5762	0.2425	36	35	7.8%	-0.58 [-1.05, -0.10]
Lin 2009	-1.0224	0.2515	36	36	7.3%	-1.02 [-1.52, -0.53]
Nahayatbin 2018	-1.7103	0.4209	16	16	2.6%	-1.71 [-2.54, -0.89]
Rewald 2020	-0.4866	0.2017	55	47	11.3%	-0.49 [-0.88, -0.09]
Segal 2015	-0.4751	0.2995	29	19	5.1%	-0.48 [-1.06, 0.11]
Sekir 2005	-1.1758	0.4708	12	10	2.1%	-1.18 [-2.10, -0.25]
Wang 2007	-0.5081	0.3307	20	18	4.2%	-0.51 [-1.16, 0.14]
Wang 2011	-0.2189	0.2782	26	26	6.0%	-0.22 [-0.76, 0.33]
Wortley 2013	-0.147	0.5008	12	6	1.8%	-0.15 [-1.13, 0.83]
Ye 2019	-0.277	0.2843	25	25	5.7%	-0.28 [-0.83, 0.28]
Ye 2020	-0.1932	0.2679	28	28	6.4%	-0.19 [-0.72, 0.33]
Total (95% CI)			**532**	**417**	**100.0%**	**-0.50 [-0.63, -0.36]**

Heterogeneity: Chi² = 26.27, df = 16 (P = 0.05); I² = 39%

Test for overall effect: Z = 7.30 (P < 0.00001)

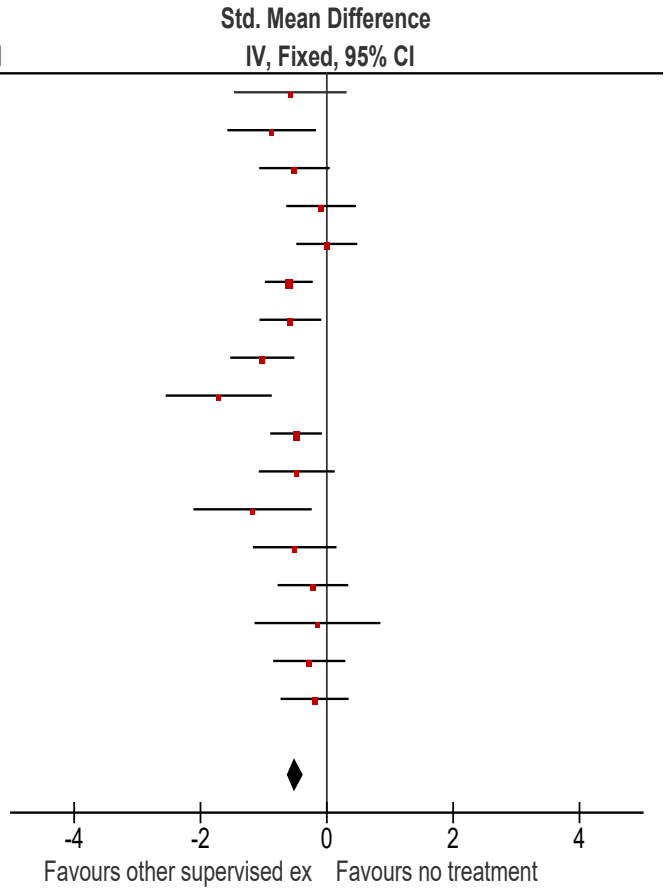


Figure 144: Pain (KOOS, 0-100, high is good, change scores) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Munukka 2016	4	10	42	1	10	42	82.3%	3.00 [-1.28, 7.28]
Segal 2015	10.1	14.2	24	2.8	15.7	18	17.7%	7.30 [-1.91, 16.51]
Total (95% CI)			**66**			**60**	**100.0%**	**3.76 [-0.12, 7.64]**

Heterogeneity: Chi² = 0.69, df = 1 (P = 0.41); I² = 0%
Test for overall effect: Z = 1.90 (P = 0.06)

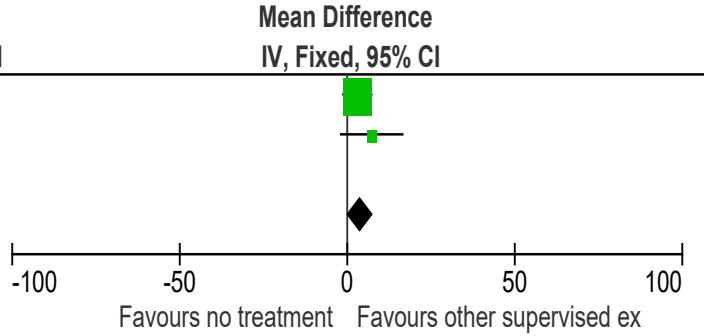


Figure 145: Pain (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Cochrane 2005	8.49	3.94	152	8.88	3.45	158	50.5%	-0.11 [-0.33, 0.12]
Lin 2004	8.62	4.34	56	9.32	2.84	38	14.7%	-0.18 [-0.60, 0.23]
Patrick 2001	1.382	0.737	98	1.462	0.619	117	34.8%	-0.12 [-0.39, 0.15]
Total (95% CI)			**306**			**313**	**100.0%**	**-0.12 [-0.28, 0.04]**

Heterogeneity: Chi² = 0.10, df = 2 (P = 0.95); I² = 0%
Test for overall effect: Z = 1.50 (P = 0.13)

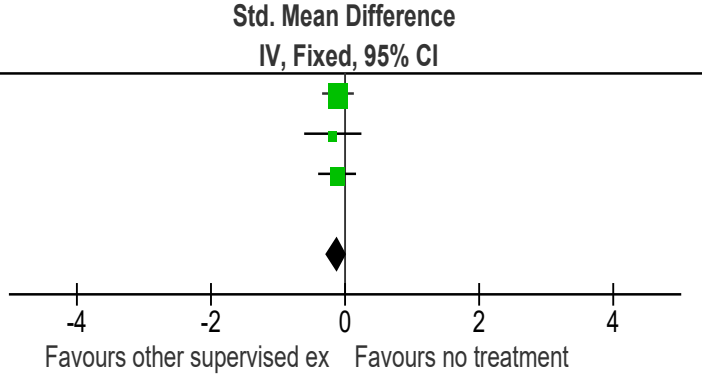


Figure 146: Physical function (WOMAC, 0-68, high is poor, change scores) at ≤3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lee 2009	-9.4	14.4	29	-2.7	10.8	15	35.5%	-6.70 [-14.27, 0.87]
Mcilroy 2017	-10.7	8.9	7	2	9.6	6	19.9%	-12.70 [-22.82, -2.58]
Song 2003	-11.09	12	22	-1.33	10.6	21	44.6%	-9.76 [-16.52, -3.00]
Total (95% CI)			**58**			**42**	**100.0%**	**-9.26 [-13.77, -4.74]**

Heterogeneity: Chi² = 0.90, df = 2 (P = 0.64); I² = 0%

Test for overall effect: Z = 4.02 (P < 0.0001)

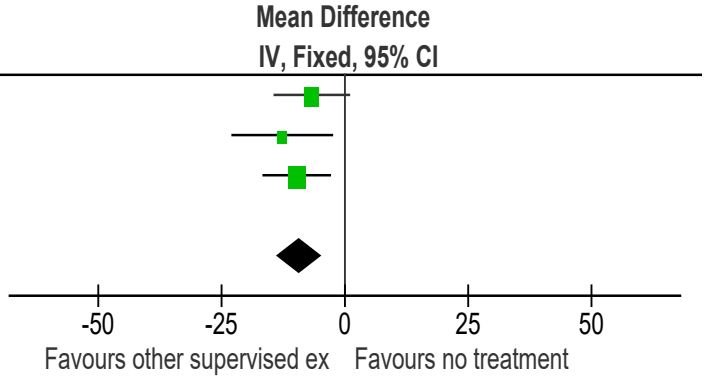


Figure 147: Physical function (KOOS, WOMAC, Multidimensional Health Assessment Questionnaire [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Other supervised ex Total	No treatment Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
An 2008	-0.4126	0.4426	11	10	2.5%	-0.41 [-1.28, 0.45]
Cheung 2014	-0.419	0.3374	18	18	4.4%	-0.42 [-1.08, 0.24]
Cheung 2017	-0.8215	0.2852	32	23	6.1%	-0.82 [-1.38, -0.26]
Duman 2012	0.0977	0.274	30	24	6.6%	0.10 [-0.44, 0.63]
Foley 2003	-0.2614	0.2401	35	35	8.6%	-0.26 [-0.73, 0.21]
Fransen 2007	-0.6558	0.1867	111	41	14.2%	-0.66 [-1.02, -0.29]
Hinman 2007	-0.1889	0.2379	36	35	8.8%	-0.19 [-0.66, 0.28]
Lin 2009	-0.9473	0.2493	36	36	8.0%	-0.95 [-1.44, -0.46]
Nahayatbin 2018	-1.1326	0.3845	16	16	3.4%	-1.13 [-1.89, -0.38]
Rewald 2020	-0.4371	0.2011	55	47	12.3%	-0.44 [-0.83, -0.04]
Wang 2007	-0.2176	0.326	20	18	4.7%	-0.22 [-0.86, 0.42]
Wang 2011	-0.4049	0.2804	26	26	6.3%	-0.40 [-0.95, 0.14]
Wortley 2013	0.203	0.5015	12	6	2.0%	0.20 [-0.78, 1.19]
Ye 2019	-0.1523	0.2833	25	25	6.2%	-0.15 [-0.71, 0.40]
Ye 2020	-0.9526	0.2831	28	28	6.2%	-0.95 [-1.51, -0.40]
Total (95% CI)			**491**	**388**	**100.0%**	**-0.47 [-0.61, -0.33]**

Heterogeneity: Chi² = 22.27, df = 14 (P = 0.07); I² = 37%

Test for overall effect: Z = 6.70 (P < 0.00001)

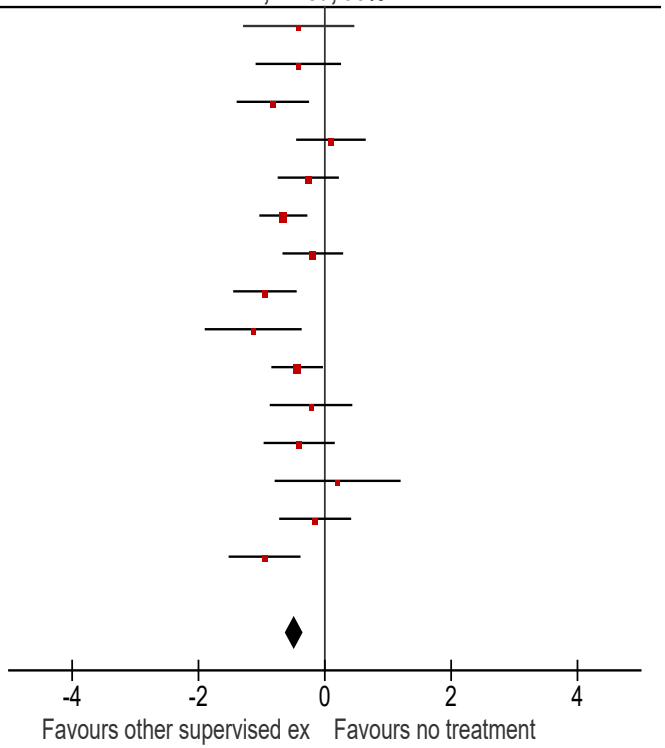


Figure 148: Physical function (KOOS, 0-100, high is good, change scores) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Munukka 2016	4	10	42	0	8	42	4.00 [0.13, 7.87]

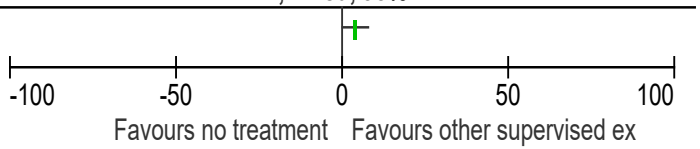


Figure 149: Physical function (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Other supervised ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Cochrane 2005	29.73	14.62	150	31.15	12.73	156	49.9%	-0.10 [-0.33, 0.12]
Lin 2004	30.16	14.03	56	34.96	9.87	38	14.5%	-0.38 [-0.80, 0.04]
Patrick 2001	0.933	0.55	101	1.127	0.671	121	35.5%	-0.31 [-0.58, -0.05]
Total (95% CI)			**307**			**315**	**100.0%**	**-0.22 [-0.38, -0.06]**

Heterogeneity: Chi² = 2.07, df = 2 (P = 0.36); I² = 3%
Test for overall effect: Z = 2.70 (P = 0.007)

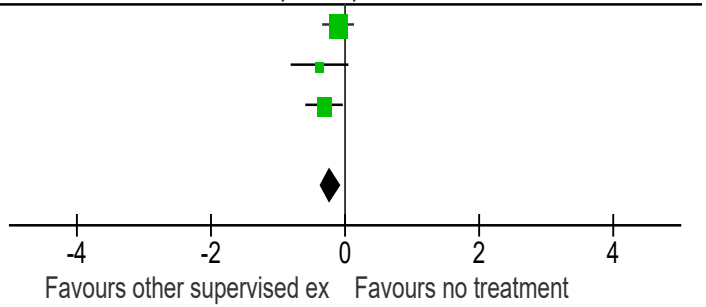


Figure 150: Psychological distress (HADS anxiety subscale, DAS scale anxiety subscale [different scale ranges], high is poor, final values) at ≤3 months

	Other supervised ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Cheung 2017	3.8	2.6	32	4.4	2.8	23	31.1%	-0.22 [-0.76, 0.32]
Fransen 2007	4.9	5.6	111	7.3	7.8	41	68.9%	-0.38 [-0.74, -0.02]
Total (95% CI)			**143**			**64**	**100.0%**	**-0.33 [-0.63, -0.03]**

Heterogeneity: Chi² = 0.24, df = 1 (P = 0.63); I² = 0%

Test for overall effect: Z = 2.17 (P = 0.03)

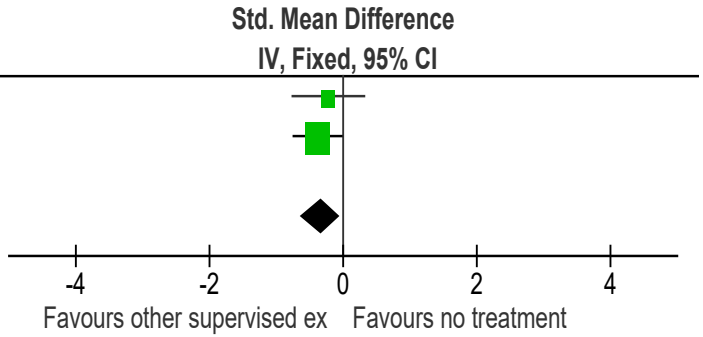


Figure 151: Psychological distress (HADS depression subscale, DAS scale depression subscale [different scale ranges], high is poor, final values) at ≤3 months

	Other supervised ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Cheung 2017	3.8	2	32	3.7	2.1	23	31.2%	0.05 [-0.49, 0.58]
Fransen 2007	5.9	7.4	111	9	11	41	68.8%	-0.36 [-0.72, -0.00]
Total (95% CI)			**143**			**64**	**100.0%**	**-0.23 [-0.53, 0.06]**

Heterogeneity: Chi² = 1.55, df = 1 (P = 0.21); I² = 36%

Test for overall effect: Z = 1.54 (P = 0.12)

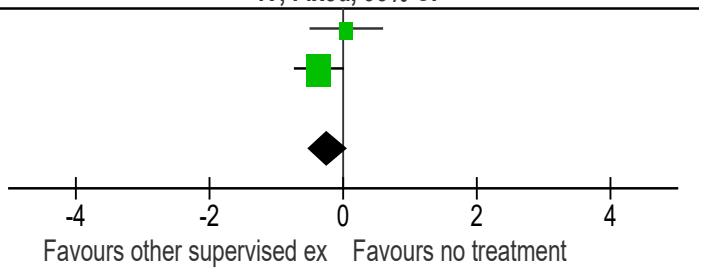


Figure 152: Psychological distress (DAS scale stress subscale, 0-48, high is poor, final value) at ≤3 months

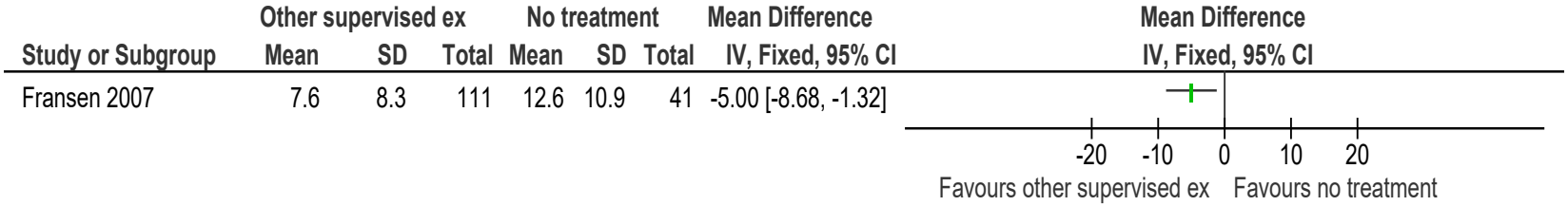


Figure 153: Psychological distress (Centre for Epidemiological Studies Depression Scale, 0-60, high is poor, final value) at >3 months

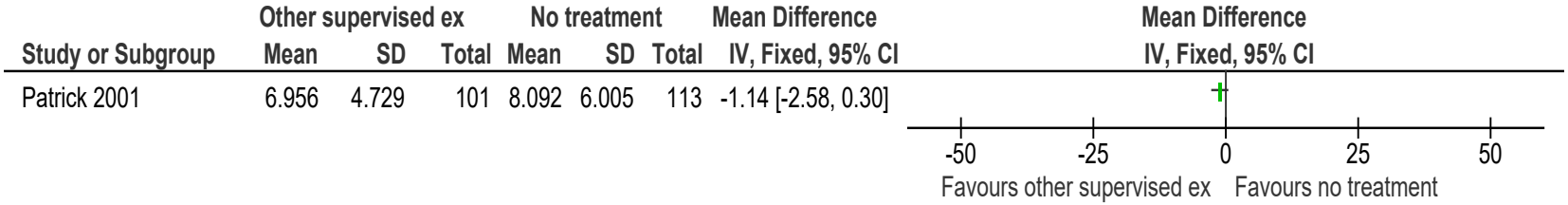


Figure 154: Serious adverse events at ≤3 months

Study or Subgroup	Other supervised ex Events	Total	No treatment Events	Total	Weight	Risk Difference M-H, Random, 95% CI
An 2008	0	14	0	14	26.2%	0.00 [-0.13, 0.13]
Cheung 2014	0	18	0	18	27.7%	0.00 [-0.10, 0.10]
Mcilroy 2017	0	7	0	7	19.5%	0.00 [-0.24, 0.24]
Rewald 2020	15	55	0	47	26.7%	0.27 [0.15, 0.39]
Total (95% CI)		**94**		**86**	**100.0%**	**0.07 [-0.10, 0.25]**
Total events	15		0			

Heterogeneity: Tau² = 0.03; Chi² = 19.67, df = 3 (P = 0.0002); I² = 85%

Test for overall effect: Z = 0.82 (P = 0.41)

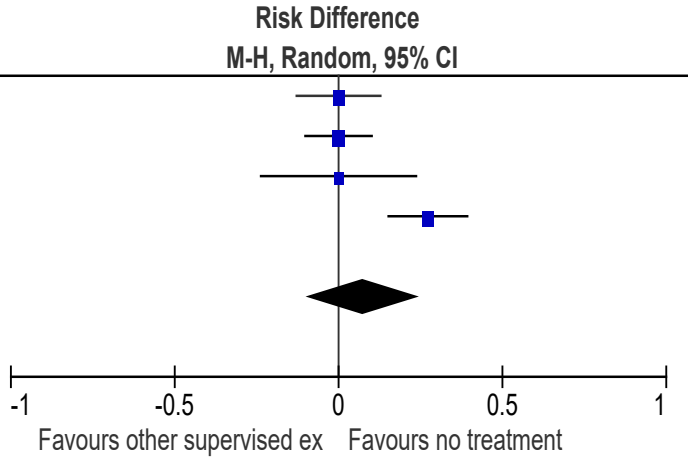
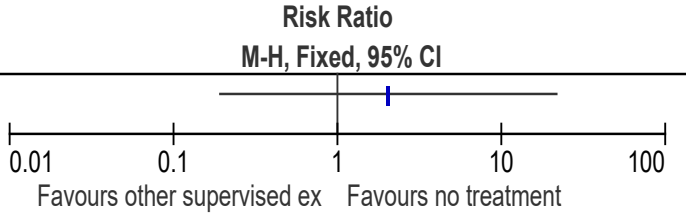


Figure 155: Serious adverse events at >3 months

Study or Subgroup	Other supervised ex Events	Total	No treatment Events	Total	Risk Ratio M-H, Fixed, 95% CI
Munukka 2016	2	43	1	44	2.05 [0.19, 21.75]



E.13 Other unsupervised exercise compared to unsupervised strength exercise

Figure 156: Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months

Study or Subgroup	Other unsupervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Chaipinyo 2009	6	16	24	23	20	18	-17.00 [-28.24, -5.76]

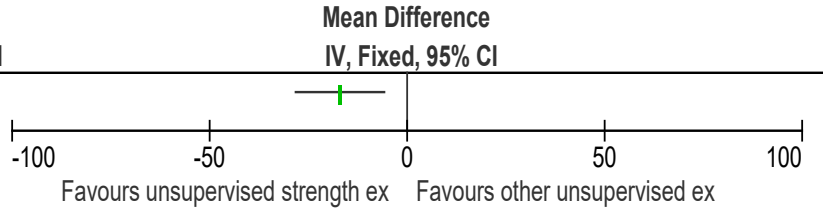


Figure 157: Pain (KOOS, 0-100, high is good, change score) at ≤3 months

Study or Subgroup	Other unsupervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Chaipinyo 2009	8	8	24	11	17	18	-3.00 [-11.48, 5.48]

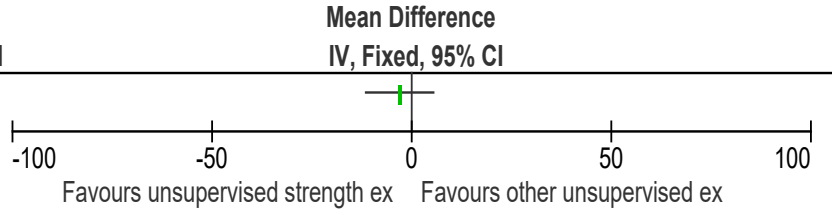
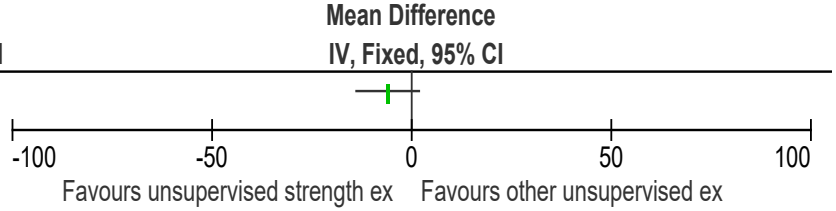


Figure 158: Physical function (KOOS, 0-100, high is good, change score) at ≤3 months

Study or Subgroup	Other unsupervised ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Chaipinyo 2009	7	14	24	13	12	18	-6.00 [-13.88, 1.88]



E.14 Supervised mixed modality exercise compared to supervised strength exercise

Figure 159: Quality of life (AQoL, 0-1, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Pazit 2018	0.71	0.16	10	0.72	0.19	10	-0.01 [-0.16, 0.14]

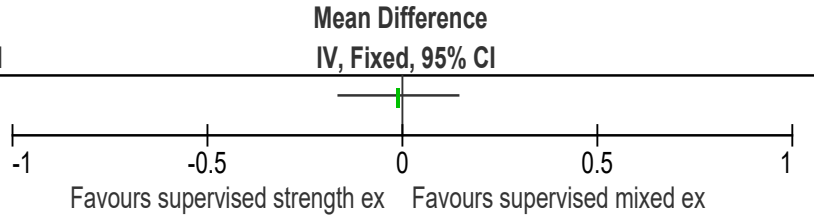


Figure 160: Quality of life (SF-36 physical function, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Diracoglu 2005	69.33	17.8	30	56.25	16.7	30	44.2%	13.08 [4.35, 21.81]
Vaghela 2020	7.6	1.41	43	7.55	1.1	40	55.8%	0.05 [-0.49, 0.59]
Total (95% CI)			**73**			**70**	**100.0%**	**5.81 [-6.88, 18.49]**

Heterogeneity: Tau² = 74.92; Chi² = 8.52, df = 1 (P = 0.004); I² = 88%

Test for overall effect: Z = 0.90 (P = 0.37)

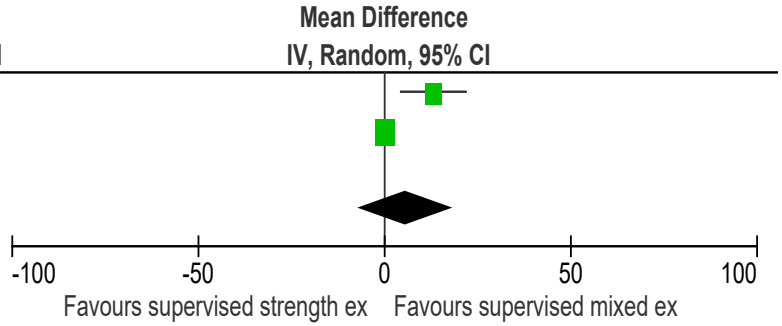


Figure 161: Quality of life (SF-36 role physical, 0-100, high is good, final values) at ≤3 months

	Supervised mixed ex			Supervised strength ex				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Diracoglu 2005	77.5	34.9	30	57.14	45	30	34.5%	20.36 [-0.02, 40.74]
Vaghela 2020	28.58	1.35	43	26.85	2.23	40	65.5%	1.73 [0.93, 2.53]
Total (95% CI)			**73**			**70**	**100.0%**	**8.15 [-9.20, 25.50]**

Heterogeneity: Tau² = 119.40; Chi² = 3.21, df = 1 (P = 0.07); I² = 69%
Test for overall effect: Z = 0.92 (P = 0.36)

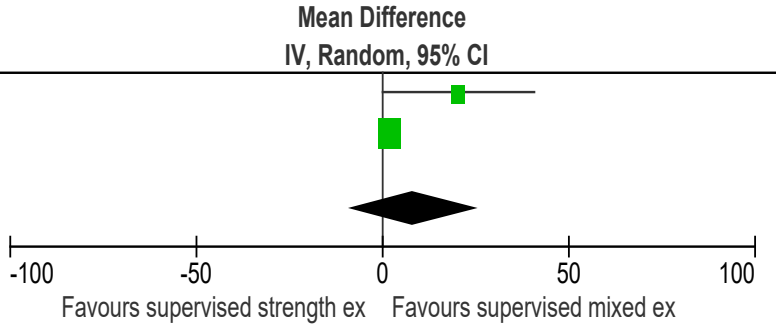


Figure 162: Quality of life (SF-36 vitality, 0-100, high is good, final values) at ≤3 months

	Supervised mixed ex			Supervised strength ex				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Diracoglu 2005	54	19.5	30	43.5	18.3	30	23.0%	10.50 [0.93, 20.07]
Diracoglud 2008	54	19.58	33	45.54	18.33	33	24.2%	8.46 [-0.69, 17.61]
Vaghela 2020	10.69	0.46	43	8.9	0.84	40	52.8%	1.79 [1.50, 2.08]
Total (95% CI)			**106**			**103**	**100.0%**	**5.40 [-0.70, 11.51]**

Heterogeneity: Tau² = 18.32; Chi² = 5.21, df = 2 (P = 0.07); I² = 62%
Test for overall effect: Z = 1.74 (P = 0.08)

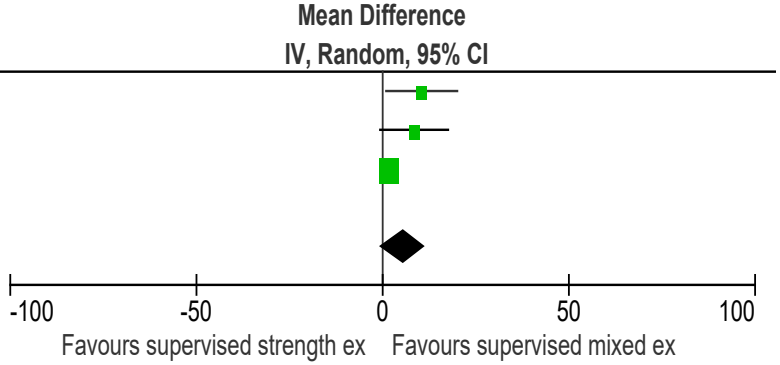


Figure 163: Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months

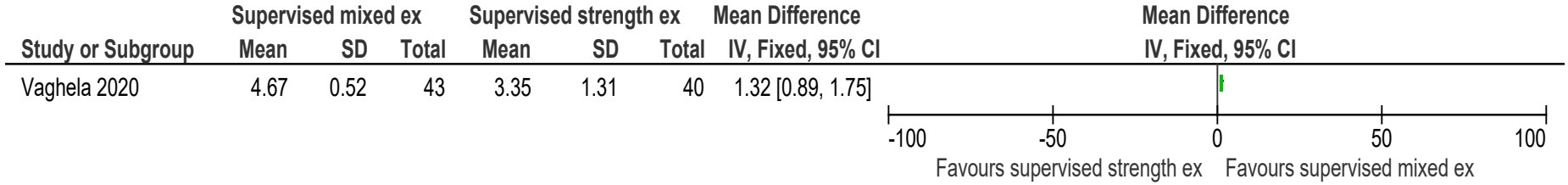


Figure 164: Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months

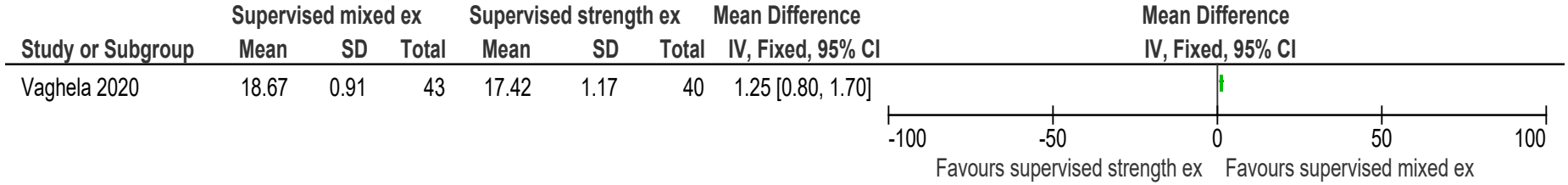


Figure 165: Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months

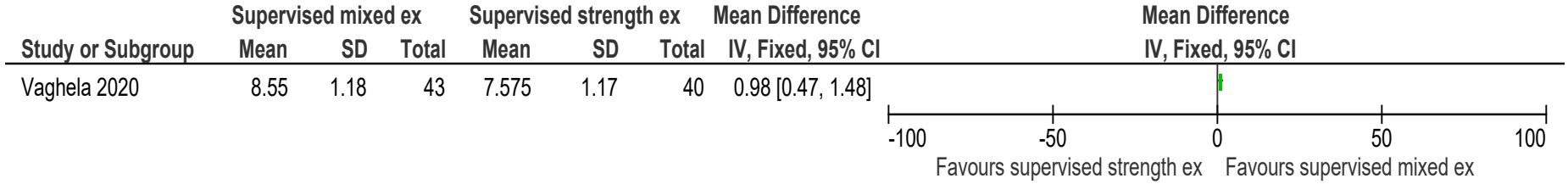


Figure 166: Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Supervised strength ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Vaghela 2020	7.79	0.41	43	7.35	0.8	40	0.44 [0.16, 0.72]	

Figure 167: Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Supervised strength ex			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI	IV, Fixed, 95% CI
Vaghela 2020	31.76	5.35	43	31.15	8.61	40	0.61 [-2.50, 3.72]	

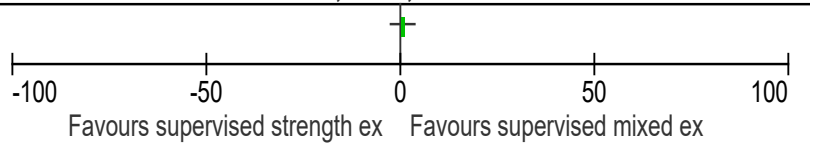


Figure 168: Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Diracoglud 2008	51.07	15.54	33	43.82	11.93	33	7.25 [0.57, 13.93]

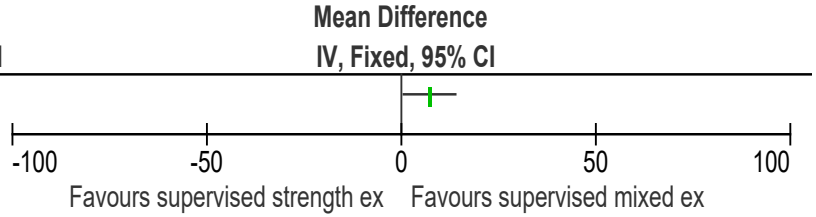


Figure 169: Pain (WOMAC, NRS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Cantero-tellez 2021	3.67	0.81	6	3.5	1.05	6	6.9%	0.17 [-0.97, 1.30]
Diracoglud 2008	1.33	0.58	33	1.5	0.57	33	11.5%	-0.29 [-0.78, 0.19]
Hernandez 2019	2.42	2.35	25	4	2.83	22	10.8%	-0.60 [-1.19, -0.01]
Joshi 2019	2.85	0.88	21	5.2	1.17	21	9.3%	-2.23 [-3.01, -1.44]
Knoop 2013	2.8	2.1	80	3.3	2.1	79	12.6%	-0.24 [-0.55, 0.08]
Kumar 2013	2.18	0.66	22	2.91	0.81	22	10.5%	-0.97 [-1.60, -0.34]
Pazit 2018	97.3	127.1	10	117	132.6	10	8.6%	-0.15 [-1.02, 0.73]
Rogers 2012	5	3.35	11	4.25	3.45	11	8.9%	0.21 [-0.63, 1.05]
Sedaghatnezhad 2021	13.67	10.06	15	25.89	16.69	15	9.5%	-0.86 [-1.62, -0.11]
Vaghela 2020	4.32	1.93	43	7.3	2.13	40	11.5%	-1.46 [-1.94, -0.97]
Total (95% CI)			**266**			**259**	**100.0%**	**-0.67 [-1.09, -0.24]**

Heterogeneity: Tau² = 0.35; Chi² = 43.30, df = 9 (P < 0.00001); I² = 79%
Test for overall effect: Z = 3.08 (P = 0.002)

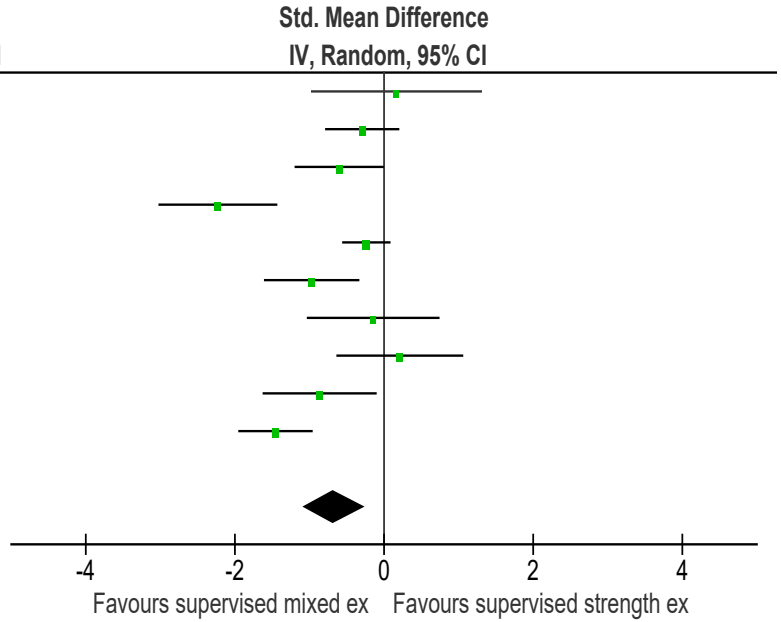


Figure 170: Pain (WOMAC, VAS, NRS [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised strength ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Diracoglud 2008	1.26	0.62	33	1.7	0.69	33	29.8%	-0.66 [-1.16, -0.17]
Hernandez 2019	3.8	2.97	23	3.63	2.8	20	23.3%	0.06 [-0.54, 0.66]
Knoop 2013	3.1	2.5	80	3.7	2.4	79	46.9%	-0.24 [-0.56, 0.07]
Total (95% CI)			**136**			**132**	**100.0%**	**-0.30 [-0.65, 0.05]**

Heterogeneity: Tau² = 0.04; Chi² = 3.54, df = 2 (P = 0.17); I² = 44%
Test for overall effect: Z = 1.68 (P = 0.09)

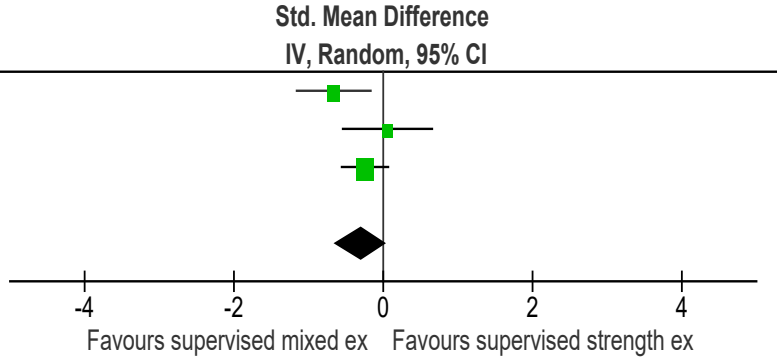


Figure 171: Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Supervised mixed ex Total	Supervised strength ex Total	Weight	Std. Mean Difference IV, Random, 95% CI
Diracoglu 2005	-0.4596	0.2618	30	30	12.1%	-0.46 [-0.97, 0.05]
Diracoglud 2008	-0.5779	0.2516	33	33	12.2%	-0.58 [-1.07, -0.08]
Hernandez 2019	-0.7721	0.3039	25	22	11.5%	-0.77 [-1.37, -0.18]
Joshi 2019	-3.9344	0.5464	21	21	8.1%	-3.93 [-5.01, -2.86]
Knoop 2013	-0.1644	0.1589	80	79	13.3%	-0.16 [-0.48, 0.15]
Kumar 2013	-1.0081	0.3219	22	22	11.2%	-1.01 [-1.64, -0.38]
Pazit 2018	-0.3986	0.4527	10	10	9.4%	-0.40 [-1.29, 0.49]
Rogers 2012	-0.2047	0.4278	11	11	9.7%	-0.20 [-1.04, 0.63]
Vaghela 2020	-0.8478	0.2298	43	40	12.5%	-0.85 [-1.30, -0.40]
Total (95% CI)			**275**	**268**	**100.0%**	**-0.83 [-1.30, -0.36]**

Heterogeneity: Tau² = 0.41; Chi² = 49.18, df = 8 (P < 0.00001); I² = 84%
Test for overall effect: Z = 3.48 (P = 0.0005)

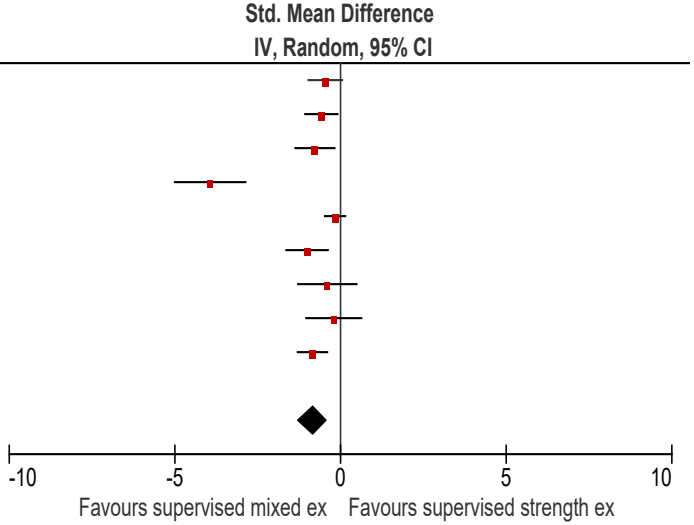


Figure 172: Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Std. Mean Difference	SE	Supervised mixed ex Total	Supervised strength ex Total	Weight	Std. Mean Difference IV, Random, 95% CI
Diracoglud 2008	-0.7833	0.256	33	33	32.7%	-0.78 [-1.29, -0.28]
Hernandez 2019	-0.8111	0.3192	23	20	28.8%	-0.81 [-1.44, -0.19]
Knoop 2013	-0.0225	0.1586	80	79	38.5%	-0.02 [-0.33, 0.29]
Total (95% CI)			**136**	**132**	**100.0%**	**-0.50 [-1.08, 0.08]**

Heterogeneity: Tau² = 0.20; Chi² = 9.16, df = 2 (P = 0.01); I² = 78%

Test for overall effect: Z = 1.68 (P = 0.09)

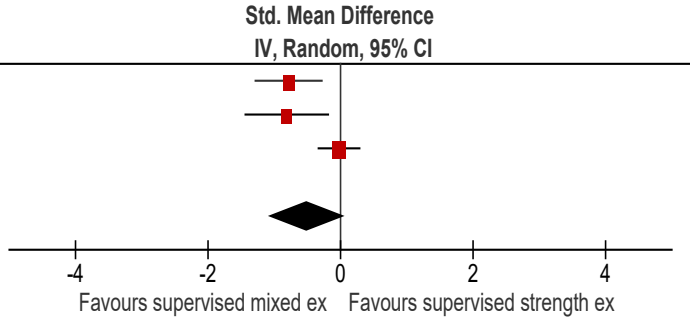


Figure 173: Serious adverse events at ≤3 months

Study or Subgroup	Supervised mixed ex Events	Total	Supervised strength ex Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Diracoglu 2005	0	30	0	30	31.2%	0.00 [-0.06, 0.06]
Hernandez 2019	0	53	0	60	58.5%	0.00 [-0.03, 0.03]
Pazit 2018	0	10	0	10	10.4%	0.00 [-0.17, 0.17]
Total (95% CI)		**93**		**100**	**100.0%**	**0.00 [-0.03, 0.03]**
Total events	0		0			

Heterogeneity: Chi² = 0.00, df = 2 (P = 1.00); I² = 0%

Test for overall effect: Z = 0.00 (P = 1.00)

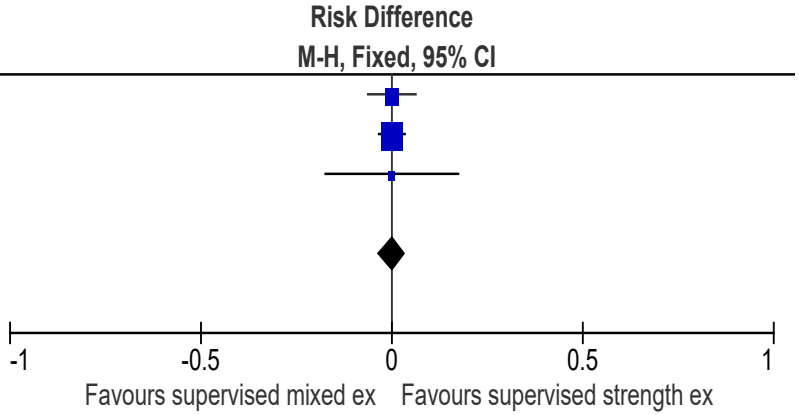
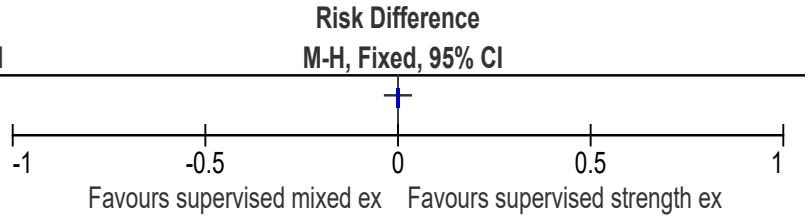


Figure 174: Serious adverse events at >3 months

Study or Subgroup	Supervised mixed ex Events	Total	Supervised strength ex Events	Total	Risk Difference M-H, Fixed, 95% CI
Hernandez 2019	0	53	0	60	0.00 [-0.03, 0.03]



E.15 Supervised mixed modality exercise compared to unsupervised strength exercise

Figure 175: Quality of life (SF-36 physical component, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Lim 2010	40.4	7.9	22	36.9	9.6	20	3.50 [-1.85, 8.85]

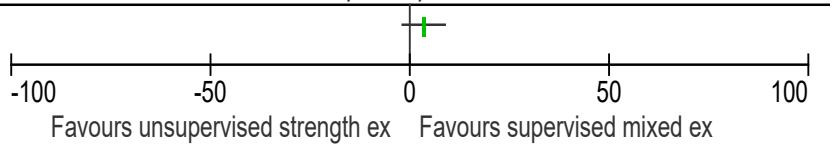


Figure 176: Quality of life (SF-36 mental component, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised strength ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Lim 2010	52.9	8.3	22	48.4	14.3	20	4.50 [-2.66, 11.66]

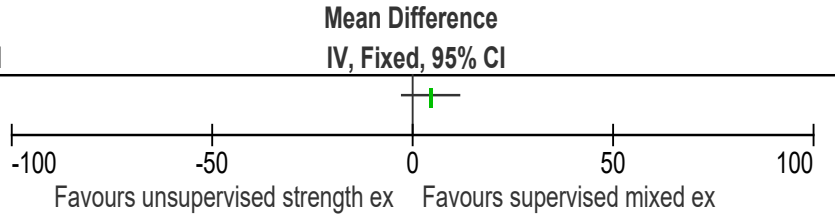
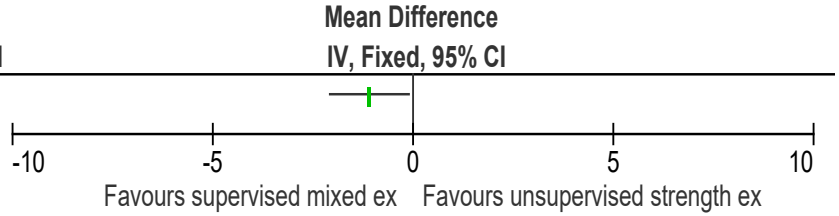


Figure 177: Pain (BPI mean pain, 0-10, high is poor, final value) at ≤3 months

	Supervised mixed ex			Unsupervised strength ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Lim 2010	3.46	1.3	22	4.55	1.88	20	-1.09 [-2.08, -0.10]



E.16 Supervised mixed modality exercise compared to supervised aerobic exercise

Figure 178: Pain (WOMAC, 0-20, high is poor, change score) at >3 months

	Supervised mixed ex			Supervised aerobic ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Hunt 2018	-2.5	3	39	-1.5	3.2	40	-1.00 [-2.37, 0.37]

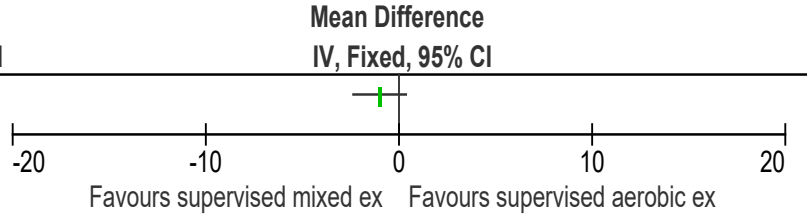
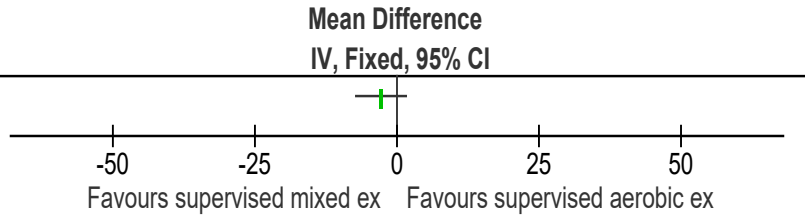


Figure 179: Physical function (WOMAC, 0-68, high is poor, change score) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Supervised aerobic ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Hunt 2018	-9.4	9.7	39	-6.6	10.3	40	-2.80 [-7.21, 1.61]



E.17 Supervised mixed modality exercise compared to other supervised exercise

Figure 180: Quality of life (KOOS, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Wang 2011	74	11	26	73	12	26	1.00 [-5.26, 7.26]

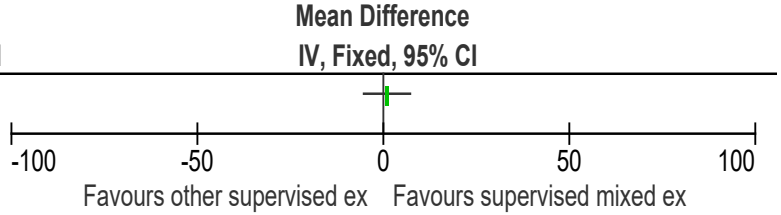


Figure 181: Quality of life (EQ-5D, -0.11-1, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Holm 2020	0.72	0.12	45	0.75	0.1	45	-0.03 [-0.08, 0.02]

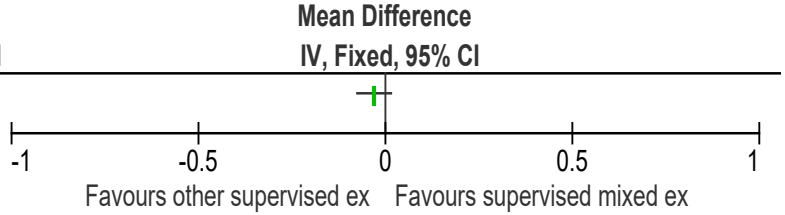


Figure 182: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months

	Supervised mixed ex			Other supervised ex				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Cheung 2017	38.8	9	28	41.5	8.5	32	50.7%	-2.70 [-7.15, 1.75]
Lim 2010	40.4	7.9	22	38.8	7.7	24	49.3%	1.60 [-2.92, 6.12]
Total (95% CI)			**50**			**56**	**100.0%**	**-0.58 [-3.75, 2.59]**

Heterogeneity: Chi² = 1.77, df = 1 (P = 0.18); I² = 43%
Test for overall effect: Z = 0.36 (P = 0.72)

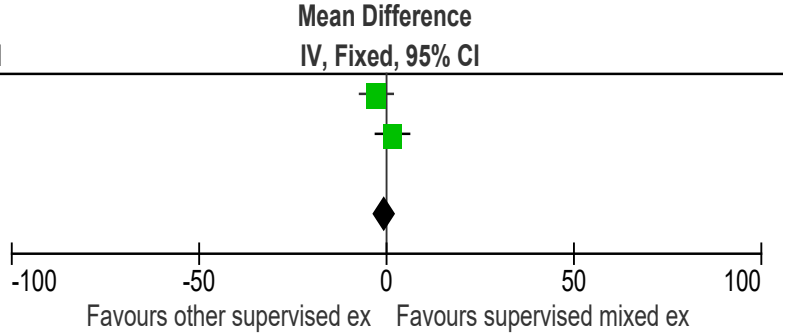


Figure 183: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months

	Supervised mixed ex			Other supervised ex				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Cheung 2017	53.8	9.2	28	55.2	8.7	32	54.1%	-1.40 [-5.95, 3.15]
Lim 2010	52.9	8.3	22	54.8	8.8	24	45.9%	-1.90 [-6.84, 3.04]
Total (95% CI)			**50**			**56**	**100.0%**	**-1.63 [-4.98, 1.72]**

Heterogeneity: Chi² = 0.02, df = 1 (P = 0.88); I² = 0%
Test for overall effect: Z = 0.95 (P = 0.34)

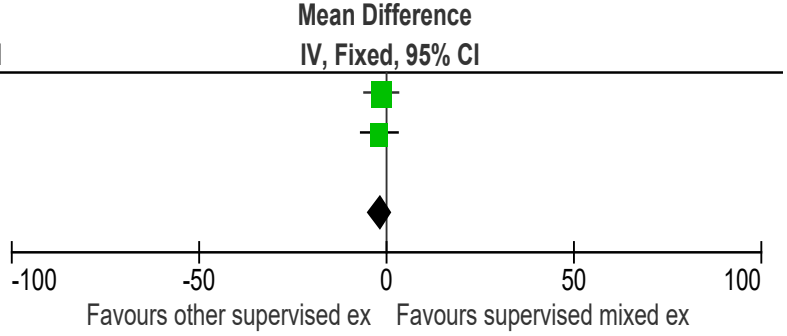


Figure 184: Pain (KOOS, WOMAC, BPI, VAS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Cheung 2017	6.5	2.6	28	5.1	2.7	32	17.4%	0.52 [0.00, 1.04]
Holm 2020	-58.5	14.7	45	-61.2	13.7	45	27.1%	0.19 [-0.23, 0.60]
Lim 2010	3.46	1.3	22	3.27	1.67	24	13.9%	0.12 [-0.46, 0.70]
Rogers 2012	5	3.35	11	4.87	3.6	11	6.7%	0.04 [-0.80, 0.87]
Silva 2008	38.4	27.5	32	37	18.1	32	19.3%	0.06 [-0.43, 0.55]
Wang 2011	-76	15	26	-72	18	26	15.6%	-0.24 [-0.78, 0.31]
Total (95% CI)			**164**			**170**	**100.0%**	**0.14 [-0.08, 0.35]**

Heterogeneity: Chi² = 4.15, df = 5 (P = 0.53); I² = 0%
Test for overall effect: Z = 1.23 (P = 0.22)

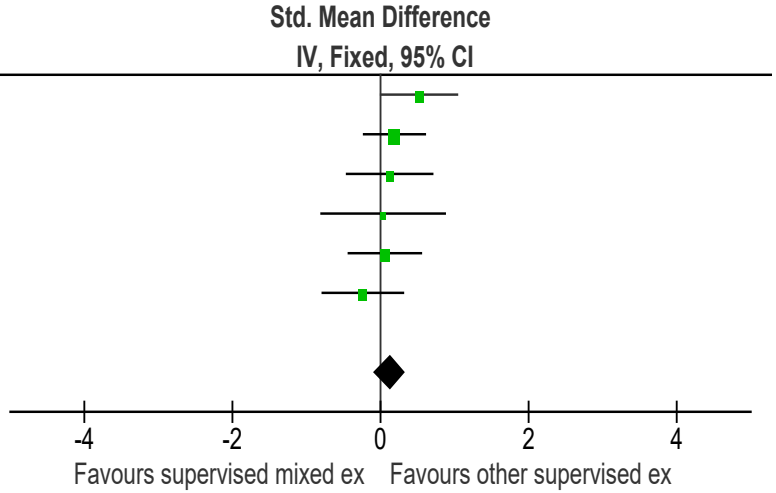


Figure 185: Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Silva 2008	37.3	27.5	32	26.7	23.1	32	47.0%	0.41 [-0.08, 0.91]
Xiao 2020	5	3.4	45	5.4	3.5	40	53.0%	-0.11 [-0.54, 0.31]
Total (95% CI)			**77**			**72**	**100.0%**	**0.13 [-0.38, 0.65]**

Heterogeneity: Tau² = 0.08; Chi² = 2.50, df = 1 (P = 0.11); I² = 60%
Test for overall effect: Z = 0.51 (P = 0.61)

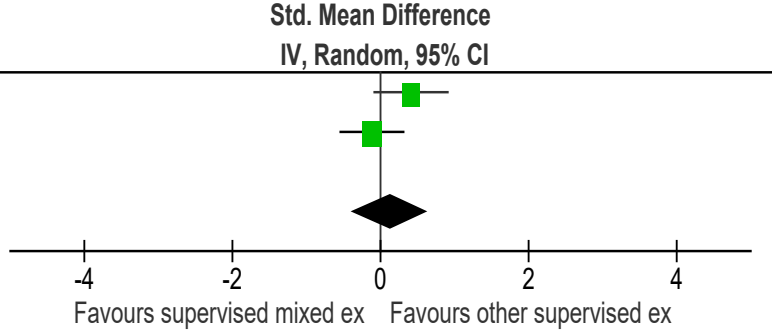


Figure 186: Physical function (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Cheung 2017	25.8	8.4	28	18.2	8.4	32	26.1%	0.89 [0.36, 1.43]
Holm 2020	-67	13	45	-68.1	14	45	28.3%	0.08 [-0.33, 0.49]
Rogers 2012	13.89	9.44	11	20	9.2	11	19.8%	-0.63 [-1.49, 0.23]
Wang 2011	-82	14	26	-76	16	26	25.8%	-0.39 [-0.94, 0.16]
Total (95% CI)			**110**			**114**	**100.0%**	**0.03 [-0.58, 0.64]**

Heterogeneity: Tau² = 0.30; Chi² = 14.36, df = 3 (P = 0.002); I² = 79%

Test for overall effect: Z = 0.09 (P = 0.93)

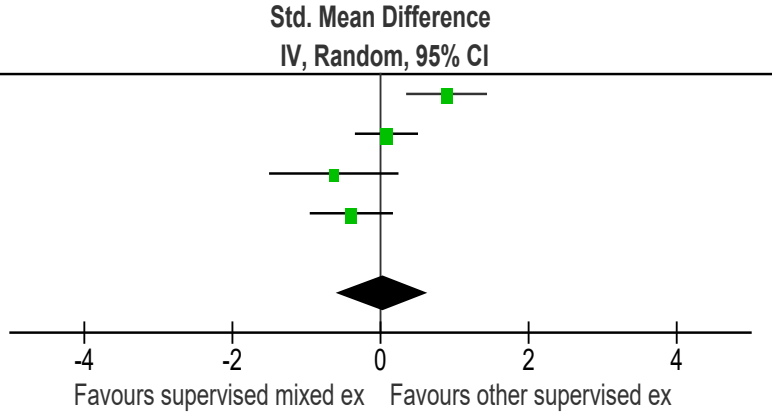


Figure 187: Physical function (WOMAC, 0-68, high is poor, final value) at <3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Other supervised ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Xiao 2020	20.7	8.7	45	18.8	7.4	40	1.90 [-1.52, 5.32]

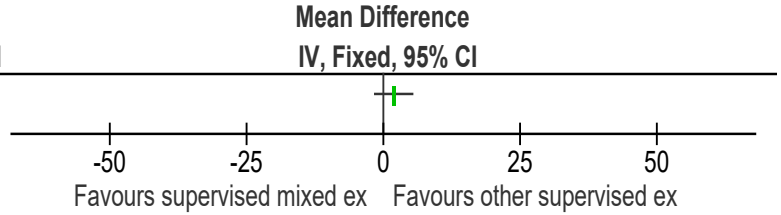


Figure 188: Psychological distress (HADS-anxiety, 0-21, high is poor, final value) at ≤3 months

	Supervised mixed ex			Other supervised ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Cheung 2017	5.2	2.7	28	3.8	2.6	32	1.40 [0.05, 2.75]

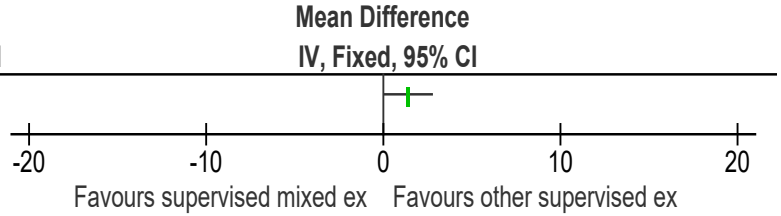


Figure 189: Psychological distress (HADS-depression, 0-21, high is poor, final value) at ≤3 months

	Supervised mixed ex			Other supervised ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Cheung 2017	4.2	2	28	3.8	2	32	0.40 [-0.61, 1.41]

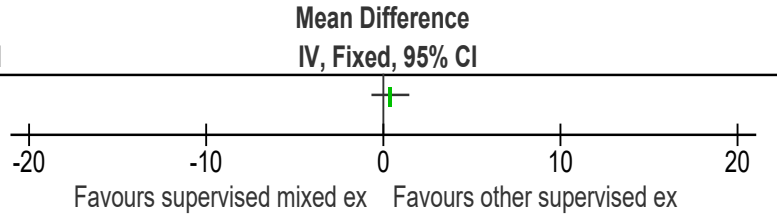


Figure 190: Serious adverse events at ≤3 months

	Supervised mixed ex		Other supervised ex		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	M-H, Fixed, 95% CI
Holm 2020	3	45	5	45	0.60 [0.15, 2.36]

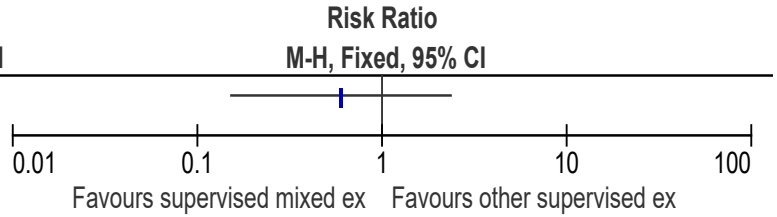
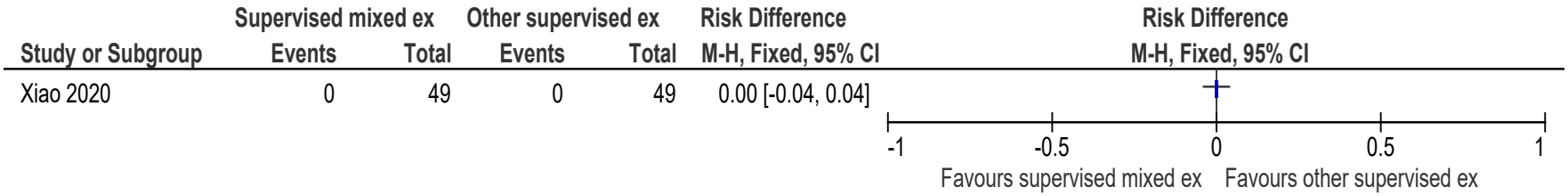


Figure 191: Serious adverse events at >3 months



E.18 Supervised mixed modality exercise compared to unsupervised mixed modality exercise

Figure 192: Quality of life (SF-36 physical function, 0-100, high is good, final value) at ≤3 months

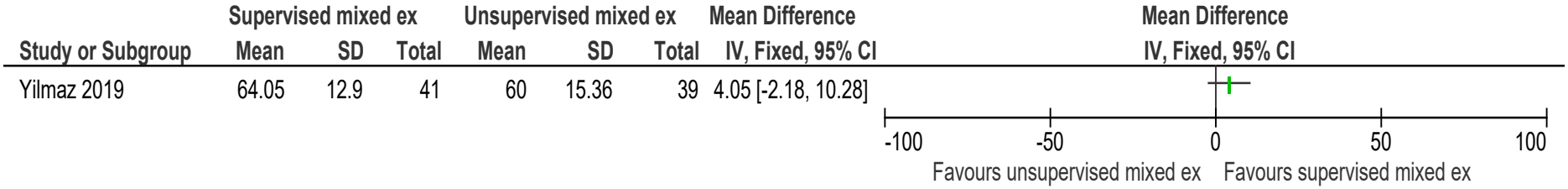


Figure 193: Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	74.1	12.63	41	64.11	21.56	39	9.99 [2.20, 17.78]

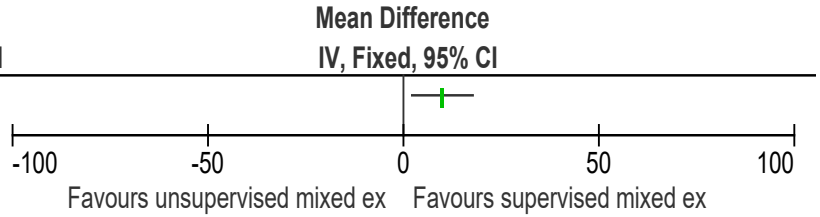


Figure 194: Quality of life (SF-36 role physical, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	77.38	28.4	41	61.84	32.66	39	15.54 [2.10, 28.98]

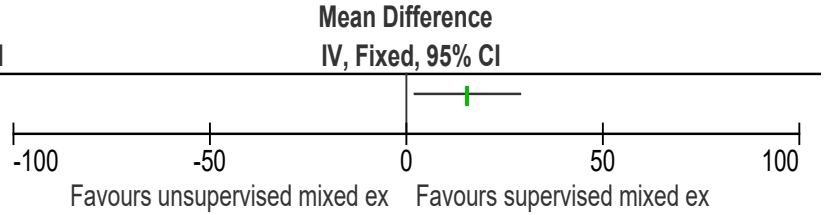


Figure 195: Quality of life (SF-36 vitality, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	51.67	23.41	41	50	22.28	39	1.67 [-8.34, 11.68]

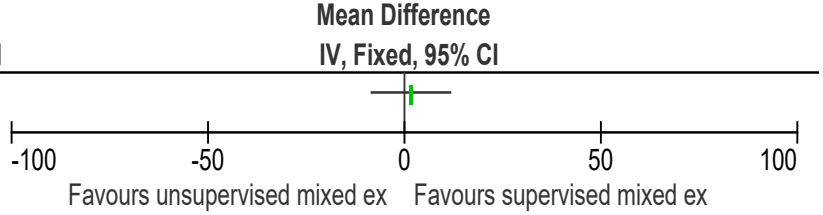


Figure 196: Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	67.62	15.4	41	57.89	16	39	9.73 [2.84, 16.62]

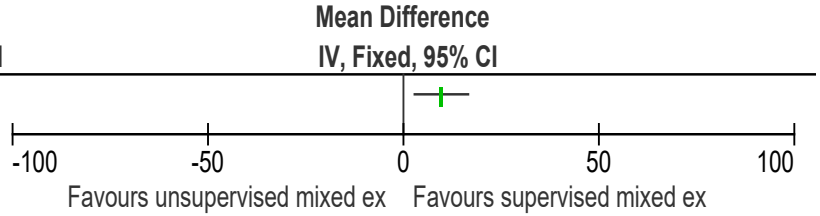


Figure 197: Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	75.62	17.1	41	75.62	17.1	39	0.00 [-7.50, 7.50]

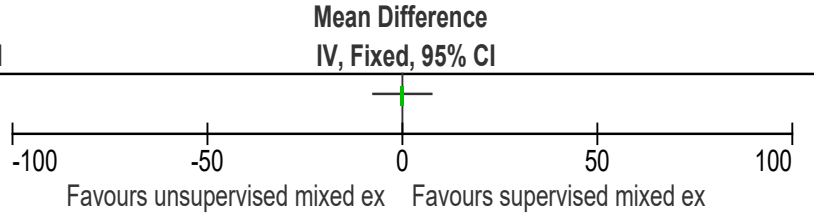


Figure 198: Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months

	Supervised mixed ex			Unsupervised mixed ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Yilmaz 2019	87.19	24.83	41	61.21	38.96	39	25.98 [11.58, 40.38]

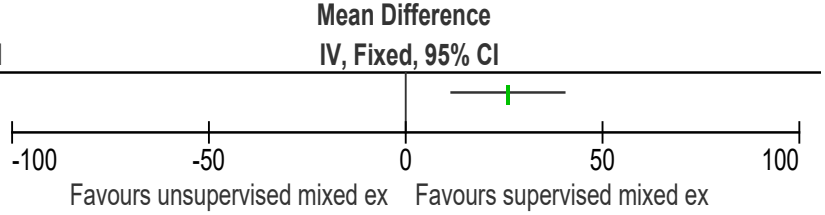


Figure 199: Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Unsupervised mixed ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Yilmaz 2019	5.95	16.31	41	65.53	17.61	39	-59.58 [-67.03, -52.13]

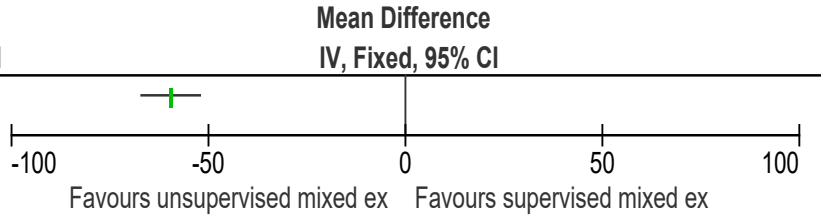


Figure 200: Pain (WOMAC, VAS [different scale ranges], high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Unsupervised mixed ex Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Tunay 2010	1.56	1.83	30	2.6	1.82	30	42.0%	-0.56 [-1.08, -0.05]
Yilmaz 2019	5.95	3.2	41	6.74	4.64	39	58.0%	-0.20 [-0.64, 0.24]
Total (95% CI)			**71**			**69**	**100.0%**	**-0.35 [-0.69, -0.02]**

Heterogeneity: Chi² = 1.11, df = 1 (P = 0.29); I² = 10%
Test for overall effect: Z = 2.05 (P = 0.04)

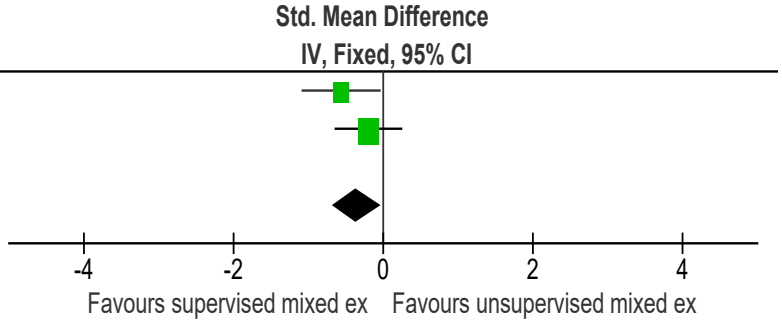
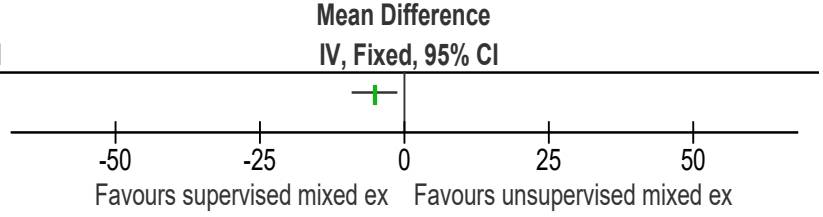


Figure 201: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Unsupervised mixed ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Yilmaz 2019	13.71	9.01	41	18.89	8.29	39	-5.18 [-8.97, -1.39]



E.19 Supervised mixed modality exercise compared to pharmacological treatment

Figure 202: Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months

	Supervised mixed ex			Pharmacological treatment			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Holsgaard-larsen 2017	3.14	12.63	47	4.5	13.05	46	-1.36 [-6.58, 3.86]

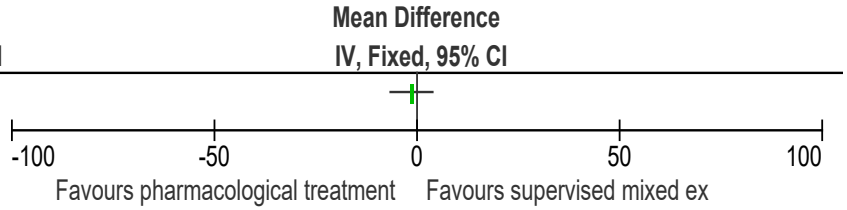


Figure 203: Quality of life (KOOS, 0-100, high is good, final value) at >3 months

	Supervised mixed ex			Pharmacological treatment			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Holsgaard-larsen 2017	10	15.1	47	8.7	15.4	46	1.30 [-4.90, 7.50]

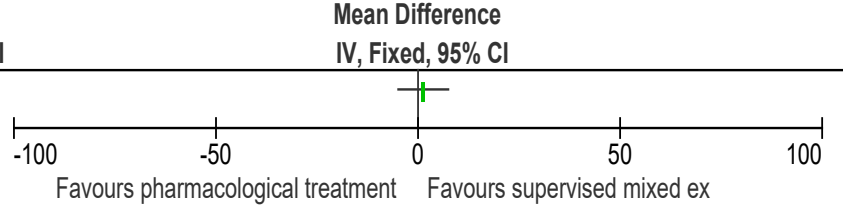


Figure 204: Pain (KOOS, 0-100, high is good, change score) at ≤3 months

	Supervised mixed ex			Pharmacological treatment			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Holsgaard-larsen 2017	7.23	10.77	47	5.15	10.68	46	2.08 [-2.28, 6.44]

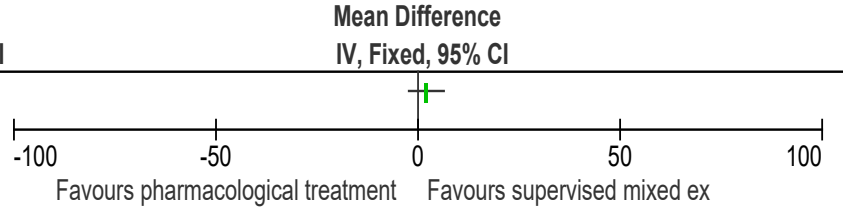


Figure 205: Pain (WOMAC, 0-500, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Saccomanno 2016	154.6	92	51	177.7	100.5	53	-23.10 [-60.11, 13.91]

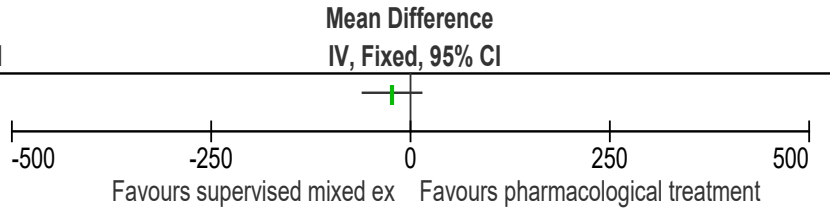


Figure 206: Pain (KOOS, 0-100, high is good, change score) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Holsgaard-larsen 2017	13.6	13.7	47	9.4	14.1	46	4.20 [-1.45, 9.85]

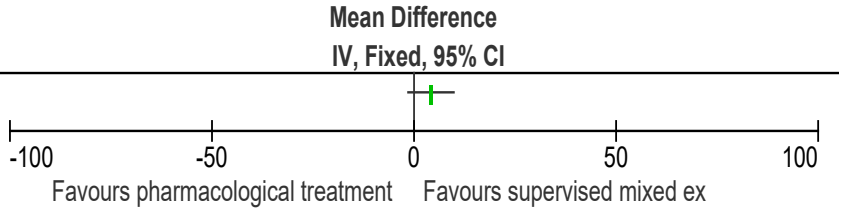


Figure 207: Pain (WOMAC, 0-500, high is poor, final value) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Saccomanno 2016	161.6	90.2	51	181.5	98	53	-19.90 [-56.08, 16.28]

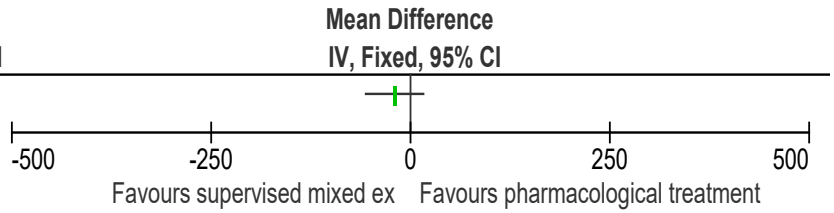


Figure 208: Physical function (KOOS, 0-100, high is good, change score) at ≤3 months

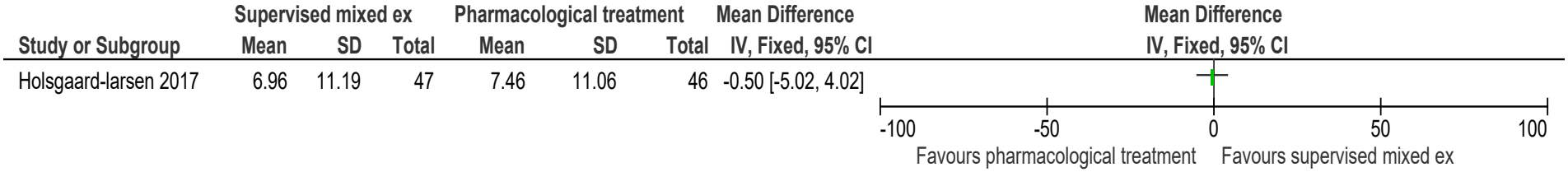


Figure 209: Physical function (WOMAC, 0-1800, high is poor, final value) at ≤3 months

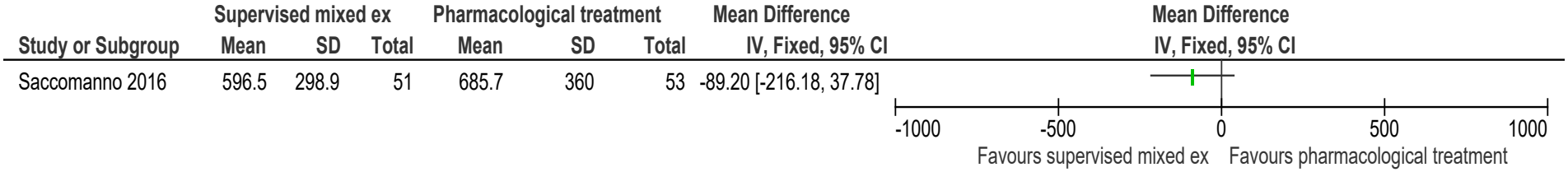


Figure 210: Physical function (KOOS, 0-100, high is good, change score) at >3 months

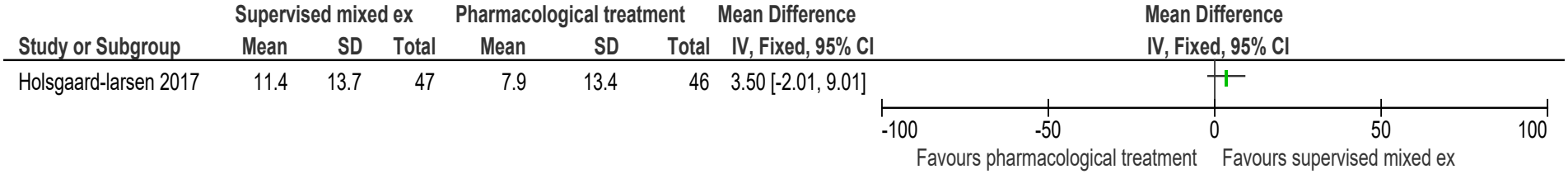


Figure 211: Physical function (WOMAC, 0-1800, high is poor, final value) at >3 months

	Supervised mixed ex			Pharmacological treatment			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Saccomanno 2016	618.5	310.4	51	691.4	363.8	53	-72.90 [-202.71, 56.91]

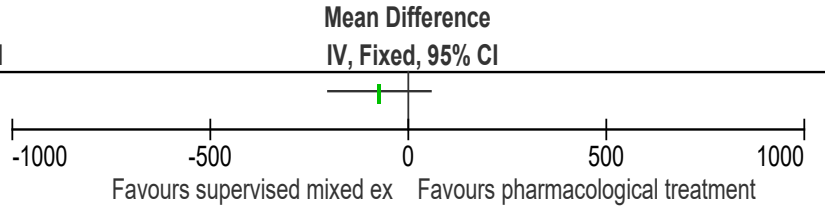
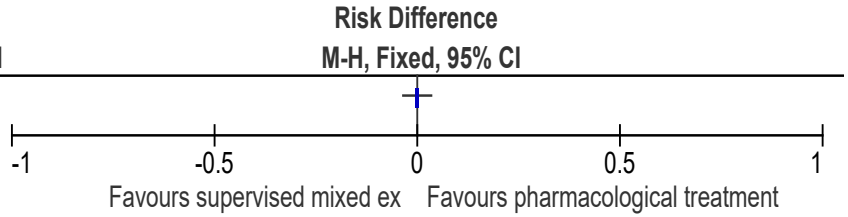


Figure 212: Serious adverse events at >3 months

	Supervised mixed ex		Pharmacological treatment		Risk Difference
Study or Subgroup	Events	Total	Events	Total	M-H, Fixed, 95% CI
Saccomanno 2016	0	55	0	55	0.00 [-0.03, 0.03]



E.20 Supervised mixed modality exercise compared to no treatment

Figure 213: Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Pazit 2018	0.71	0.16	10	0.63	0.12	10	27.8%	0.54 [-0.35, 1.44]
Wang 2011	74	11	26	67	13	26	72.2%	0.57 [0.02, 1.13]
Total (95% CI)			**36**			**36**	**100.0%**	**0.56 [0.09, 1.04]**

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.95); I² = 0%
Test for overall effect: Z = 2.34 (P = 0.02)

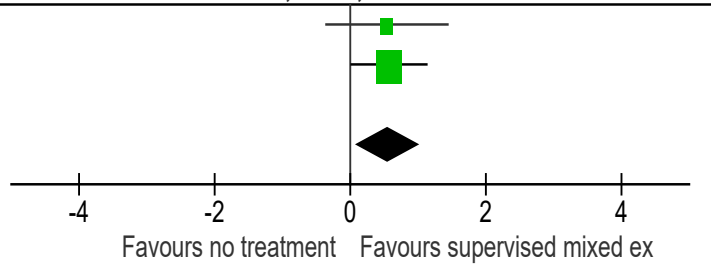


Figure 214: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Cheung 2017	38.8	9	28	39	8.4	23	45.6%	-0.20 [-4.99, 4.59]
French 2013	37.03	11.25	45	33.82	9.67	43	54.4%	3.21 [-1.17, 7.59]
Total (95% CI)			**73**			**66**	**100.0%**	**1.66 [-1.57, 4.89]**

Heterogeneity: Chi² = 1.06, df = 1 (P = 0.30); I² = 6%
Test for overall effect: Z = 1.01 (P = 0.31)

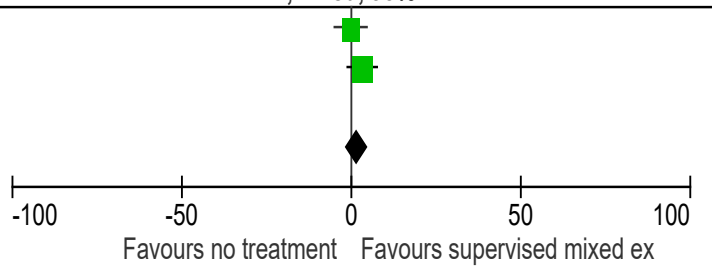


Figure 215: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Cheung 2017	53.8	9.2	28	52.8	8.8	23	55.2%	1.00 [-3.95, 5.95]
French 2013	48.92	12.5	45	48.52	13.75	43	44.8%	0.40 [-5.10, 5.90]
Total (95% CI)			**73**			**66**	**100.0%**	**0.73 [-2.95, 4.41]**

Heterogeneity: Chi² = 0.03, df = 1 (P = 0.87); I² = 0%
Test for overall effect: Z = 0.39 (P = 0.70)

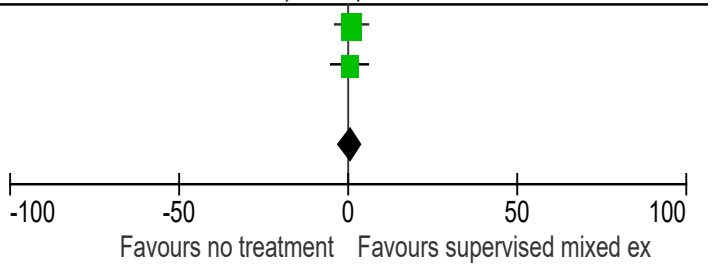


Figure 216: Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Aglamis 2008	87.2	9.7	16	36.4	7.9	9	49.9%	50.80 [43.78, 57.82]
Kraus 2014	2	14	71	2	18	69	50.1%	0.00 [-5.35, 5.35]
Total (95% CI)			**87**			**78**	**100.0%**	**25.35 [-24.44, 75.13]**

Heterogeneity: Tau² = 1280.18; Chi² = 127.31, df = 1 (P < 0.00001); I² = 99%
Test for overall effect: Z = 1.00 (P = 0.32)

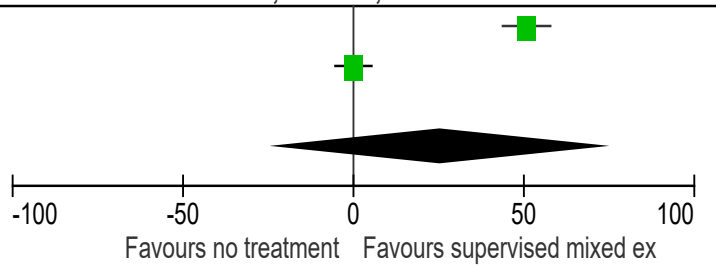


Figure 217: Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Aglamis 2008	67.5	18.1	16	20	17.5	9	48.7%	47.50 [33.03, 61.97]
Kraus 2014	5.2	17.6	70	-0.1	17.3	68	51.3%	5.30 [-0.52, 11.12]
Total (95% CI)			**86**			**77**	**100.0%**	**25.86 [-15.48, 67.20]**

Heterogeneity: Tau² = 858.76; Chi² = 28.12, df = 1 (P < 0.00001); I² = 96%
Test for overall effect: Z = 1.23 (P = 0.22)

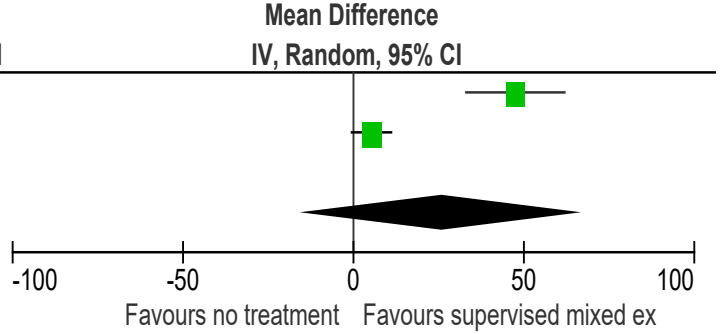


Figure 218: Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Aglamis 2008	90.6	25.6	16	5.6	11	9	49.9%	85.00 [70.54, 99.46]
Kraus 2014	2	35	71	3	33	69	50.1%	-1.00 [-12.27, 10.27]
Total (95% CI)			**87**			**78**	**100.0%**	**41.88 [-42.40, 126.15]**

Heterogeneity: Tau² = 3654.28; Chi² = 84.58, df = 1 (P < 0.00001); I² = 99%
Test for overall effect: Z = 0.97 (P = 0.33)

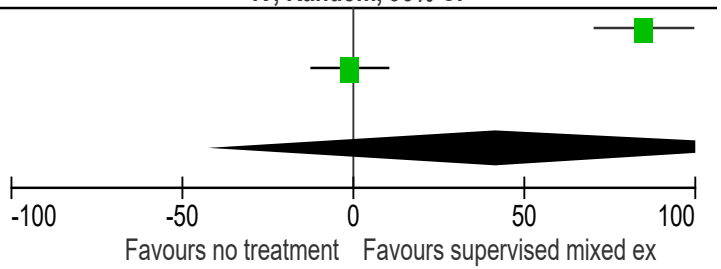


Figure 219: Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at ≤3 months

	Supervised mixed ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Aglamis 2008	76.3	9.9	16	24.4	25.2	9	48.7%	51.90 [34.74, 69.06]
Kraus 2014	-1	15	71	0	12	69	51.3%	-1.00 [-5.49, 3.49]
Total (95% CI)			**87**			**78**	**100.0%**	**24.77 [-27.05, 76.60]**

Heterogeneity: Tau² = 1358.23; Chi² = 34.15, df = 1 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.94 (P = 0.35)

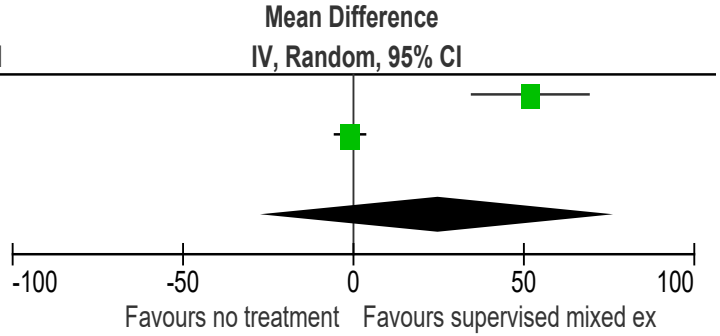


Figure 220: Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at ≤3 months

	Supervised mixed ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Aglamis 2008	77.5	10.2	16	40	20.5	9	48.0%	37.50 [23.20, 51.80]
Kraus 2014	3	14	71	0	16	69	52.0%	3.00 [-1.99, 7.99]
Total (95% CI)			**87**			**78**	**100.0%**	**19.57 [-14.21, 53.36]**

Heterogeneity: Tau² = 565.29; Chi² = 19.95, df = 1 (P < 0.00001); I² = 95%
Test for overall effect: Z = 1.14 (P = 0.26)

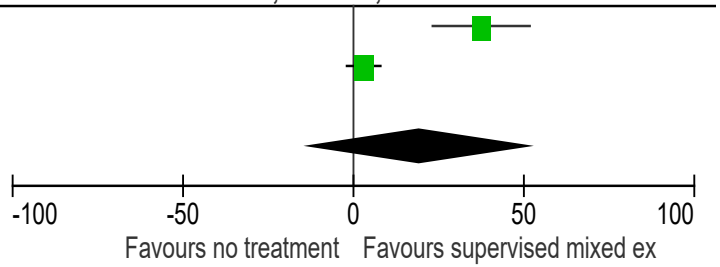


Figure 221: Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at ≤3 months

	Supervised mixed ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Aglamis 2008	79.3	8	16	46.4	13.8	9	48.9%	32.90 [23.07, 42.73]
Kraus 2014	-1	11	71	-2	10	69	51.1%	1.00 [-2.48, 4.48]
Total (95% CI)			**87**			**78**	**100.0%**	**16.61 [-14.65, 47.86]**

Heterogeneity: Tau² = 494.65; Chi² = 35.94, df = 1 (P < 0.00001); I² = 97%
Test for overall effect: Z = 1.04 (P = 0.30)

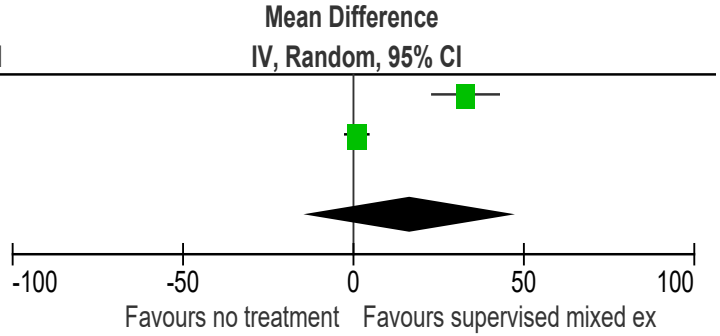


Figure 222: Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at ≤3 months

	Supervised mixed ex			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Aglamis 2008	87.5	26.9	16	14.7	33.7	9	48.5%	72.80 [47.14, 98.46]
Kraus 2014	1	27	71	2	14	69	51.5%	-1.00 [-8.10, 6.10]
Total (95% CI)			**87**			**78**	**100.0%**	**34.83 [-37.46, 107.12]**

Heterogeneity: Tau² = 2630.96; Chi² = 29.52, df = 1 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.94 (P = 0.35)

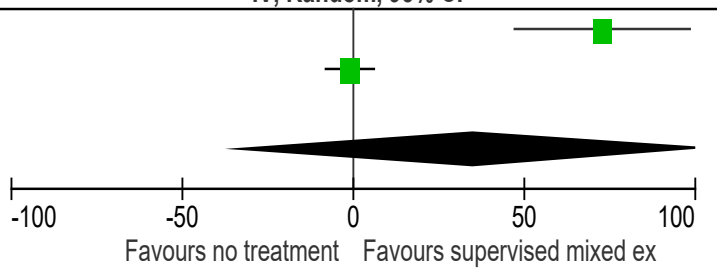


Figure 223: Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Aglamis 2008	96.9	7.2	16	38.6	35.9	9	47.9%	58.30 [34.58, 82.02]
Kraus 2014	-2	13	71	-2	15	69	52.1%	0.00 [-4.66, 4.66]
Total (95% CI)			**87**			**78**	**100.0%**	**27.94 [-29.14, 85.03]**

Heterogeneity: Tau² = 1623.40; Chi² = 22.35, df = 1 (P < 0.00001); I² = 96%

Test for overall effect: Z = 0.96 (P = 0.34)

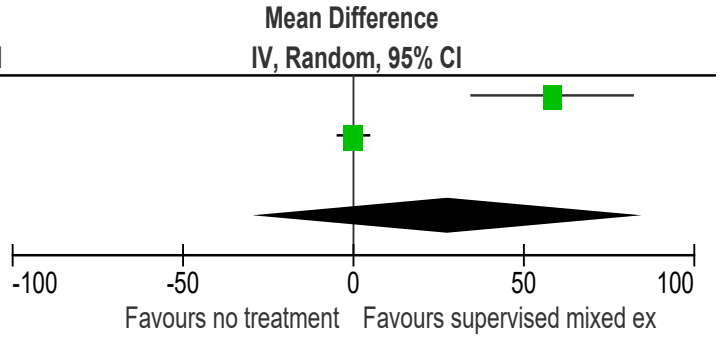


Figure 224: Quality of life (AIMS2 arm function, 0-10, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Peloquin 1999	0.26	0.61	59	0.39	1.1	65	-0.13 [-0.44, 0.18]

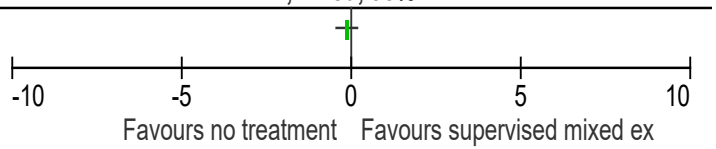


Figure 225: Quality of life (AIMS2 arthritis pain, 0-10, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Peloquin 1999	3.09	1.54	59	3.94	2.22	65	-0.85 [-1.52, -0.18]

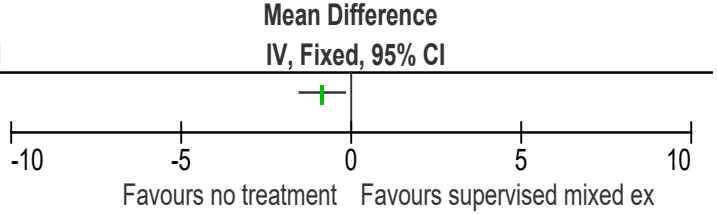


Figure 226: Quality of life (AIMS2 hand and finger function, 0-10, high is good, final value) at ≤3 months

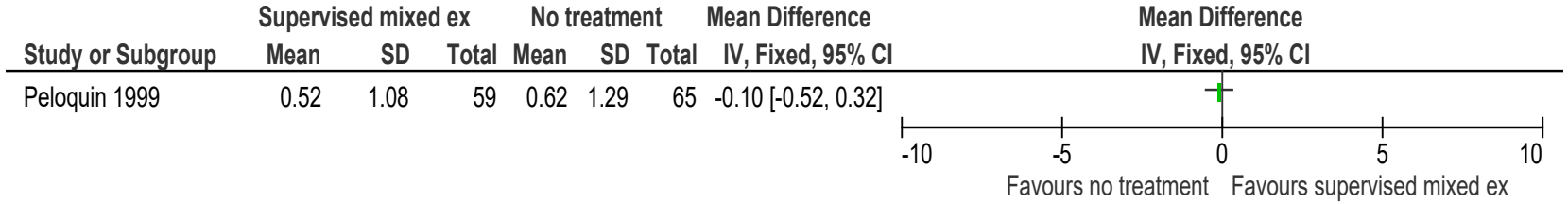


Figure 227: Quality of life (AIMS2 household tasks, 0-10, high is good, final value) at ≤3 months

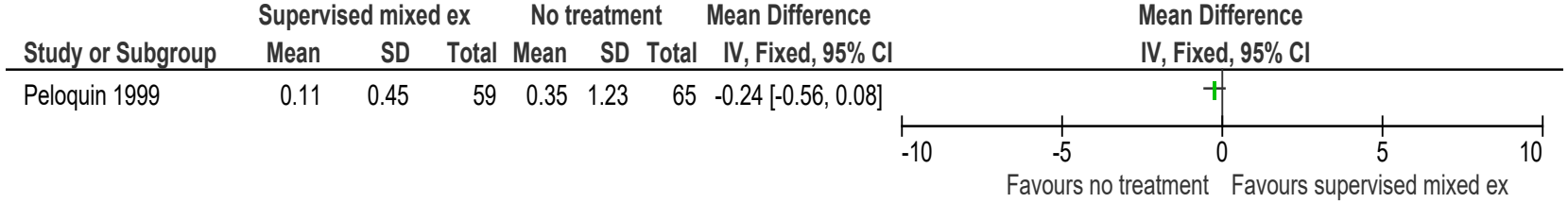


Figure 228: Quality of life (AIMS2 level of tension, 0-10, high is good, final value) at ≤3 months

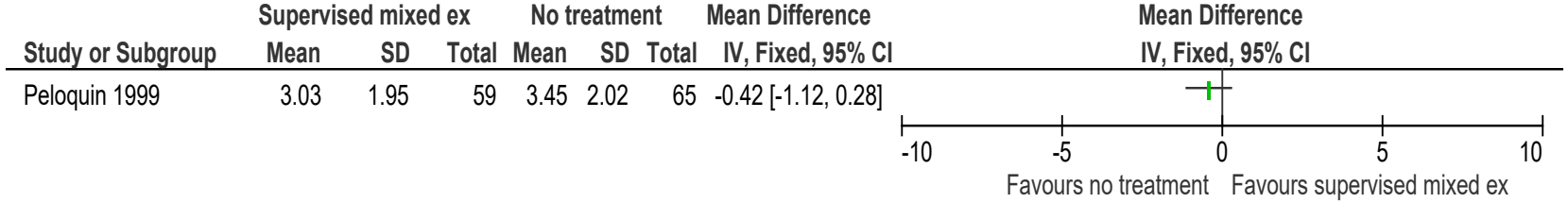


Figure 229: Quality of life (AIMS2 mobility level, 0-10, high is good, final value) at ≤3 months

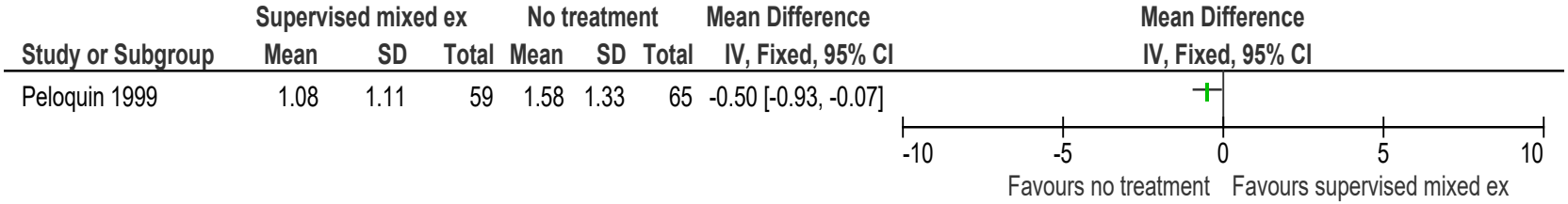


Figure 230: Quality of life (AIMS2 mood, 0-10, high is good, final value) at ≤3 months

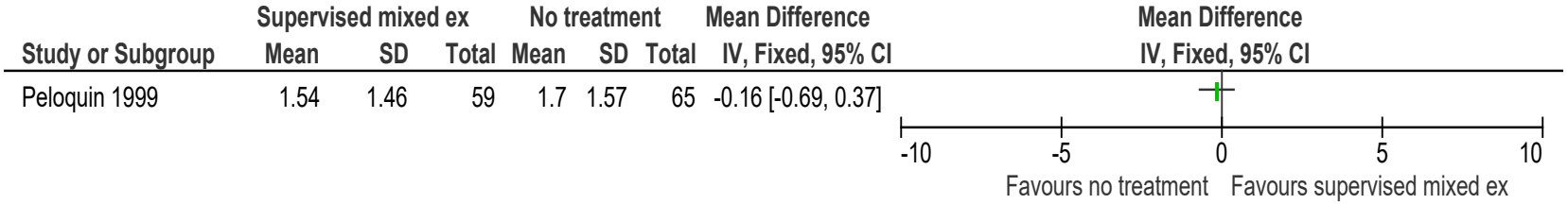


Figure 231: Quality of life (AIMS2 self-care tasks, 0-10, high is good, final value) at ≤3 months

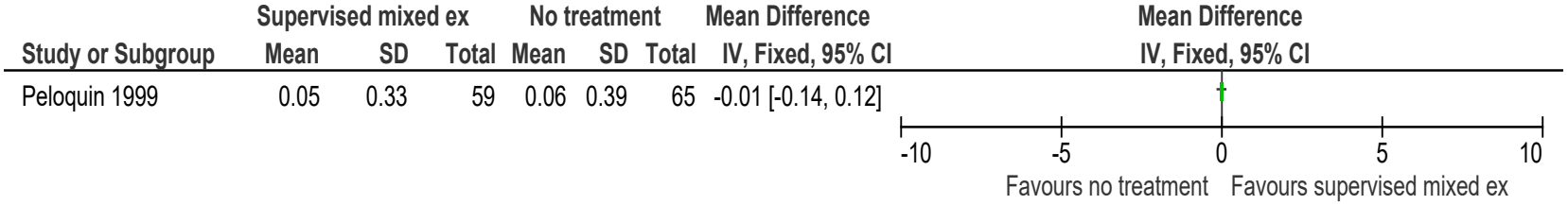


Figure 232: Quality of life (AIMS2 social activity, 0-10, high is good, final value) at ≤3 months

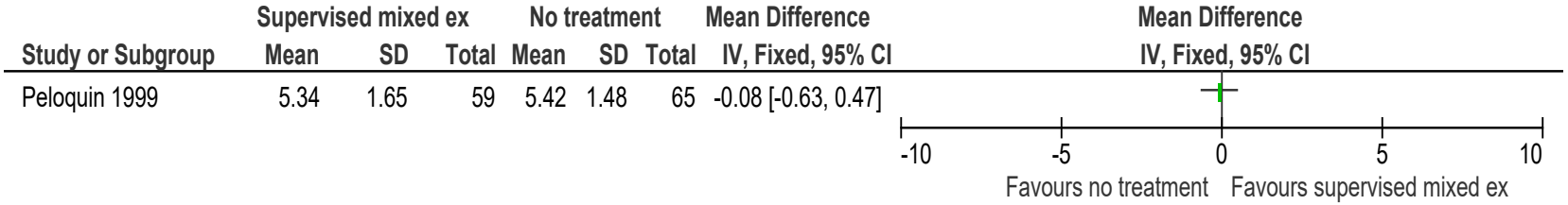


Figure 233: Quality of life (AIMS2 support from family and friends, 0-10, high is good, final value) at ≤3 months

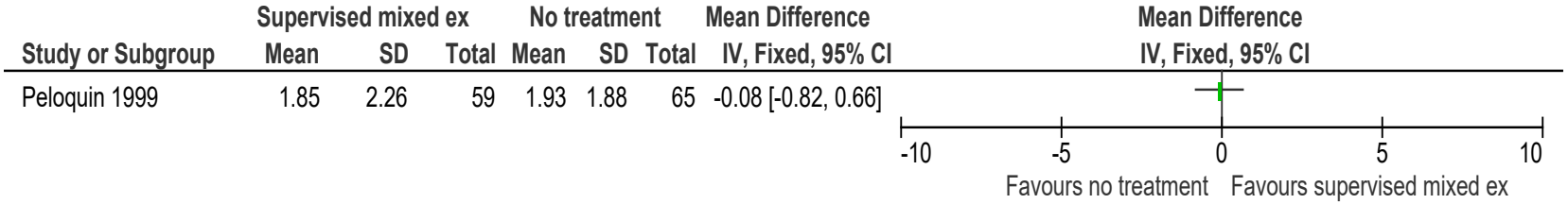


Figure 234: Quality of life (AIMS2 walking and bending, 0-10, high is good, final value) at ≤3 months

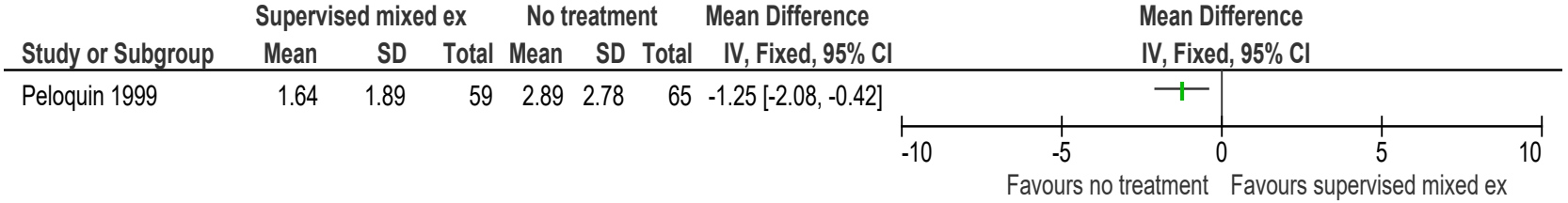


Figure 235: Quality of life (AIMS2 work, 0-10, high is good, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Peloquin 1999	0.89	1.13	59	1.28	1.64	65	-0.39 [-0.88, 0.10]

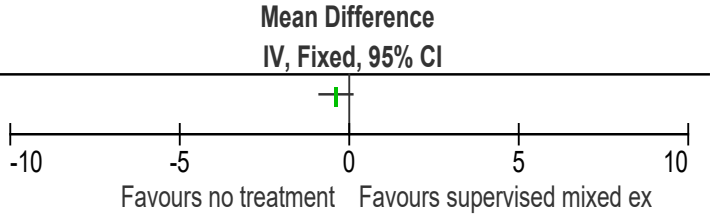


Figure 236: Pain (WOMAC, VAS, 0-100, high is poor, change scores) at ≤3 months

Study or Subgroup	Mean Difference	SE	Supervised mixed ex Total	No treatment Total	Weight	Mean Difference IV, Random, 95% CI
Kraus 2014	-7.2	2.4726	71	69	52.7%	-7.20 [-12.05, -2.35]
Van baar 2001	-17	3.3674	93	98	47.3%	-17.00 [-23.60, -10.40]
Total (95% CI)			**164**	**167**	**100.0%**	**-11.83 [-21.42, -2.24]**

Heterogeneity: Tau² = 39.29; Chi² = 5.50, df = 1 (P = 0.02); I² = 82%
Test for overall effect: Z = 2.42 (P = 0.02)

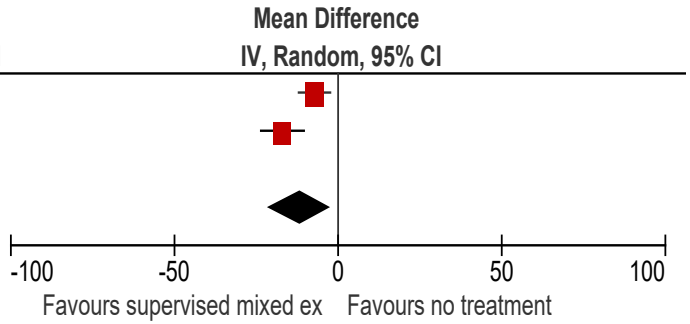


Figure 237: Pain (KOOS, WOMAC, AIMS, VAS, NRS [different scale ranges], high is poor, final values) at ≤3 months

	Supervised mixed ex			No treatment				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Aglamis 2008	0.7	1	16	7.7	2.3	9	4.3%	-4.29 [-5.82, -2.76]
Cheung 2017	6.5	2.6	28	6.5	2.7	23	11.5%	0.00 [-0.55, 0.55]
De matos brunelli braghin 2018	8	10.99	15	22.33	23.59	15	9.5%	-0.76 [-1.50, -0.01]
De rooij 2017	8.4	3	60	9.1	3.6	55	13.3%	-0.21 [-0.58, 0.16]
French 2013	4.02	2.88	45	5.62	2.84	43	12.8%	-0.55 [-0.98, -0.13]
Keefe 2004	3.19	1.85	16	4.03	2.08	18	10.2%	-0.42 [-1.10, 0.27]
Pazit 2018	97.3	127.1	10	249.7	309.3	10	8.1%	-0.62 [-1.52, 0.28]
Rogind 1998	4	2.2	12	6	2.3	13	8.8%	-0.86 [-1.68, -0.03]
Takacs 2017	2.8	1.7	17	4.6	2.3	19	10.1%	-0.86 [-1.55, -0.18]
Wang 2011	-76	15	26	-68	18	26	11.5%	-0.48 [-1.03, 0.08]
Total (95% CI)			**245**			**231**	**100.0%**	**-0.67 [-1.04, -0.29]**

Heterogeneity: Tau² = 0.24; Chi² = 31.50, df = 9 (P = 0.0002); I² = 71%

Test for overall effect: Z = 3.47 (P = 0.0005)

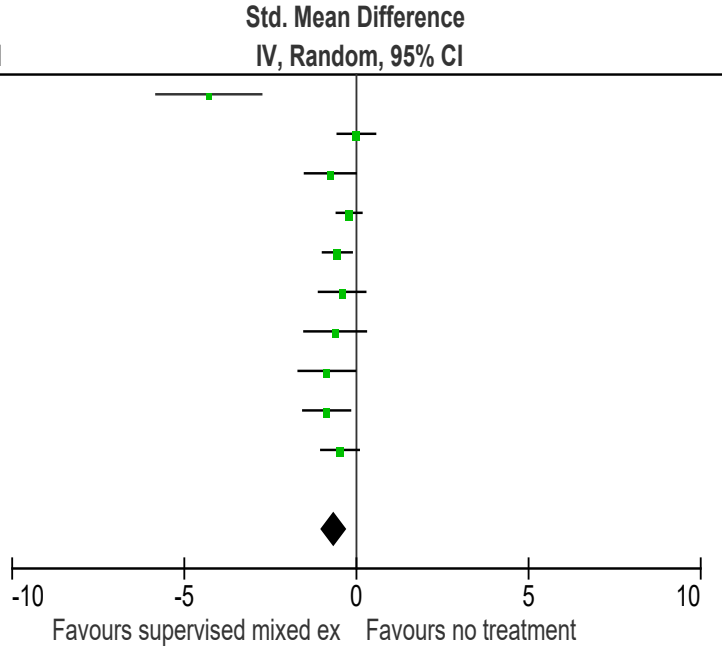


Figure 238: Pain (VAS, 0-100, high is poor, change scores) at >3 months

Study or Subgroup	Mean Difference	SE	Supervised mixed ex Total	No treatment Total	Weight	Mean Difference IV, Fixed, 95% CI
Abbott 2013	-9	4.8532	51	51	42.0%	-9.00 [-18.51, 0.51]
Van baar 2001	-6.6	4.1327	90	92	58.0%	-6.60 [-14.70, 1.50]
Total (95% CI)			**141**	**143**	**100.0%**	**-7.61 [-13.78, -1.44]**

Heterogeneity: Chi² = 0.14, df = 1 (P = 0.71); I² = 0%
Test for overall effect: Z = 2.42 (P = 0.02)

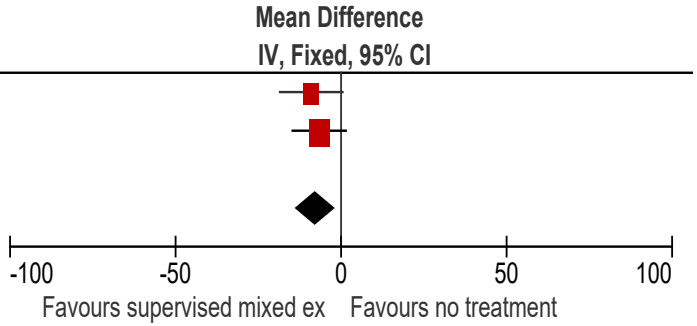


Figure 239: Pain (KOOS, NRS [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
De rooij 2017	6.6	3.6	51	8.6	3.6	56	82.4%	-0.55 [-0.94, -0.16]
Rogind 1998	4	2.7	12	7	3.2	13	17.6%	-0.98 [-1.81, -0.14]
Total (95% CI)			**63**			**69**	**100.0%**	**-0.63 [-0.98, -0.27]**

Heterogeneity: Chi² = 0.81, df = 1 (P = 0.37); I² = 0%
Test for overall effect: Z = 3.49 (P = 0.0005)

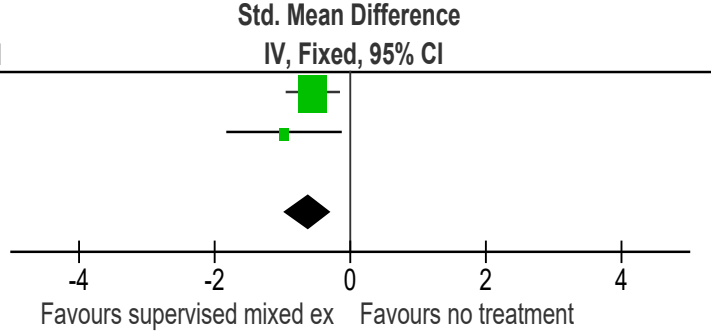


Figure 240: Physical function (WOMAC, 0-100, high is poor, change score) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Kraus 2014	-8.4	13.4	71	-2.1	12.9	68	-6.30 [-10.67, -1.93]

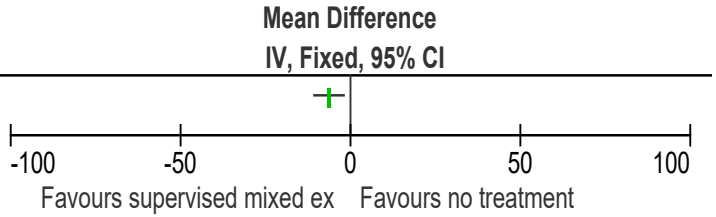


Figure 241: Physical function (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Cheung 2017	25.8	8.4	28	25.2	8.4	23	13.4%	0.07 [-0.48, 0.62]
De matos brunelli braghin 2018	6.57	8.28	15	15.2	21.73	15	7.7%	-0.51 [-1.24, 0.22]
De rooij 2017	30.4	11.6	60	32.9	11.2	55	30.3%	-0.22 [-0.58, 0.15]
French 2013	28.08	15.48	45	36.09	16.41	43	22.6%	-0.50 [-0.92, -0.07]
Pazit 2018	277.8	237	10	565.7	282.5	10	4.5%	-1.06 [-2.01, -0.11]
Takacs 2017	20	11	17	28	10	19	8.8%	-0.75 [-1.43, -0.07]
Wang 2011	-82	14	26	-69	18	26	12.7%	-0.79 [-1.36, -0.23]
Total (95% CI)			**201**			**191**	**100.0%**	**-0.42 [-0.62, -0.22]**

Heterogeneity: Chi² = 8.68, df = 6 (P = 0.19); I² = 31%
Test for overall effect: Z = 4.10 (P < 0.0001)

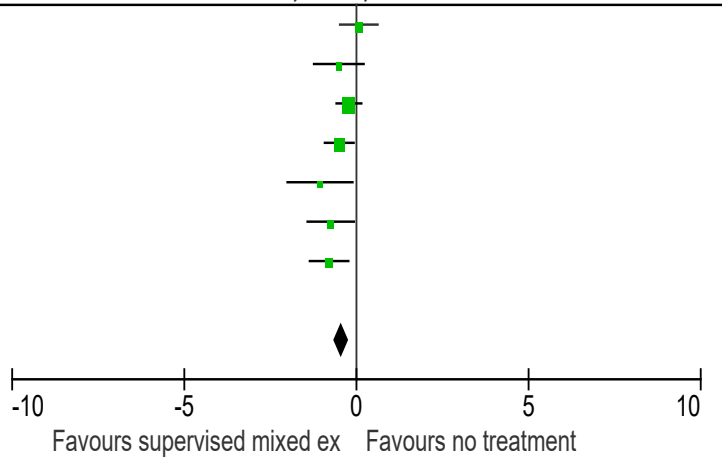


Figure 242: Physical function (WOMAC, 0-68, high is poor, final value) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
De rooij 2017	23.5	13.1	51	31.4	12.6	56	-7.90 [-12.78, -3.02]

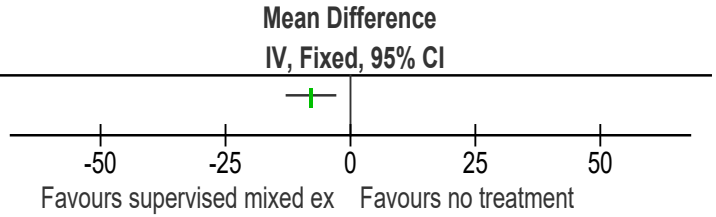


Figure 243: Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Cheung 2017	5.2	2.7	28	4.4	2.8	23	56.2%	0.80 [-0.72, 2.32]
French 2013	6.74	4.27	45	6.14	3.97	43	43.8%	0.60 [-1.12, 2.32]
Total (95% CI)			**73**			**66**	**100.0%**	**0.71 [-0.43, 1.85]**

Heterogeneity: Chi² = 0.03, df = 1 (P = 0.86); I² = 0%
Test for overall effect: Z = 1.23 (P = 0.22)

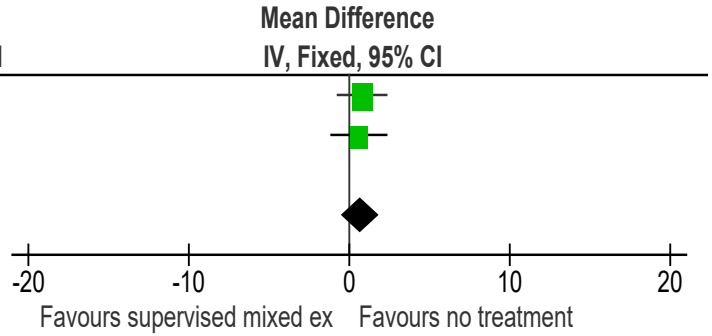


Figure 244: Psychological distress (HADS depression, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Cheung 2017	4.2	2	28	3.7	2.1	23	61.4%	0.50 [-0.63, 1.63]
French 2013	5.02	3.39	45	5.58	3.45	43	38.6%	-0.56 [-1.99, 0.87]
Total (95% CI)			**73**			**66**	**100.0%**	**0.09 [-0.80, 0.98]**

Heterogeneity: Chi² = 1.30, df = 1 (P = 0.25); I² = 23%

Test for overall effect: Z = 0.20 (P = 0.84)

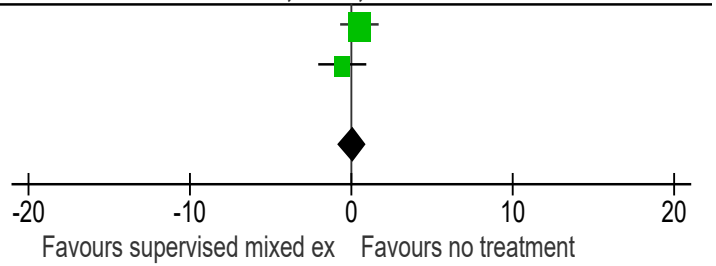


Figure 245: Psychological distress (AIMS psychological disability, 0-10, high is poor, final value) at ≤3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Keefe 2004	1.88	0.87	16	1.8	1.04	18	0.08 [-0.56, 0.72]

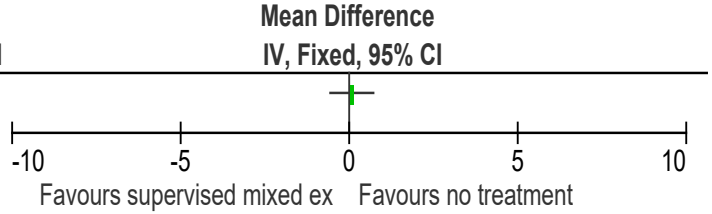


Figure 246: Psychological distress (HADS, 0-21, high is poor, final value) at >3 months

Study or Subgroup	Supervised mixed ex Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
De rooij 2017	9.6	6.5	51	8	6.7	55	1.60 [-0.91, 4.11]

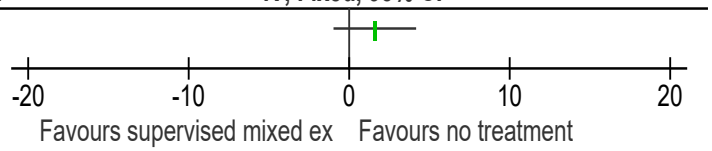


Figure 247: Serious adverse events at ≤3 months

Study or Subgroup	Supervised mixed ex Events	Total	No treatment Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Kraus 2014	0	71	0	69	49.3%	0.00 [-0.03, 0.03]
Pazit 2018	0	10	0	10	7.1%	0.00 [-0.17, 0.17]
Peloquin 1999	1	59	0	65	43.6%	0.02 [-0.03, 0.06]
Total (95% CI)		**140**		**144**	**100.0%**	**0.01 [-0.02, 0.04]**
Total events	1		0			

Heterogeneity: Chi² = 0.46, df = 2 (P = 0.79); I² = 0%

Test for overall effect: Z = 0.52 (P = 0.60)

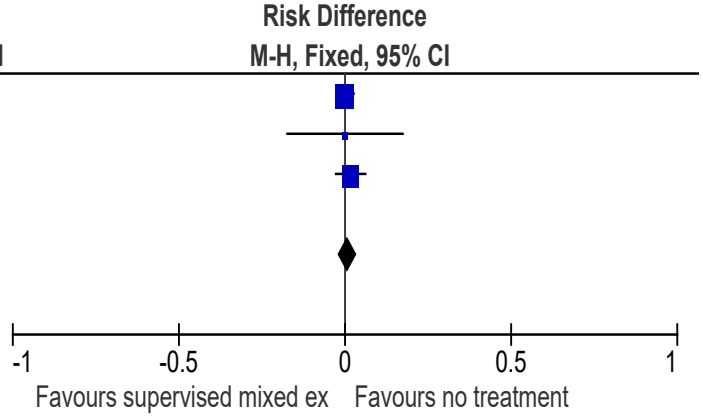
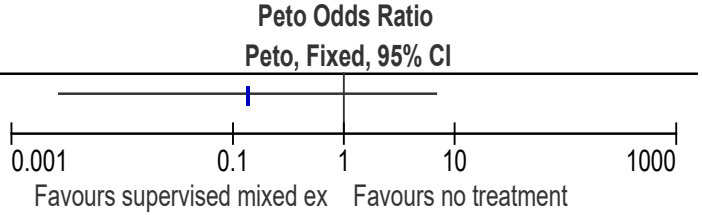


Figure 248: Serious adverse events at >3 months

Study or Subgroup	Supervised mixed ex Events	Total	No treatment Events	Total	Peto Odds Ratio Peto, Fixed, 95% CI
Abbott 2013	0	51	1	51	0.14 [0.00, 6.82]



E.21 Unsupervised mixed modality exercise compared to unsupervised strength exercise

Figure 249: Pain (WOMAC, 0-20, high is poor, change score) at ≤3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Chen 2021	-3	1.67	16	-1.88	1.03	16	-1.12 [-2.08, -0.16]

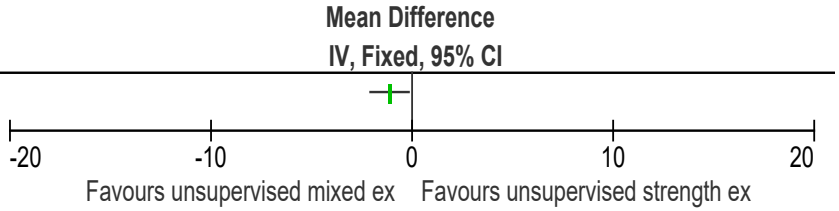


Figure 250: Pain (VAS, NRS, 0-10, high is poor, final values) at ≤3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Fitzgerald 2011	3.5	2.4	75	4.1	2.6	84	52.1%	-0.60 [-1.38, 0.18]
Gondhalekar 2013	4.07	1.18	15	3.53	1.33	15	47.9%	0.54 [-0.36, 1.44]
Total (95% CI)			**90**			**99**	**100.0%**	**-0.05 [-1.17, 1.06]**

Heterogeneity: Tau² = 0.47; Chi² = 3.53, df = 1 (P = 0.06); I² = 72%

Test for overall effect: Z = 0.09 (P = 0.93)

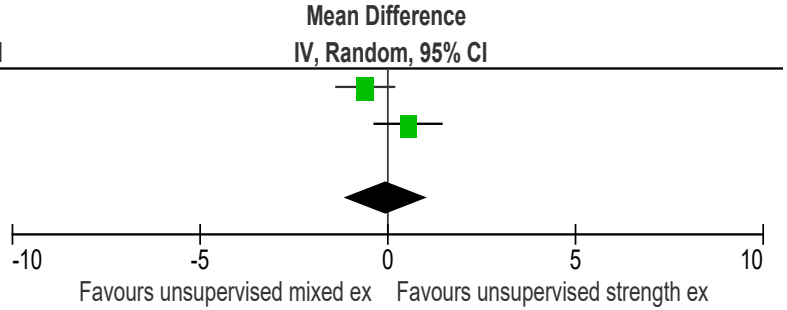


Figure 251: Pain (NRS, 0-10, high is poor, final value) at >3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Fitzgerald 2011	3.6	2.7	76	3.5	3.1	66	0.10 [-0.86, 1.06]

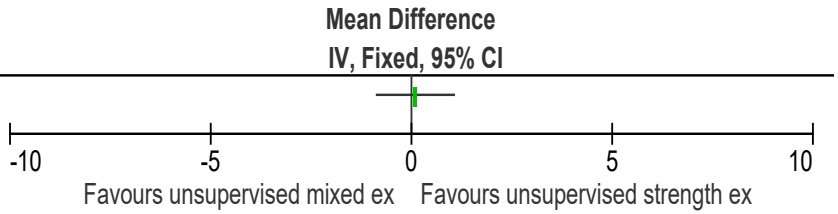
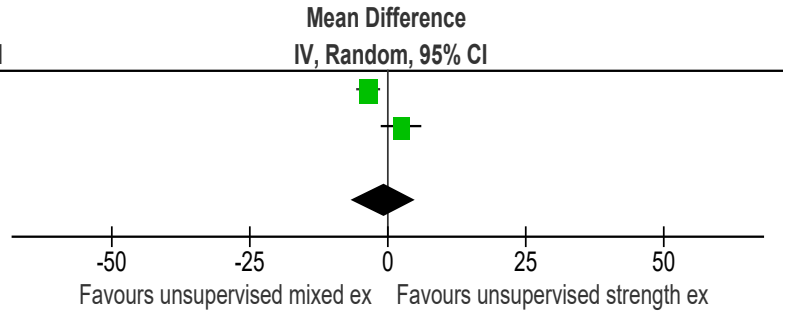


Figure 252: Physical function (WOMAC, 0-68, high is poor, change score and final value) at ≤3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Unsupervised strength ex Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Chen 2021	-6.44	3.69	16	-2.88	1.78	16	53.1%	-3.56 [-5.57, -1.55]
Fitzgerald 2011	15.2	11.5	84	12.8	11.1	75	46.9%	2.40 [-1.12, 5.92]
Total (95% CI)			**100**			**91**	**100.0%**	**-0.76 [-6.59, 5.07]**

Heterogeneity: Tau² = 15.63; Chi² = 8.33, df = 1 (P = 0.004); I² = 88%

Test for overall effect: Z = 0.26 (P = 0.80)



E.22 Unsupervised mixed modality exercise compared to other unsupervised exercise

Figure 253: Pain (WOMAC, 0-100, high is poor, change score) at ≤3 months

	Unsupervised mixed ex			Other unsupervised ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Petrella 2000	18	9	91	11	7	88	7.00 [4.64, 9.36]

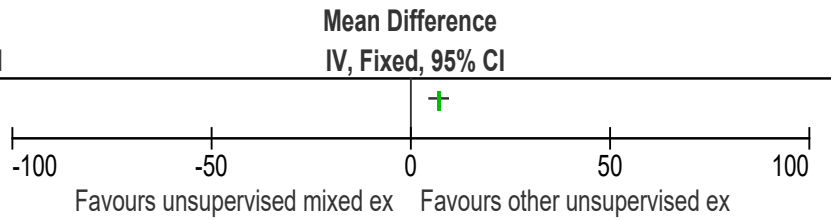


Figure 254: Physical function (WOMAC, 0-100, high is poor, change score) at ≤3 months

	Unsupervised mixed ex			Other unsupervised ex			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Petrella 2000	14	6	91	5	3	88	9.00 [7.62, 10.38]

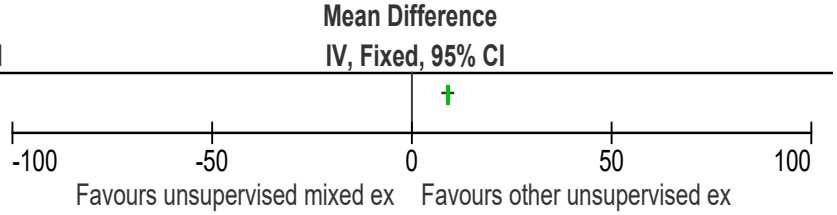
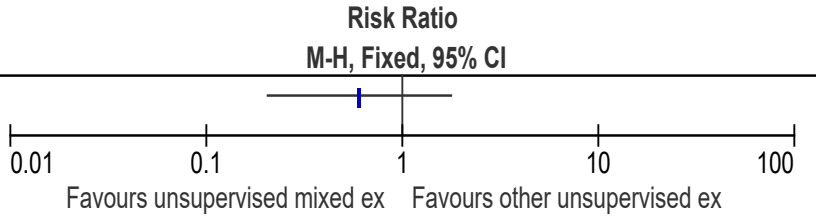


Figure 255: Serious adverse events at ≤3 months

	Unsupervised mixed ex		Other unsupervised ex		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	M-H, Fixed, 95% CI
Petrella 2000	5	91	8	88	0.60 [0.21, 1.78]



E.23 Unsupervised mixed modality exercise compared to pharmacological treatment

Figure 256: Pain (HSS pain during activity, VAS [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Karatosun 2006	-12.1	3.1	53	-12.9	3.4	52	78.0%	0.24 [-0.14, 0.63]
Karatosun 2008	2.4	3.1	15	1.4	1.9	15	22.0%	0.38 [-0.34, 1.10]
Total (95% CI)			**68**			**67**	**100.0%**	**0.27 [-0.07, 0.61]**

Heterogeneity: Chi² = 0.10, df = 1 (P = 0.75); I² = 0%
Test for overall effect: Z = 1.58 (P = 0.11)

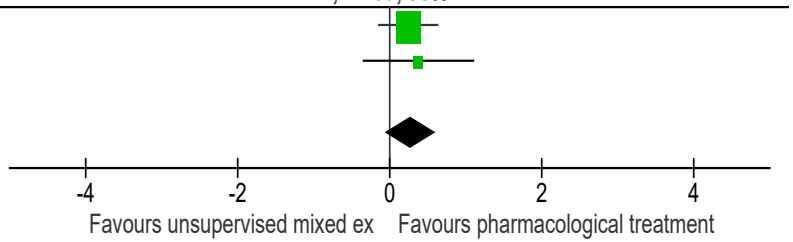


Figure 257: Pain (VAS, 0-100, high is poor, change score) at >3 months

Study or Subgroup	Unsupervised mixed ex Mean	SD	Total	Pharmacological treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Kawasaki 2009	-21.29	27.6	60	-20.46	36.04	60	-0.83 [-12.32, 10.66]

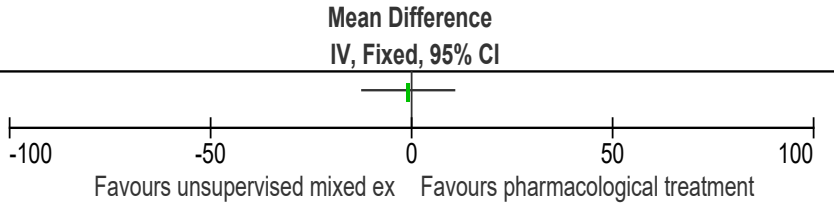
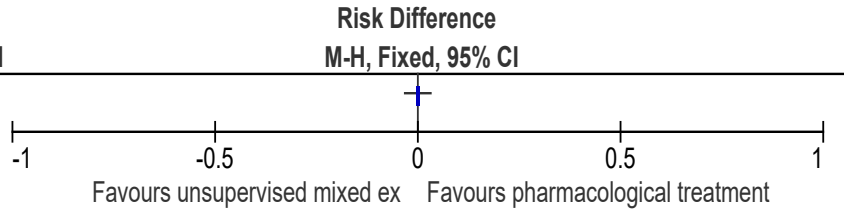


Figure 258: Serious adverse events at >3 months

Study or Subgroup	Unsupervised mixed ex Events	Total	Pharmacological treatment Events	Total	Risk Difference M-H, Fixed, 95% CI
Kawasaki 2009	0	60	0	60	0.00 [-0.03, 0.03]



E.24 Unsupervised mixed modality exercise compared to no treatment

Figure 259: Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at ≤3 months

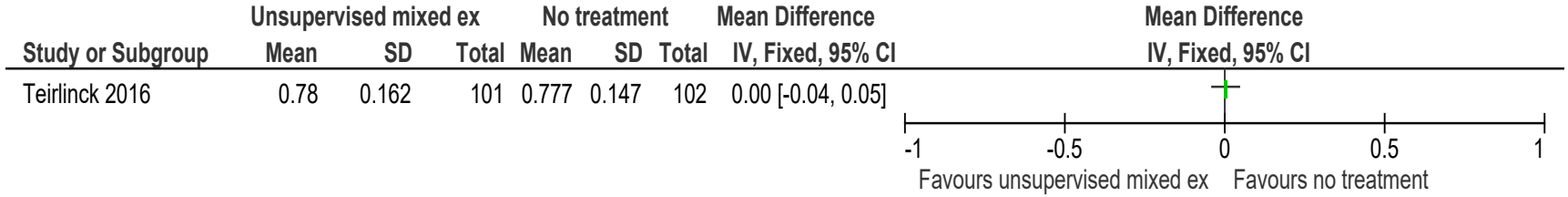


Figure 260: Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at >3 months

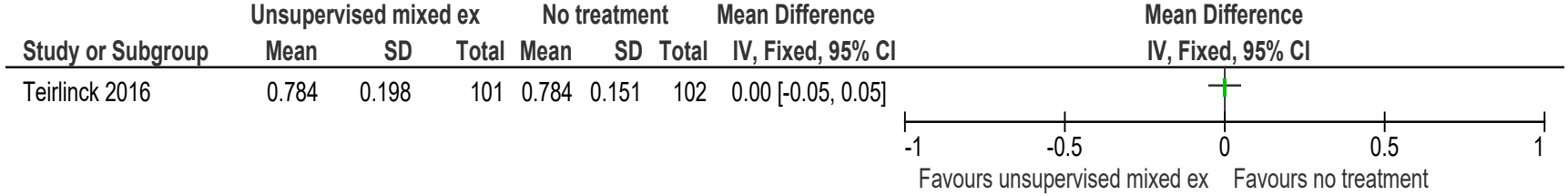


Figure 261: Pain (HOOS, 0-100, high is poor, final value) at ≤3 months

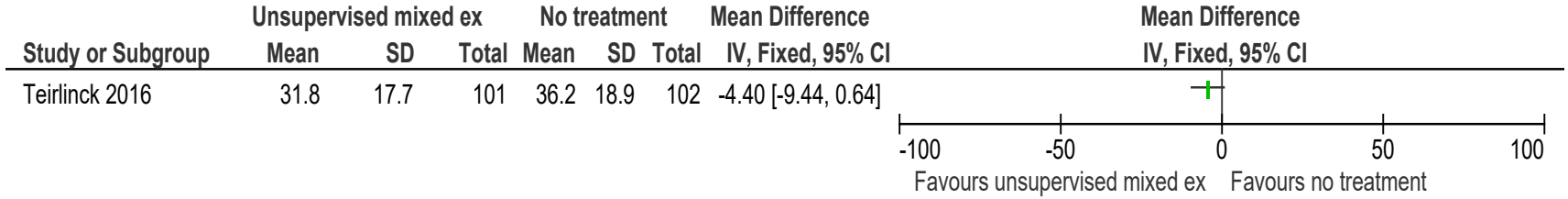


Figure 262: Pain (WOMAC, 0-20, high is poor, change score) at >3 months

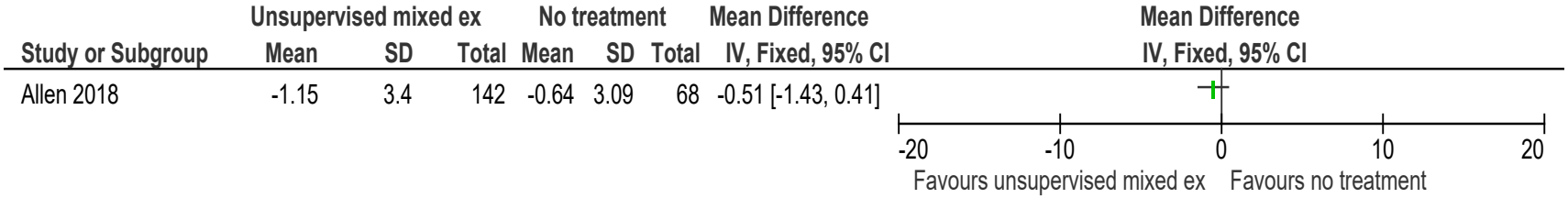


Figure 263: Pain (HOOS, 0-100, high is poor, final value) at >3 months

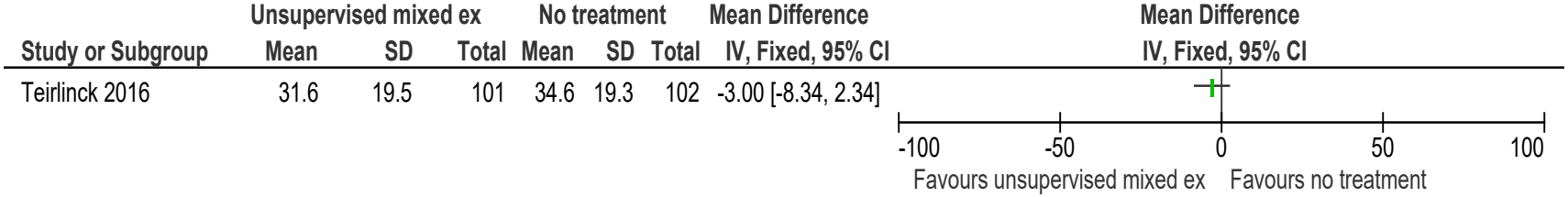


Figure 264: Physical function (HOOS, 0-100, high is poor, final value) at ≤3 months

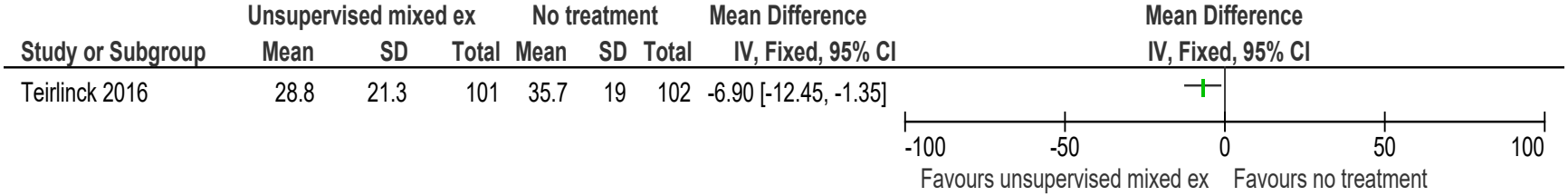


Figure 265: Physical function (WOMAC, 0-68, high is poor, change score) at >3 months

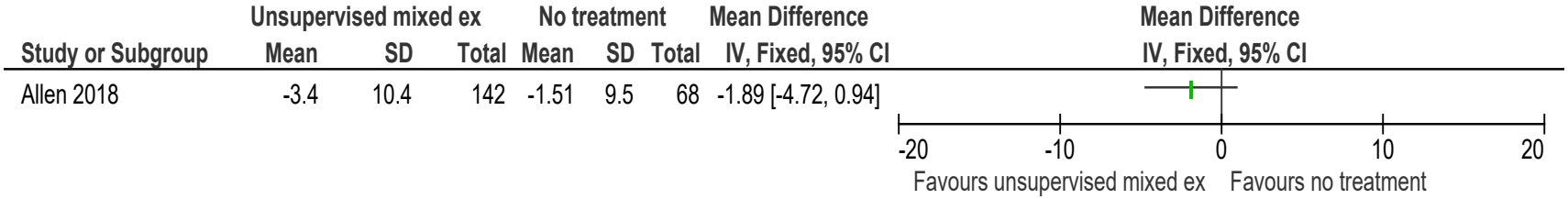


Figure 266: Physical function (HOOS, 0-100, high is poor, final value) at >3 months

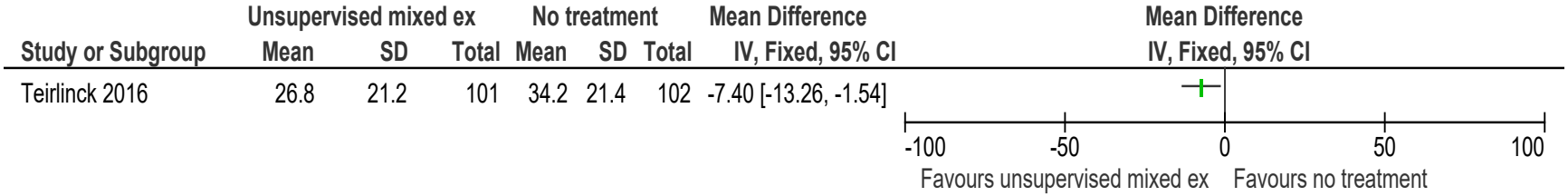
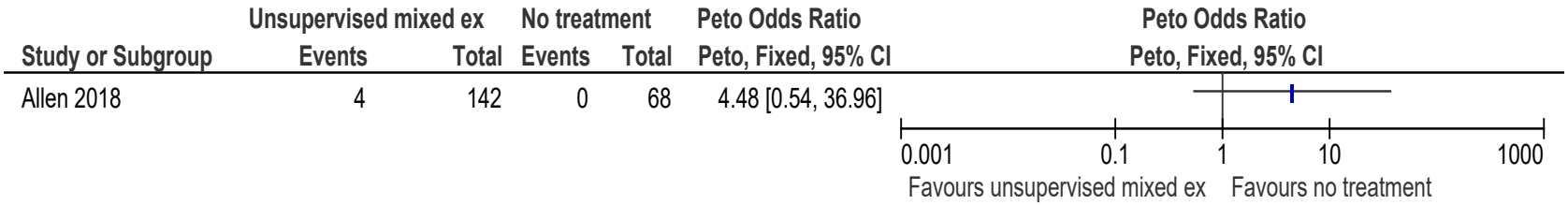


Figure 267: Serious adverse events at >3 months



Appendix F – GRADE tables

F.1 Supervised strength exercise compared to unsupervised strength exercise, supervised aerobic exercise and no treatment

Table 58: Clinical evidence profile: supervised strength exercise compared to unsupervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS, 0-100, high is poor, final values) at <3 months (follow-up: 7 weeks; assessed with: VAS; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	58	57	-	MD **19.77 lower** (22.32 lower to 17.23 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Pain (VAS, 0-10, high is poor, final value) at >3 months (follow-up: 6 months; assessed with: VAS; Scale from: 0 to 10)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	18	18	-	MD **2.3 lower** (2.47 lower to 2.13 lower)	⨁⨁⨁◯ MODERATE	CRITICAL

CI: confidence interval; **MD:** mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Table 59: Clinical evidence profile: supervised strength exercise compared to supervised aerobic exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	supervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 7 weeks; assessed with: WOMAC)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	33	33	-	SMD **0.45 SD higher** (0.04 lower to 0.94 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (Arthritis Self-Efficacy pain subscale, 0-100, high is poor) at >3 months (follow-up: 12 months; assessed with: Arthritis Self-Efficacy pain subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	50	50	-	MD **11.1 higher** (0.1 higher to 22.1 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (Intermittent and constant osteoarthritis pain total pain, 0-20, high is poor, final value) at >3 months (follow-up: 18 weeks; assessed with: Intermittent and constant osteoarthritis pain total pain,; Scale from: 0 to 20)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	14	16	-	MD **2.17 lower** (8.41 lower to 4.07 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at <3 months (follow-up: 6 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	15	14	-	MD **1.51 higher** (6.88 lower to 9.9 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (Arthritis Self-Efficacy function subscale, 0-100, high is poor) at >3 months (follow-up: 12 months; assessed with: Arthritis Self-Efficacy function subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	50	50	-	MD **7.6 higher** (0.7 higher to 14.5 higher)	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 60: Clinical evidence profile: supervised strength exercise compared to pharmacological treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	pharmacological treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 total, scale range unclear, high is good, final values) at <3 months (follow-up: 12 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	92	74	-	MD **22 higher** (17.5 higher to 26.5 higher)	⨁⨁◯◯ LOW	CRITICAL

CI: confidence interval; **MD:** mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

Table 61: Clinical evidence profile: supervised strength exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, change scores) at <3 months (follow-up: mean 10 weeks; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	50	48	-	MD **15.94 higher** (4.44 lower to 36.32 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (EQ-5D, KOOS, HOOS, Assessment of Quality of Life Scale, AIMS [different scale ranges], high is good, final values) at <3 months (follow-up: mean 10 weeks; assessed with: EQ-5D, KOOS, HOOS, Assessment of Quality of Life Scale, AIMS)												
6	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	285	284	-	SMD **0.42 SD higher** (0.01 lower to 0.86 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical component summary, SF-12 physical score, 0-100, high is good, final values) at <3 months (follow-up: mean 11 weeks; assessed with: SF-36 physical component summary, SF-12 physical score; Scale from: 0 to 100)												
4	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	122	96	-	MD **5.33 higher** (8.19 lower to 18.85 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental component summary, SF-12 mental score, 0-100, high is good, final values) at <3 months (follow-up: mean 11 weeks; assessed with: SF-36 mental component summary, SF-12 mental score; Scale from: 0 to 100)												
4	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	122	96	-	MD **9.45 higher** (0.79 higher to 18.11 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at <3 months (follow-up: 10 weeks; assessed with: SF-36 physical function; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	79	81	-	MD **16.35 higher** (9.1 higher to 23.61 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	79	81	-	MD **14.47 higher** (5.21 higher to 23.73 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	79	81	-	MD **26.19 higher** (11.79 higher to 40.58 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	79	81	-	MD **9.83 higher** (0.44 higher to 19.22 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	serious[c]	not serious	very serious[b]	none	79	81	-	MD **7.57 higher** (3.53 lower to 18.67 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	serious[c]	not serious	not serious	none	79	81	-	MD **10.12 higher** (3.98 lower to 24.22 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	79	81	-	MD **16.9 higher** (0.14 higher to 33.67 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	79	81	-	MD **15.4 higher** (4.24 higher to 26.56 higher)	⨁⨁◯◯ Low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (EQ-5D, KOOS [different scale ranges], high is good, final values) at >3 months (follow-up: mean 15 months; assessed with: EQ-5D, KOOS)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	204	203	-	SMD **0.06 lower** (0.25 lower to 0.14 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 physical component summary, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 physical component summary; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	30	31	-	MD **3.7 higher** (12.19 lower to 19.59 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental component summary, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 mental component summary; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	30	31	-	MD **1.4 higher** (19.26 lower to 22.06 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (KOOS, WOMAC, NRS, VAS [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 8 weeks; assessed with: KOOS, WOMAC, NRS, VAS)												
7	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	266	265	-	SMD **0.55 SD lower** (0.73 lower to 0.37 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (KOOS, HOOS, AUSCAN, WOMAC, NRS, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, HOOS, AUSCAN, WOMAC, NRS, VAS)												
23	randomised trials	very serious[a]	serious[c]	not serious	not serious	none	970	763	-	SMD **0.81 SD lower** (1.06 lower to 0.57 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (KOOS, VAS [different scale ranges], high is poor, change score and final values) at >3 months (follow-up: mean 11 months; assessed with: KOOS, VAS)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
7	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	462	380	-	SMD **1 SD lower** (1.76 lower to 0.23 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, WOMAC [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, WOMAC)												
5	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	151	150	-	SMD **0.58 SD lower** (0.83 lower to 0.33 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, HOOS, AUSCAN, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, HOOS, AUSCAN, WOMAC, Modified Bandi's criteria of functional incapacity)												
19	randomised trials	very serious[a]	very serious[c]	not serious	not serious	none	761	620	-	SMD **1 SD lower** (1.37 lower to 0.63 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS ADL, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: KOOS ADL; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	30	31	-	MD **2.8 higher** (4.36 lower to 9.96 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, WOMAC, Modified Bandi's criteria of functional incapacity [different scale ranges], high is poor, final scores) at >3 months (follow-up: mean 10 months; assessed with: KOOS, WOMAC, Modified Bandi's criteria of functional incapacity)												
3	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	268	251	-	SMD **0.31 lower** (1.09 lower to 0.48 higher)	⨁◯◯◯ Very low	CRITICAL
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: HADS anxiety; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	60	61	-	MD **0.54 higher** (1.1 lower to 2.18 higher)	⨁⨁◯◯ Low	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Psychological distress (HADS depression, 0-21, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: HADS depression; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	60	61	-	MD **0.38 lower** (1.7 lower to 0.94 higher)	⨁⨁◯◯ Low	IMPORTANT
Serious adverse events at <3 months												
3	randomised trials	serious[a]	serious[d]	not serious	very serious[e]	none	5/89 (5.6%)	0/91 (0.0%)	**RD 0.06** (0.00 to 0.12)	**60 more per 1,000** (from 0 fewer to 120 more)[f]	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

e. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

f. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

F.2 Unsupervised strength exercise compared to unsupervised aerobic exercise and no treatment

Table 62: Clinical evidence profile: unsupervised strength exercise compared to unsupervised aerobic exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	unsupervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile physical mobility subscale; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	27	28	-	MD **20.9 higher** (18.56 higher to 23.24 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile pain subscale; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	27	28	-	MD **0.8 higher** (0.89 lower to 2.49 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile energy subscale; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	27	28	-	MD **18.8 higher** (17.87 higher to 19.73 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile sleep subscale; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	27	28	-	MD **12.3 higher** (9.93 higher to 14.67 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile emotional reactions subscale; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	unsupervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	27	28	-	MD **12.2 higher** (10.81 higher to 13.59 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: Nottingham Health Profile social isolation subscale; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	27	28	-	MD **0.2 lower** (2.32 lower to 1.92 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	27	28	-	MD **0.4 lower** (1.2 lower to 0.4 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	27	28	-	MD **0.6 higher** (0.52 lower to 1.72 higher)	⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 63: Clinical evidence profile: unsupervised strength exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (EQ-5D, Arthritis Impact Measurement Scale 2 - Short form [different scale ranges], high is good, final values) at <3 months (follow-up: mean 12 weeks; assessed with: EQ-5D, Arthritis Impact Measurement Scale 2 - Short form)												
2	randomised trials	very serious[a]	serious[b]	not serious	serious[c]	none	136	135	-	SMD **0.2 SD higher** (0.23 lower to 0.63 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile physical mobility subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	27	26	-	MD **7.1 lower** (10.06 lower to 4.14 lower)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile pain subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	27	26	-	MD **10.6 lower** (12.3 lower to 8.9 lower)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile energy subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	27	26	-	MD **15.9 lower** (16.93 lower to 14.87 lower)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile sleep subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	27	26	-	MD **3.4 lower** (5.91 lower to 0.89 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile emotional reactions subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	27	26	-	MD **8.8 lower** (10.72 lower to 6.88 lower)	⨁⨁◯◯ LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile social isolation subscale; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	27	26	-	MD **2.1 lower** (4.48 lower to 0.28 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (EQ-5D, 0-1, high is good, final value) at >3 months (follow-up: 12 months; assessed with: EQ-5D; Scale from: 0 to 1)												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	65	65	-	MD **0.07 higher** (0 to 0.14 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 physical functioning, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 physical functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **4.31 higher** (0.41 lower to 9.03 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **4.81 higher** (2.3 lower to 11.92 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 role physical; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **10.78 higher** (0.54 lower to 22.1 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 vitality; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	113	78	-	MD **1.91 higher** (3.53 lower to 7.35 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 general health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **2.63 higher** (1.55 lower to 6.81 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 mental health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **2.7 higher** (1.8 lower to 7.2 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	113	78	-	MD **1.37 higher** (14.87 lower to 17.61 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, change score) at >3 months (follow-up: 6 months; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	113	78	-	MD **0.01 lower** (10.3 lower to 10.28 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC, NRS)												
5	randomised trials	very serious[a]	not serious	not serious	not serious	none	192	187	-	SMD **1.1 SD lower** (1.32 lower to 0.88 lower)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, NRS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC, NRS)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
4	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	223	219	-	SMD **0.37 SD lower** (0.81 lower to 0.08 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at >3 months (follow-up: mean 6 months; assessed with: WOMAC, VAS)												
2	randomised trials	very serious[a]	very serious[b]	not serious	not serious	none	848	838	-	SMD **0.08 lower** (0.18 lower to 0.01 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, NRS, 0-100, high is poor, final values) at >3 months (follow-up: 15 months; assessed with: WOMAC, NRS; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	125	123	-	MD **1.75 lower** (7.31 lower to 3.8 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC, FIHOA)												
5	randomised trials	very serious[a]	not serious	not serious	not serious	none	192	187	-	SMD **0.93 SD lower** (1.14 lower to 0.72 lower)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 13 weeks; assessed with: WOMAC, FIHOA)												
3	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	152	149	-	SMD **0.85 SD lower** (2.15 lower to 0.44 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months (follow-up: mean 6 months; assessed with: WOMAC)												
2	randomised trials	very serious[a]	very serious[b]	not serious	not serious	none	848	838	-	SMD **0.1 lower** (0.2 lower to 0.01 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, FIHOA [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 15 months; assessed with: WOMAC, FIHOA)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised strength exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[a]	not serious	not serious	not serious	none	125	123	-	SMD **0.06 lower** (0.31 lower to 0.19 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Psychological distress (HADS anxiety, 0-21, high is poor, change score) at >3 months (follow-up: 6 months; assessed with: HADS anxiety; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	113	78	-	MD **0.63 lower** (1.54 lower to 0.28 higher)	⨁⨁◯◯ LOW	IMPORTANT
Psychological distress (HADS depression, 0-21, high is poor, change score) at >3 months (follow-up: 6 months; assessed with: HADS depression; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	113	78	-	MD **0.68 lower** (1.3 lower to 0.06 lower)	⨁◯◯◯ VERY LOW	IMPORTANT
Serious adverse events at <3 months (follow-up: 12 weeks)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	5/45 (11.1%)	0/44 (0.0%)	**Peto OR 7.94** (1.32 to 47.77)	**110 more per 1,000** (from 10 more to 210 more)[d]	⨁⨁⨁◯ MODERATE	IMPORTANT
Serious adverse events at >3 months (follow-up: 6 months)												
1	randomised trials	not serious	not serious	not serious	not serious	none	8/65 (12.3%)	0/65 (0.0%)	**Peto OR 8.29** (1.99 to 34.46)	**120 more per 1,000** (from 40 more to 210 more)[d]	⨁⨁⨁⨁ HIGH	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

c. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

F.3 Supervised aerobic exercise compared to no treatment

Table 64: Clinical evidence profile: supervised aerobic exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 12 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	13	15	-	MD **6.8 higher** (6.32 lower to 19.92 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 physical component, 0-100, high is good, change score and final value) at >3 months (follow up: mean 17 months; assessed with: SF-36 physical component; Scale from: 0 to 100)												
2	randomised trials	serious [a]	not serious	not serious	serious [b]	none	108	100	-	MD **1.15 lower** (3.41 lower to 1.11 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 mental component, 0-100, high is good, change score) at >3 months (follow up: mean 17 months; assessed with: SF-36 mental component; Scale from: 0 to 100)												
2	randomised trials	serious [a]	not serious	not serious	serious [b]	none	108	100	-	MD **1.18 lower** (3.46 lower to 1.11 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 12 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	13	15	-	MD **13.3 higher** (2.97 higher to 23.63 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months (follow up: 6 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	14	13	-	MD **4.02 lower** (6.01 lower to 2.03 lower)	⨁⨁◯◯ LOW	CRITICAL
Pain (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months (follow up: mean 17 months; assessed with: KOOS, WOMAC; Scale from: 0 to 100)												
2	randomised trials	serious [a]	not serious	not serious	not serious	none	107	99	-	MD **1.3 higher** (3 lower to 5.59 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 12 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	13	15	-	MD **11.1 higher** (2.9 lower to 25.1 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months (follow up: 6 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	14	13	-	MD **15.35 lower** (24.02 lower to 6.68 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, WOMAC, 0-100, high is poor, change score and final value) at >3 months (follow up: mean 17 months; assessed with: KOOS, WOMAC; Scale from: 0 to 100)												
2	randomised trials	serious [a]	not serious	not serious	not serious	none	107	99	-	MD **1.87 lower** (5.98 lower to 2.24 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Serious adverse events at ≤3 months (follow up: 12 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [a]	not serious	not serious	very serious [c]	none	3/19 (15.8%)	0/18 (0.0%)	**OR 7.86** (0.77 to 80.77)	**160 more per 1,000** (from 20 fewer to 340 more) [d]	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **OR:** Odds ratio

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

F.4 Unsupervised aerobic exercise compared to no treatment

Table 65: Clinical evidence profile: unsupervised aerobic exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, final value) at <3 months (follow-up: 13 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	85	80	-	MD **2.1 higher** (8.86 lower to 13.06 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Quality of life (Nottingham Health Profile physical mobility subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile physical mobility subscale; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	28	26	-	MD **28 lower** (30.77 lower to 25.23 lower)	⨁⨁◯◯ Low	CRITICAL

Quality of life (Nottingham Health Profile pain subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile pain subscale; Scale from: 0 to 100)

1	randomised trials	very serious[a]	not serious	not serious	not serious	none	28	26	-	MD **11.4 lower** (13.13 lower to 9.67 lower)	⨁⨁◯◯ Low	CRITICAL

Quality of life (Nottingham Health Profile energy subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile energy subscale; Scale from: 0 to 100)

1	randomised trials	very serious[a]	not serious	not serious	not serious	none	28	26	-	MD **34.7 lower** (35.51 lower to 33.89 lower)	⨁⨁◯◯ Low	CRITICAL

Quality of life (Nottingham Health Profile sleep subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile sleep subscale; Scale from: 0 to 100)

1	randomised trials	very serious[a]	not serious	not serious	not serious	none	28	26	-	MD **15.7 lower** (17.95 lower to 13.45 lower)	⨁⨁◯◯ Low	CRITICAL

Quality of life (Nottingham Health Profile emotional reactions subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile emotional reactions subscale; Scale from: 0 to 100)

1	randomised trials	very serious[a]	not serious	not serious	not serious	none	28	26	-	MD **21 lower** (23.06 lower to 18.94 lower)	⨁⨁◯◯ Low	CRITICAL

Quality of life (Nottingham Health Profile social isolation subscale, 0-100, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Nottingham Health Profile social isolation subscale; Scale from: 0 to 100)

1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	28	26	-	MD **1.9 lower** (4.21 lower to 0.41 higher)	⨁◯◯◯ Very low	CRITICAL

Quality of life (KOOS, 0-100, high is good, final value) at >3 months (follow-up: 12 months; assessed with: KOOS; Scale from: 0 to 100)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	75	71	-	MD **1.2 higher** (10.14 lower to 12.54 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC, VAS)												
3	randomised trials	serious[a]	very serious[c]	not serious	serious[b]	none	146	140	-	SMD **1.49 SD lower** (3.11 lower to 0.14 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (VAS, 0-10, high is poor, final value) at >3 months (follow-up: 12 months; assessed with: VAS; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	76	71	-	MD **0.3 lower** (1.82 lower to 1.22 higher)	⨁⨁◯◯ Low	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC)												
3	randomised trials	serious[a]	very serious[c]	not serious	serious[b]	none	145	139	-	SMD **2.1 SD lower** (4.38 lower to 0.18 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-100, high is good, final value) at >3 months (follow-up: 12 months; assessed with: WOMAC; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	75	72	-	MD **5 higher** (7.45 lower to 17.45 higher)	⨁⨁◯◯ Low	CRITICAL
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at <3 months (follow-up: 13 weeks; assessed with: HADS anxiety; Scale from: 0 to 21)												
1	randomised trials	serious[a]	not serious	not serious	not serious	none	85	79	-	MD **0.7 lower** (2.16 lower to 0.76 higher)	⨁⨁⨁◯ Moderate	IMPORTANT
Psychological distress (HADS depression, 0-21, high is poor, final value) at <3 months (follow-up: 13 weeks; assessed with: HADS depression; Scale from: 0 to 21)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised aerobic exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	not serious	none	85	79	-	MD **0.6 lower** (2.16 lower to 0.96 higher)	⨁⨁⨁◯ Moderate	IMPORTANT
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at >3 months (follow-up: 12 months; assessed with: HADS anxiety; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	75	72	-	MD **1 lower** (2.63 lower to 0.63 higher)	⨁◯◯◯ Very low	IMPORTANT
Psychological distress (HADS depression, 0-21, high is poor, final value) at >3 months (follow-up: 12 months; assessed with: HADS depression; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	75	72	-	MD **0.6 lower** (2.23 lower to 1.03 higher)	⨁⨁◯◯ Low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

F.5 Other supervised exercise compared to supervised strength exercise, unsupervised strength exercise and no treatment

Table 66: Clinical evidence profile: other supervised exercise compared to supervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, Assessment of Quality of Life Instrument version two, WHO Quality of Life total [different scale ranges], high is good, final values) at <3 months (follow-up: mean 9 weeks; assessed with: KOOS, Assessment of Quality of Life Instrument version two, WHO Quality of Life total)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	71	77	-	SMD **0 SD** (0.32 lower to 0.32 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-12 physical score, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-12 physical score; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	35	35	-	MD **5.7 higher** (0.25 lower to 11.65 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental score, 0-100, high is good, final values) at <3 months (follow-up: mean 10 weeks; assessed with: SF-36 mental component, SF-12 mental score; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	67	69	-	MD **4.97 lower** (9.23 lower to 0.7 lower)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 physical functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **16.56 higher** (13.52 higher to 19.6 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **26.84 higher** (23.87 higher to 29.81 higher)	⨁⨁◯◯ Low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 role physical, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **28.11 higher** (19.78 higher to 36.44 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **16.85 lower** (18.46 lower to 15.24 lower)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **17.35 higher** (13.07 higher to 21.63 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **17.94 lower** (19.35 lower to 16.53 lower)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **27.66 higher** (20.17 higher to 35.15 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	125	125	-	MD **6.89 higher** (4.49 higher to 9.29 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 mental component, 0-100, high is good, final value) at >3 months (follow-up: 15 weeks; assessed with: SF-36 mental component; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	34	-	MD **0.7 lower** (6.16 lower to 4.76 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical functioning, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 physical functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **6.7 higher** (13.31 lower to 26.71 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **4.5 higher** (8.97 lower to 17.97 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **3.4 higher** (8.88 lower to 15.68 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **4.2 higher** (5.21 lower to 13.61 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **1.2 lower** (11.35 lower to 8.95 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	32	32	-	MD **8.5 higher** (0.54 lower to 17.54 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **3.5 lower** (24.37 lower to 17.37 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	32	32	-	MD **6.7 higher** (5.8 lower to 19.2 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (WHO Quality of Life Total, 0-100, high is good, final value) at >3 months (follow-up: 6 months; assessed with: WHO Quality of Life; Scale from: 0 to 100)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	17	17	-	MD **1.94 higher** (2.22 lower to 6.1 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score) at <3 months (follow-up: mean 8 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	15	18	-	MD **0.22 higher** (1.69 lower to 2.12 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 9 weeks; assessed with: KOOS, WOMAC, VAS)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
13	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	396	401	-	SMD **0.18 SD lower** (0.43 lower to 0.79 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months (follow-up: 19 weeks; assessed with: WOMAC, VAS)												
3	randomised trials	very serious[a]	very serious[c]	not serious	not serious	none	82	84	-	SMD **0.37 SD higher** (0.03 higher to 0.71 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score) at <3 months (follow-up: mean 7 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	13	17	-	MD **4.92 lower** (13.86 lower to 4.02 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, WOMAC, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 8 weeks; assessed with: KOOS, WOMAC, VAS)												
10	randomised trials	very serious[a]	serious[c]	not serious	not serious	none	231	238	-	SMD **0.03 SD lower** (0.4 lower to 0.33 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months (follow-up: 15 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	32	34	-	MD **0.4 higher** (5.18 lower to 5.98 higher)	⨁◯◯◯ Very low	CRITICAL
Serious adverse events at <3 months (follow-up: mean 11 weeks)												
2	randomised trials	very serious[a]	serious[d]	not serious	very serious[b]	none	13/46 (28.3%)	10/44 (22.7%)	**RD 0.05** (-0.11 to 0.20)	**50 more per 1,000** (from 110 fewer to 200 more)[e]	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at >3 months (follow-up: 16 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	1/32 (3.1%)	0/32 (0.0%)	**OR 7.39** (0.15 to 372.38)	**30 more per 1,000** (from 50 fewer to 110 more)[e]	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **OR:** odds ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 67: Clinical evidence profile: other supervised exercise compared to unsupervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 physical component, 0-100, high is good, final value) at <3 months (follow-up: 8 weeks; assessed with: SF-36 physical component; Scale from: 0 to 100)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	24	20	-	MD **1.9 higher** (3.31 lower to 7.11 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental component, 0-100, high is good, final value) at <3 months (follow-up: 8 weeks; assessed with: SF-36 mental component; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	24	20	-	MD **6.4 higher** (0.79 lower to 13.59 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (NRS, 0-10, high is poor, change score) at <3 months (follow-up: 4 weeks; assessed with: NRS; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	40	40	-	MD **0.5 lower** (1.29 lower to 0.29 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (WOMAC, NRS [different scale ranges], high is poor, final scores) at <3 months (follow-up: mean 8 weeks; assessed with: WOMAC, NRS)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	102	98	-	SMD **1.03 SD lower** (1.33 lower to 0.74 lower)	⨁⨁◯◯ Low	CRITICAL
Pain (VAS, 0-10, high is poor, final value) at >3 months (follow-up: 6 months; assessed with: VAS; Scale from: 0 to 10)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	18	18	-	MD **0.2 lower** (0.33 lower to 0.07 lower)	⨁⨁◯◯ Low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at <3 months (follow-up: 8 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	60	60	-	MD **20.8 lower** (26.68 lower to 14.92 lower)	⨁⨁◯◯ Low	CRITICAL
Serious adverse events at <3 months (follow-up: 4 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	8/40 (20.0%)	14/40 (35.0%)	**RR 0.57** (0.27 to 1.21)	**151 fewer per 1,000** (from 256 fewer to 73 more)	⨁◯◯◯ Very low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 68: Clinical evidence profile: other supervised exercise compared to supervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	supervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (VAS, 0-100, high is poor, final value) at <3 months (follow-up: 2 months; assessed with: VAS; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	22	18	-	MD **20.5 lower** (40.01 lower to 0.99 lower)	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval; **MD:** mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 69: Clinical evidence profile: other supervised exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, AQoL)												
4	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	133	124	-	SMD **0.44 SD higher** (0.14 lower to 1.02 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change scores and final values) at <3 months (follow-up: mean 9 weeks; assessed with: SF-36 physical component, SF-12 physical component; Scale from: 0 to 100)												
6	randomised trials	very serious[a]	serious[b]	not serious	serious[c]	none	232	138	-	MD **4 higher** (0.56 higher to 7.44 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change scores and final values) at <3 months (follow-up: mean 9 weeks; assessed with: SF-36 mental component, SF-12 mental component; Scale from: 0 to 100)												
6	randomised trials	very serious[a]	serious[b]	not serious	serious[c]	none	232	138	-	MD **3.37 higher** (0.11 lower to 6.85 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, final value) at <3 months												
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	11	10	-	MD **12.1 higher** (7.12 lower to 31.32 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, final value) at <3 months												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	11	10	-	MD **9.4 higher** (0.97 lower to 19.77 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at <3 months												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	11	10	-	MD **2.5 lower** (25.05 lower to 20.05 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (KOOS, 0-100, high is good, change score) at >3 months (follow-up: 16 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	42	42	-	MD **4 higher** (2 lower to 10 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (EQ-5D VAS, Quality of Well-being scale [different scale ranges], high is good, final values) at >3 months (follow-up: mean 11 months; assessed with: EQ-5D VAS, Quality of Well-being scale)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	251	278	-	SMD **0.1 higher** (0.07 lower to 0.27 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, NRS [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC, NRS)												
4	randomised trials	serious[a]	not serious	not serious	not serious	none	144	128	-	SMD **0.79 SD lower** (1.04 lower to 0.54 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, WOMAC)												
17	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	532	417	-	SMD **0.5 SD lower** (0.63 lower to 0.36 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (KOOS, 0-100, high is good, change scores) at >3 months (follow-up: mean 8 months; assessed with: KOOS; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	66	60	-	MD **3.76 higher** (0.12 lower to 7.64 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 12 months; assessed with: WOMAC, HAQ)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	306	313	-	SMD **0.12 lower** (0.28 lower to 0.04 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change scores) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC)												
3	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	58	42	-	MD **9.26 lower** (13.77 lower to 4.74 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, WOMAC, Multidimensional Health Assessment Questionnaire [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 9 weeks; assessed with: KOOS, WOMAC, Multidimensional Health Assessment Questionnaire)												
15	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	491	388	-	SMD **0.47 SD lower** (0.61 lower to 0.33 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, 0-100, high is good, change scores) at >3 months (follow-up: 16 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	42	42	-	MD **4 higher** (0.13 higher to 7.87 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, HAQ [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 12 months; assessed with: WOMAC, HAQ)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	307	315	-	SMD **0.22 lower** (0.38 lower to 0.06 lower)	⨁⨁◯◯ LOW	CRITICAL
Psychological distress (HADS anxiety subscale, DAS scale anxiety subscale [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: HADS anxiety subscale, DAS scale anxiety subscale)												
2	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	143	64	-	SMD **0.33 lower** (0.63 lower to 0.03 lower)	⨁◯◯◯ VERY LOW	IMPORTANT
Psychological distress (HADS depression subscale, DAS scale depression subscale [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: HADS depression subscale, DAS scale depression subscale)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other supervised exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	143	64	-	SMD **0.23 lower** (0.53 lower to 0.06 higher)	⨁◯◯◯ VERY LOW	IMPORTANT
Psychological distress (DAS scale stress subscale, 0-48, high is poor, final value) at <3 months												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	111	41	-	MD **5 lower** (8.68 lower to 1.32 lower)	⨁◯◯◯ VERY LOW	IMPORTANT
Psychological distress (Centre for Epidemiological Studies Depression Scale, 0-60, high is poor, final value) at >3 months												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	101	113	-	MD **1.14 lower** (2.58 lower to 0.3 higher)	⨁⨁◯◯ LOW	IMPORTANT
Serious adverse events at <3 months (follow-up: mean 9 weeks)												
4	randomised trials	very serious[a]	serious[d]	not serious	not serious	none	15/94 (16.0%)	0/86 (0.0%)	**RD 0.07** (-0.10 to 0.25)	**70 more per 1,000** (from 100 fewer to 250 more)[e]	⨁◯◯◯ VERY LOW	IMPORTANT
Serious adverse events at >3 months (follow-up: 16 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[c]	none	2/43 (4.7%)	1/44 (2.3%)	**RR 2.05** (0.19 to 21.75)	**24 more per 1,000** (from 18 fewer to 472 more)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

c. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

F.6 Other unsupervised exercise compared to unsupervised strength exercise

Table 70: Clinical evidence profile: other unsupervised exercise compared to unsupervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	other unsupervised exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 4 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	24	18	-	MD **17 lower** (28.24 lower to 5.76 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 4 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	24	18	-	MD **3 lower** (11.48 lower to 5.48 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 4 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	24	18	-	MD **6 lower** (13.88 lower to 1.88 higher)	⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

F.7 Supervised mixed modality exercise compared to supervised strength exercise, unsupervised strength exercise, supervised aerobic exercise, other supervised exercise, unsupervised mixed modality exercise, pharmacological treatment and no treatment

Table 71: Clinical evidence profile: supervised mixed modality exercise compared to supervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (AQoL, 0-1, high is good, final value) at <3 months (follow-up: 8 weeks; assessed with: AQoL; Scale from: 0 to 1)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	10	10	-	MD **0.01 lower** (0.16 lower to 0.14 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 physical function, 0-100, high is good, final values) at <3 months (follow-up: mean 6 weeks; assessed with: SF-36 physical function; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	serious[c]	not serious	very serious[b]	none	73	70	-	MD **5.81 higher** (6.88 lower to 18.49 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, final values) at <3 months (follow-up: mean 6 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	serious[c]	not serious	very serious[b]	none	73	70	-	MD **8.15 higher** (9.2 lower to 25.5 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, final values) at <3 months (follow-up: mean 7 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												
3	randomised trials	very serious[a]	serious[c]	not serious	serious[b]	none	106	103	-	MD **5.4 higher** (0.7 lower to 11.51 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at <3 months (follow-up: 6 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	43	40	-	MD **1.32 higher** (0.89 higher to 1.75 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, final value) at <3 months (follow-up: 6 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	43	40	-	MD **1.25 higher** (0.8 higher to 1.7 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, final value) at <3 months (follow-up: 6 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	43	40	-	MD **0.98 higher** (0.47 higher to 1.48 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at <3 months (follow-up: 6 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	43	40	-	MD **0.44 higher** (0.16 higher to 0.72 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at <3 months (follow-up: 6 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	43	40	-	MD **0.61 higher** (2.5 lower to 3.72 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, final value) at >3 months (follow-up: 8 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	33	33	-	MD **7.25 higher** (0.57 higher to 13.93 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, VAS, NRS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 8 weeks; assessed with: WOMAC, NRS)												
10	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	266	259	-	SMD **0.67 SD lower** (1.09 lower to 0.24 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, VAS, NRS [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 39 weeks; assessed with: WOMAC, NRS)												
3	randomised trials	very serious[a]	serious[c]	not serious	serious[b]	none	136	132	-	SMD **0.3 SD lower** (0.65 lower to 0.05 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 8 weeks; assessed with: WOMAC)												
9	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	275	268	-	SMD **0.83 SD lower** (1.3 lower to 0.36 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 39 weeks; assessed with: WOMAC)												
3	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	136	132	-	SMD **0.5 SD lower** (1.08 lower to 0.08 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse events at <3 months (follow-up: mean 9 weeks)												
3	randomised trials	very serious[a]	not serious	not serious	serious[d]	none	0/93 (0.0%)	0/100 (0.0%)	**RD 0.00** (-0.07 to 0.07)	**0 fewer per 1,000** (from 70 fewer to 70 more)[e]	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Serious adverse events at >3 months (follow-up: 6 months)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[d]	none	0/53 (0.0%)	0/60 (0.0%)	**RD 0.00** (-0.03 to 0.03)	**0 fewer per 1,000** (from 30 fewer to 30 more)[e]	⨁◯◯◯ VERY LOW	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

d. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 72: Clinical evidence profile: supervised mixed modality exercise compared to unsupervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 physical component, 0-100, high is good, final value) at ≤3 months (follow up: 8 weeks; assessed with: SF-36 physical component; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	22	20	-	MD **3.5 higher** (1.85 lower to 8.85 higher)	⨁⨁◯◯ LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 mental component, 0-100, high is good, final value) at ≤3 months (follow up: 8 weeks; assessed with: SF-36 mental component; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	22	20	-	MD **4.5 higher** (2.66 lower to 11.66 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (BPI mean pain, 0-10, high is poor, final value) at ≤3 months (follow up: 8 weeks; assessed with: NRS; Scale from: 0 to 10)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	22	20	-	MD **1.09 lower** (2.08 lower to 0.1 lower)	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 73: Clinical evidence profile: supervised mixed modality exercise compared to supervised aerobic exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (WOMAC, 0-20, high is poor, change score) at >3 months (follow up: 5 months; assessed with: WOMAC; Scale from: 0 to 20)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	supervised aerobic exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	39	40	-	MD **1 lower** (2.37 lower to 0.37 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months (follow up: 5 months; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	39	40	-	MD **2.8 lower** (7.21 lower to 1.61 higher)	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 74: Clinical evidence profile: supervised mixed modality exercise compared to other supervised exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	other supervised exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	26	26	-	MD **1 higher** (5.26 lower to 7.26 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (EQ-5D, -0.11-1, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: EQ-5D; Scale from: -0.11 to 1)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	other supervised exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	45	45	-	MD **0.03 lower** (0.08 lower to 0.02 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at <3 months (follow-up: mean 8 weeks; assessed with: SF-36 physical component, SF-12 physical component; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	50	56	-	MD **0.58 lower** (3.75 lower to 2.59 higher)	⨁⨁⨁◯ Moderate	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at <3 months (follow-up: mean 8 weeks; assessed with: SF-36 mental component, SF-12 mental component; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	serious[b]	none	50	56	-	MD **1.63 lower** (4.98 lower to 1.72 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (KOOS, WOMAC, BPI, VAS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, WOMAC, BPI, VAS)												
6	randomised trials	very serious[a]	not serious	not serious	not serious	none	164	170	-	SMD **0.14 SD higher** (0.08 lower to 0.35 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (WOMAC, VAS [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 21 weeks; assessed with: VAS; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	serious[c]	not serious	serious[b]	none	77	72	-	SMD **0.13 SD higher** (0.38 lower to 0.65 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, WOMAC)												
4	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	110	114	-	SMD **0.03 SD higher** (0.58 lower to 0.64 higher)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months (follow-up: 26 weeks; assessed with: WOMAC; Scale from: 0 to 68)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	other supervised exercise	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	45	40	-	MD **1.9 higher** (1.52 lower to 5.32 higher)	⨁◯◯◯ Very low	CRITICAL
Psychological distress (HADS-anxiety, 0-21, high is poor, final value) at <3 months (follow-up: 8 weeks; assessed with: HADS-anxiety; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	28	32	-	MD **1.4 higher** (0.05 higher to 2.75 higher)	⨁◯◯◯ Very low	IMPORTANT
Psychological distress (HADS-depression, 0-21, high is poor, final value) at <3 months (follow-up: 8 weeks; assessed with: HADS-depression; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	28	32	-	MD **0.4 higher** (0.61 lower to 1.41 higher)	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at <3 months (follow-up: 12 weeks)												
1	randomised trials	serious[a]	not serious	not serious	very serious[b]	none	3/45 (6.7%)	5/45 (11.1%)	**RR 0.60** (0.15 to 2.36)	**44 fewer per 1,000** (from 94 fewer to 151 more)	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at >3 months (follow-up: 24 weeks)												
1	randomised trials	serious[a]	not serious	not serious	serious[d]	none	0/49 (0.0%)	0/49 (0.0%)	**RD 0.00** (-0.04 to 0.04)	**0 fewer per 1,000** (from 40 fewer to 40 more)[e]	⨁⨁◯◯ Low	IMPORTANT

CI: confidence interval; **MD:** mean difference; **RR:** risk ratio; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

d. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

e. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 75: Clinical evidence profile: supervised mixed modality exercise compared to unsupervised mixed modality exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	unsupervised mixed modality exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 physical function, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 physical function; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	41	39	-	MD **4.05 higher** (2.18 lower to 10.28 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	41	39	-	MD **9.99 higher** (2.2 higher to 17.78 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	41	39	-	MD **15.54 higher** (2.10 higher to 28.98 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	very serious [b]	none	41	39	-	MD **1.67 higher** (8.34 lower to 11.68 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	unsupervised mixed modality exercise	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 general health, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	41	39	-	MD **9.73 higher** (2.84 higher to 16.62 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	very serious [b]	none	41	39	-	MD 0 **higher** (7.50 lower to 7.50 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	41	39	-	MD **25.98 higher** (11.58 higher to 40.38 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 social functioning, 0-100, high is good, final value) at ≤3 months (follow up: 6 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	41	39	-	MD **59.58 lower** (67.03 lower to 52.13 lower)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, VAS, [different scale ranges], high is poor, final value) at ≤3 months (follow up: mean 6 weeks; assessed with: WOMAC, VAS)												
2	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	71	69	-	**SMD 0.35 SD lower** (0.69 lower to 0.02 lower)	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	unsupervised mixed modality exercise	Relative (95% CI)	Absolute (95% CI)		
Physical function (WOMAC, 0-20, high is poor, final value) at ≤3 months (follow up: 6 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	41	39	-	MD **5.18 lower** (8.97 lower to 1.39 lower)	⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **SMD:** Standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

Table 76: Clinical evidence profile: supervised mixed modality exercise compared to pharmacological treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	pharmacological treatments	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 8 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	47	46	-	MD **1.36 lower** (6.58 lower to 3.86 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (KOOS, 0-100, high is good, final value) at >3 months (follow up: 52 weeks; assessed with: KOOS; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	pharmacological treatments	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [a]	not serious	not serious	not serious	none	47	46	-	MD **1.3 higher** (4.9 lower to 7.5 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Pain (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 8 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	47	46	-	MD **2.08 higher** (2.28 lower to 6.44 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, 0-500, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: WOMAC; Scale from: 0 to 500)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	51	53	-	MD **23.1 lower** (60.11 lower to 13.91 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (KOOS, 0-100, high is good, change score) at >3 months (follow up: 52 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	47	46	-	MD **4.2 higher** (1.45 lower to 9.85 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, 0-500, high is poor, final value) at >3 months (follow up: 26 weeks; assessed with: WOMAC; Scale from: 0 to 500)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	51	53	-	MD **19.9 lower** (56.08 lower to 16.28 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, 0-100, high is good, change score) at ≤3 months (follow up: 8 weeks; assessed with: KOOS; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	pharmacological treatments	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [a]	not serious	not serious	not serious	none	47	46	-	MD **0.5 lower** (5.02 lower to 4.02 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (WOMAC, 0-1800, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: WOMAC; Scale from: 0 to 1800)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	51	53	-	MD **89.2 lower** (216.18 lower to 37.78 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (KOOS, 0-100, high is good, change score) at >3 months (follow up: 52 weeks; assessed with: KOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	47	46	-	MD **3.5 higher** (2.01 lower to 9.01 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-1800, high is poor, final value) at >3 months (follow up: 6 months; assessed with: WOMAC; Scale from: 0 to 1800)												
1	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	51	53	-	MD **72.9 lower** (202.71 lower to 56.91 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse events at >3 months (follow up: 26 weeks)												
1	randomised trials	serious [a]	not serious	not serious	serious [c]	none	0/55 (0.0%)	0/55 (0.0%)	**RD 0.00** (-0.03 to 0.03)	**0 fewer per 1,000** (from 30 fewer to 30 more) [d]	⨁⨁◯◯ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 77: Clinical evidence profile: supervised mixed modality exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (KOOS, AQoL [different scale ranges], high is good, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, AQoL)												
2	randomised trials	serious[a]	not serious	not serious	serious[b]	none	36	36	-	SMD **0.56 higher** (0.09 higher to 1.04 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at <3 months (follow-up: mean 9 weeks; assessed with: SF-36 physical component, SF-12 physical component; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	serious[b]	none	73	66	-	MD **1.66 higher** (1.57 lower to 4.89 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at <3 months (follow-up: mean 9 weeks; assessed with: SF-36 mental component, SF-12 mental component; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	73	66	-	MD **0.73 higher** (2.95 lower to 4.41 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical function, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 physical function; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **25.35 higher** (24.44 lower to 75.13 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 bodily pain, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 bodily pain; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	86	77	-	MD **25.86 higher** (15.48 lower to 67.2 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 role physical, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 role physical; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **41.88 higher** (42.4 lower to 126.15 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 vitality, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 vitality; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **24.77 higher** (27.05 lower to 76.6 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 general health, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 general health; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **19.57 higher** (14.21 lower to 53.36 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 mental health, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 mental health; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **16.61 higher** (14.65 lower to 47.86 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 role emotional, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 role emotional; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **34.83 higher** (37.46 lower to 107.12 higher)	⨁◯◯◯ Very low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 social functioning, 0-100, high is good, change score and final value) at <3 months (follow-up: 12 weeks; assessed with: SF-36 social functioning; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[c]	not serious	very serious[b]	none	87	78	-	MD **27.94 higher** (29.14 lower to 85.03 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 arm function, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 arm function; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.13 lower** (0.44 lower to 0.18 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 arthritis pain, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 arthritis pain; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **0.85 lower** (1.52 lower to 0.18 lower)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 hand and finger function, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 hand and finger function; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.1 lower** (0.52 lower to 0.32 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 household tasks, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 household tasks; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **0.24 lower** (0.56 lower to 0.08 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 level of tension, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 level of tension; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **0.42 lower** (1.12 lower to 0.28 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 mobility level, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 mobility level; Scale from: 0 to 10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **0.5 lower** (0.93 lower to 0.07 lower)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 mood, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 mood; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.16 lower** (0.69 lower to 0.37 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 self-care tasks, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 self-care tasks; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.01 lower** (0.14 lower to 0.12 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 social activity, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 social activity; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.08 lower** (0.63 lower to 0.47 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 support from family and friends, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 support from family and friends; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	59	65	-	MD **0.08 lower** (0.82 lower to 0.66 higher)	⨁⨁◯◯ Low	CRITICAL
Quality of life (AIMS2 walking and bending, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 walking and bending; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **1.25 lower** (2.08 lower to 0.42 lower)	⨁◯◯◯ Very low	CRITICAL
Quality of life (AIMS2 work, 0-10, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS2 work; Scale from: 0 to 10)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	59	65	-	MD **0.39 lower** (0.88 lower to 0.1 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (WOMAC, VAS, 0-100, high is poor, change scores) at <3 months (follow-up: 12 weeks; assessed with: WOMAC, VAS; Scale from: 0 to 100)												
2	randomised trials	serious[a]	very serious[c]	not serious	serious[b]	none	164	167	-	MD **11.83 lower** (21.42 lower to 2.24 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (KOOS, WOMAC, AIMS, VAS, NRS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: KOOS, WOMAC, AIMS, VAS, NRS)												
10	randomised trials	very serious[a]	very serious[c]	not serious	serious[b]	none	245	231	-	SMD **0.67 SD lower** (1.04 lower to 0.29 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (VAS, 0-100, high is poor, change scores) at >3 months (follow-up: mean 44 weeks; assessed with: VAS; Scale from: 0 to 100)												
2	randomised trials	serious[a]	not serious	not serious	serious[b]	none	141	143	-	MD **7.61 lower** (13.78 lower to 1.44 lower)	⨁⨁◯◯ Low	CRITICAL
Pain (KOOS, NRS [different scale ranges], high is poor, final values) at >3 months (follow-up: 42 weeks; assessed with: KOOS, NRS)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	63	69	-	SMD **0.63 lower** (0.98 lower to 0.27 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-100, high is poor, change score) at <3 months (follow-up: 12 weeks; assessed with: WOMAC; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	71	68	-	MD **6.3 lower** (10.67 lower to 1.93 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (KOOS, WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 9 weeks; assessed with: KOOS, WOMAC)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
7	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	201	191	-	SMD **0.42 lower** (0.62 lower to 0.22 lower)	⨁◯◯◯ Very low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months (follow-up: 32 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	51	56	-	MD **7.9 lower** (12.78 lower to 3.02 lower)	⨁◯◯◯ Very low	CRITICAL
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at <3 months (follow-up: 9 weeks; assessed with: HADS anxiety; Scale from: 0 to 21)												
2	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	73	66	-	MD **0.71 higher** (0.43 lower to 1.85 higher)	⨁◯◯◯ Very low	IMPORTANT
Psychological distress (HADS depression, 0-21, high is poor, final value) at <3 months (follow-up: 9 weeks; assessed with: HADS depression; Scale from: 0 to 21)												
2	randomised trials	very serious[a]	not serious	not serious	not serious	none	73	66	-	MD **0.09 higher** (0.8 lower to 0.98 higher)	⨁⨁◯◯ Low	IMPORTANT
Psychological distress (AIMS psychological disability, 0-10, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: AIMS psychological disability; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	16	18	-	MD **0.08 higher** (0.56 lower to 0.72 higher)	⨁◯◯◯ Very low	IMPORTANT
Psychological distress (HADS, 0-21, high is poor, final value) at >3 months (follow-up: 32 weeks; assessed with: HADS; Scale from: 0 to 21)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	51	55	-	MD **1.6 higher** (0.91 lower to 4.11 higher)	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at <3 months (follow-up: mean 11 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	supervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	very serious[a]	serious[d]	not serious	very serious[e]	none	1/140 (0.7%)	0/144 (0.0%)	**RD 0.01** (-0.02 to 0.04)	**10 fewer per 1,000** (from 40 fewer to 20 more)	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at >3 months (follow-up: 52 weeks)												
1	randomised trials	not serious	not serious	not serious	serious[b]	none	0/51 (0.0%)	1/51 (2.0%)	**Peto OR 0.14** (0.00 to 6.82)	**20 fewer per 1,000** (from 70 fewer to 30 more)[f]	⨁⨁⨁◯ Moderate	IMPORTANT

CI: confidence interval; **MD:** mean difference; **SMD:** standardised mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

d. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

e. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

f. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

F.8 Unsupervised mixed modality exercise compared to unsupervised strength exercise, other unsupervised exercise, pharmacological treatment and no treatment

Table 78: Clinical evidence profile: unsupervised mixed modality exercise compared to unsupervised strength exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised mixed modality exercise	unsupervised strength exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (WOMAC, 0-20, high is poor, change score) at <3 months (follow-up: 4 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
1	randomised trials	very serious[a]	not serious	not serious	serious[b]	none	16	16	-	MD **1.12 lower** (2.08 lower to 0.16 lower)	⨁◯◯◯ Very low	CRITICAL
Pain (VAS, NRS, 0-10, high is poor, final values) at <3 months (follow-up: 6 weeks; assessed with: VAS, NRS; Scale from: 0 to 10)												
2	randomised trials	very serious[a]	serious[c]	not serious	very serious[b]	none	90	99	-	MD **0.05 lower** (1.17 lower to 1.06 higher)	⨁◯◯◯ Very low	CRITICAL
Pain (NRS, 0-10, high is poor, final value) at >3 months (follow-up: 12 months; assessed with: NRS; Scale from: 0 to 10)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	76	66	-	MD **0.1 higher** (0.86 lower to 1.06 higher)	⨁⨁◯◯ Low	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score and final value) at <3 months (follow-up: mean 6 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	very serious[a]	not serious	not serious	very serious[b]	none	100	91	-	MD **0.76 lower** (6.59 lower to 5.07 higher)	⨁◯◯◯ Very low	CRITICAL

CI: confidence interval; **MD:** mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

Table 79: Clinical evidence profile: unsupervised mixed modality exercise compared to other unsupervised exercise

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised mixed modality exercise	other unsupervised exercise	Relative (95% CI)	Absolute (95% CI)		
Pain (WOMAC, 0-100, high is poor, change score) at ≤3 months (follow up: 8 weeks; assessed with: WOMAC; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	91	88	-	MD **7 higher** (4.64 higher to 9.36 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-100, high is poor, change score) at ≤3 months (follow up: 8 weeks; assessed with: WOMAC; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	91	88	-	MD **9 higher** (7.62 higher to 10.38 higher)	⨁⨁◯◯ LOW	CRITICAL
Serious adverse events at ≤3 months (follow up: 8 weeks)												
1	randomised trials	serious [a]	not serious	not serious	very serious [b]	none	5/91 (5.5%)	8/88 (9.1%)	**RR 0.60** (0.21 to 1.78)	**36 fewer per 1,000** (from 72 fewer to 71 more)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

Table 80: Clinical evidence profile: unsupervised mixed modality exercise compared to pharmacological treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised mixed modality exercise	pharmacological treatments	Relative (95% CI)	Absolute (95% CI)		
Pain (HSS pain during activity, VAS [different scale ranges], high is poor, final values) at >3 months (follow up: mean 15 months; assessed with: HSS pain during activity, VAS)												
2	randomised trials	very serious [a]	not serious	not serious	serious [b]	none	68	67	-	SMD **0.27 higher** (0.07 lower to 0.61 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (VAS, 0-100, high is poor, change score) at >3 months (follow up: 24 weeks; assessed with: VAS; Scale from: 0 to 100)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	60	60	-	MD **0.83 lower** (12.32 lower to 10.66 higher)	⨁⨁◯◯ LOW	CRITICAL
Serious adverse events at >3 months (follow up: 24 weeks)												
1	randomised trials	not serious	not serious	not serious	serious [c]	none	0/60 (0.0%)	0/60 (0.0%)	**RD 0.00** (-0.03 to 0.03)	**0 fewer per 1,000** (from 30 fewer to 30 more) [d]	⨁⨁⨁◯ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference

Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size

d. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Table 81: Clinical evidence profile: unsupervised mixed modality exercise compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at ≤3 months (follow up: 12 weeks; assessed with: EQ-5D; Scale from: -0.329 to 1.0)												
1	randomised trials	serious [a]	not serious	not serious	very serious [b]	none	101	102	-	MD **0** (0.04 lower to 0.05 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (EQ-5D, -0.329-1.0, high is good, final value) at >3 months (follow up: 12 months; assessed with: EQ-5D; Scale from: -0.329 to 1.0)												
1	randomised trials	serious [a]	not serious	not serious	very serious [b]	none	101	102	-	MD **0** (0.05 lower to 0.05 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (HOOS, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: HOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	101	102	-	MD **4.4 lower** (9.44 lower to 0.64 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score) at >3 months (follow up: 12 months; assessed with: WOMAC; Scale from: 0 to 20)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	142	68	-	MD **0.51 lower** (1.43 lower to 0.41 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (HOOS, 0-100, high is poor, final value) at >3 months (follow up: 12 months; assessed with: HOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	not serious	none	101	102	-	MD **3 lower** (8.34 lower to 2.34 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (HOOS, 0-100, high is poor, final value) at ≤3 months (follow up: 12 weeks; assessed with: HOOS; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	unsupervised mixed modality exercise	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	101	102	-	MD **6.9 lower** (12.45 lower to 1.35 lower)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months (follow up: 12 months; assessed with: WOMAC; Scale from: 0 to 68)												
1	randomised trials	very serious [a]	not serious	not serious	not serious	none	142	68	-	MD **1.89 lower** (4.72 lower to 0.94 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (HOOS, 0-100, high is poor, final value) at >3 months (follow up: 12 months; assessed with: HOOS; Scale from: 0 to 100)												
1	randomised trials	serious [a]	not serious	not serious	serious [b]	none	101	102	-	MD **7.4 lower** (13.26 lower to 1.54 lower)	⨁⨁◯◯ LOW	CRITICAL
Serious adverse events at >3 months (follow up: 12 months)												
1	randomised trials	serious [a]	not serious	not serious	very serious [b]	none	4/142 (2.8%)	0/68 (0.0%)	**Peto OR 4.48** (0.54 to 36.96)	**30 more per 1,000** (from 10 fewer to 60 more) [c]	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference

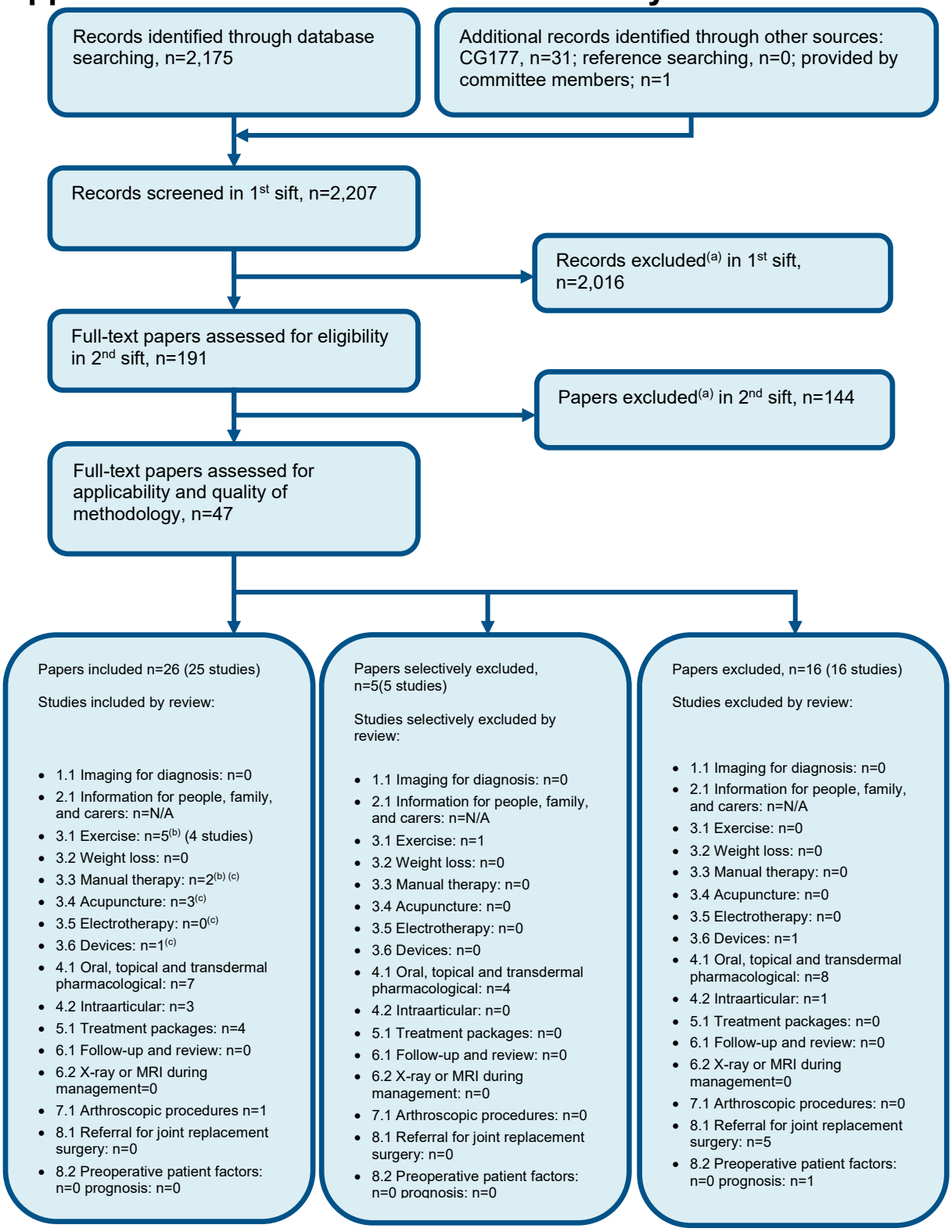
Explanations

a. Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

b. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs

c. Absolute effect calculated by risk difference due to zero events in at least one arm of one study

Appendix G – Economic evidence study selection



(a) Non-relevant population, intervention, comparison, design or setting; non-English language.
(b) Two articles identified were applicable to Q3.1 and Q3.3, for the purposes of this diagram they have been included under Q3.1 only.
(c) One article identified was applicable to Q3.3, Q3.4, Q3.5 and Q3.6, for the purposes of this diagram it has been included under Q3.3 only.

Appendix H – Economic evidence tables

Study	Abbott 2019[3] (Pinto 2013[357])			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Within-trial analysis (Abbott 2013[2] **Approach to analysis:** Analysis of individual level quality of life and resource use data adjusted for age, sex, primary OA joint (hip or knee), BMI, years since symptom onset, and baseline WOMAC, quadricep muscle strength, mental health, self-efficacy, and SF-6D score. Unit costs applied. **Perspective:** New Zealand healthcare (public and private) and societal - only public healthcare perspective reported here. **Follow-up:** 2 years	**Population:** People with hip or knee osteoarthritis meeting American College of Rheumatology clinical diagnostic criteria for hip or knee OA with no previous history of RA or joint replacement, no recent initiation (30 days) of opioid or corticosteroid. **Patient characteristics:** Age: 66 Male: 45% **Intervention 1:** Usual medical care (no trial physiotherapy) **Intervention 2:** Supervised exercise physiotherapy (increasing strength, neuromuscular control and flexibility of the muscles of the lower extremities) in addition to usual care* **Intervention 3:** Manual physiotherapy (improving joint mobility	**Total costs (mean per patient):** Intervention 1: £3,577 Intervention 2: £3,550 Intervention 3: £4,602 Intervention 4: £3,744 Incremental (2−1): saves £27 Incremental (3–2): £1,052 Incremental (4–3): saves £858 (95% CI: NR; p=NR) Intervention costs only: Intervention 1: £0 Intervention 2: £503 Intervention 3: £486 Intervention 4: £507 **Currency & cost year:** 2009 NZ dollars (presented here as 2009 UK pounds[(b)])] **Cost components incorporated:** Medical and other healthcare consumed by participants during the trial.	**QALYs (mean per patient):** Intervention 1: 1.31 Intervention 2: 1.46 Intervention 3: 1.39 Intervention 4: 1.38 Incremental (2−1): 0.15 Incremental (3–2): -0.07 Incremental (4–3): -0.01 (95% CI: NR; p=NR)	Intervention 2 dominates all other interventions. Probability Intervention 2 cost effective (£20K/30K threshold): NR **Analysis of uncertainty:** A sensitivity analysis was undertaken for participants with complete case data only – costs reported for this also include private healthcare costs, but intervention 2 remains dominant. A sensitivity analysis was also undertaken excluding participants who underwent joint replacement surgery – costs reported for this also include private healthcare costs, but intervention 2 remains dominant. Another sensitivity analysis was undertaken excluding productivity losses from the societal perspective analysis (results not informative to UK NHS context and so not reported here)

Discounting: Costs: 3.5%; Outcomes: 3.5%	through manually administered forces to the target joint and surrounding soft tissue) in addition to usual care* **Intervention 4:** **C**ombination of exercise and manual physiotherapy in addition to usual care* *10 individual, supervised 50-minute sessions (7 sessions over a 9-week programme, with 2 booster sessions at week 16 and 54)			
Data sources				
Health outcomes: QALYs calculated by using the time-weighted averages at the beginning and end of each measurement period. SF-12 version 2 questionnaire administered at baseline, 6 months, 1 year and 2 years. **Quality-of-life weights:** SF-6D UK tariff. **Cost sources:** Public healthcare costs - New Zealand case-mix framework for publicly funded hospitals. New Zealand Pharmaceutical Schedule, Otago District Health Board finance pricing, average fees from Dunedin metropolitan area.				
Comments				
Source of funding: Health Research Council of New Zealand and the New Zealand Lottery Grants Board. **Limitations:** Not double-blinded. 2009 New Zealand resource use and unit costs may not reflect current UK NHS practice. Within trial analysis may not reflect full body of evidence available. **Other:** None.				
Overall applicability:(c) Partially applicable **Overall quality:**(d)				

Abbreviations: CCA= cost–consequences analysis; CEA= cost-effectiveness analysis; 95% CI= 95% confidence interval; CUA= cost–utility analysis; da= deterministic analysis; EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); ICER= incremental cost-effectiveness ratio; NR= not reported; pa= probabilistic analysis; QALYs= quality-adjusted life years

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Converted using 2009 purchasing power parities[334]

(c) Directly applicable / Partially applicable / Not applicable

(d) Minor limitations / Potentially serious limitations / Very serious limitations

<table>
<tr><th>Study</th><th colspan="4">Kigozi 2018[200]</th></tr>
<tr><th>Study details</th><th>Population & interventions</th><th>Costs</th><th>Health outcomes</th><th>Cost effectiveness</th></tr>
<tr>
<td>Economic analysis: CUA (health outcome: QALYs)

Study design: Within trial analysis (BEEP RCT, Kigozi 2018[200]
Approach to analysis: Analysis of individual level data for QALYs calculated using EQ5D-3L administered at baseline 3, 6, 9 and 18 months. UK tariff applied. Resource use from trial with unit costs applied.

Perspective: UK NHS
Follow-up: 18 months
Treatment effect duration:[(a)] 18 months
Discounting: Costs: none; Outcomes: none</td>
<td>Population:
Adults with knee osteoarthritis in primary care

Cohort settings:
Start age: 63
Male: 49%

Intervention 1: Usual physical therapy care - advice and lower limb exercise provided in up to four individual one-to-one treatment sessions over 12 weeks (in line with NHS practice)

Intervention 2:
Individually tailored exercise - supervised, individually tailored and progressed lower-limb exercise programme in 6-8 one-to-one treatment sessions over 12 weeks

Intervention 3:
Targeted exercise therapy - started with a focus on lower-limb exercise with</td>
<td>Total costs (mean per patient):
Intervention 1: £383
Intervention 2: £656
Intervention 3: £524

Currency & cost year:
2012/13 UK pounds

Cost components incorporated:
Primary care consultations (GP, nurse practitioners, community physical therapists), consultations with other health-care professionals (hospital consultants, hospital physical therapists, acupuncturists), hospital-based investigations (X-ray and MRI), procedures (injections, surgery), prescribed meds
Intervention costs - sessions. Also included a 47-minute initial assessment and treatment session, followed by 28-minute face to face treatment</td>
<td>QALYs (mean per patient):
Intervention 1: 1.035
Intervention 2: 1.032
Intervention 3: 1.019</td>
<td>Full incremental analysis (b):
<table>
<tr><th>Int</th><th>Cost</th><th>QALY</th><th>Inc cost</th><th>Inc QALY</th><th>ICER</th></tr>
<tr><td>3</td><td>£524</td><td>1.019</td><td colspan="3">Dominated</td></tr>
<tr><td>2</td><td>£656</td><td>1.032</td><td colspan="3">Dominated</td></tr>
<tr><td>1</td><td>£383</td><td>1.035</td><td colspan="3">Baseline</td></tr>
</table>
Probability Intervention 2 or 3 being cost effective compared to 1 (£20K threshold): <40%

Analysis of uncertainty: Complete case analysis to assess impact of missing cost and EQ5D data. This resulted in the same conclusion that usual care was dominant.
Costs outside NHS presented but not reported here.</td>
</tr>
</table>

	an aim to support progress to increasing general physical activity adherence over 6 months. 4 individual face-to-face treatments up to week 12, and a further 4-6 follow-up contacts from week 12-26. All received an information booklet providing information about benefits of exercise and physical activity and a home exercise programme.	session, 11 min telephone call contacts (where applicable)		
Data sources				
Health outcomes: Quality of life taken from within trial analysis of RCT, associated paper Kigozi 2018[200]. Incremental QALY estimates were adjusted (to control for imbalances in baseline utility between the interventions) and imputed for missing data. **Quality-of-life weights:** EQ5D-3L administered at baseline 3, 6, 9 and 18 months. UK tariff applied. **Cost sources:** Resource use from within trial. Unit costs from BNF, NHS reference costs, PSSRU 2012-13				
Comments				
Source of funding: NIHR **Limitations:** Study does not include all exercise treatment options. Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison. **Other:**				
Overall applicability: Partially applicable(c) **Overall quality:** Potentially serious limitations(d)				

Abbreviations:; 95% CI= 95% confidence interval; CUA= cost–utility analysis; EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); ICER= incremental cost-effectiveness ratio; NR= not reported; pa= probabilistic analysis; QALYs= quality-adjusted life years

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Intervention number in order of least to most effective in terms of QALYs.

(c) Directly applicable / Partially applicable / Not applicable

(d) Minor limitations / Potentially serious limitations / Very serious limitations

<table>
<tr><th>Study</th><th colspan="4">Oppong 2014[333]</th></tr>
<tr><th>Study details</th><th>Population & interventions</th><th>Costs</th><th>Health outcomes</th><th>Cost effectiveness</th></tr>
<tr>
<td>Economic analysis: CUA (health outcome: QALYs)

Study design: Within trial analysis (RCT, associated paper Dziedzic 2015[114])
Approach to analysis: Analysis of individual level data for QALYs calculated using EQ5D measured at baseline, 3, 6 and 12 months. UK general population weights applied. Resource use from trial with unit costs applied.
Perspective: UK NHS
Follow-up: 1 year
Treatment effect duration:[(a)] 1 year
Discounting: Costs: n/a; Outcomes: n/a</td>
<td>Population:
Adults aged 50 years or older with hand osteoarthritis

Cohort settings:
Start age: 66
Male: 44%

Intervention 1:
Leaflet and advice only

Intervention 2:
Joint protection only

Intervention 3:
Hand exercises only (stretching and strengthening hand and thumb exercises)

Intervention 4:
Joint protection and hand exercises

Joint protection, hand exercises and joint protection combined with hand exercises were all delivered by an occupational therapist in a</td>
<td>Total costs (mean per patient):
Intervention 1: £58
Intervention 2: £92
Intervention 3: £65
Intervention 4: £112

Currency & cost year:
2010/11 UK pounds
Cost components incorporated:
Intervention, primary care (general practice and nurse); secondary care (orthopaedic surgeon, rheumatologist, plastic surgeon, physiotherapist, occupational therapist), other health care staff and prescribed medication. As all participants received the leaflet and advice, this cost was not included in the analysis</td>
<td>QALYs (mean per patient):
Intervention 1: 0.662
Intervention 2: 0.659
Intervention 3: 0.681
Intervention 4: 0.658</td>
<td>Full incremental analysis (b):
<table>
<tr><th>Int</th><th>Cost</th><th>QALY</th><th>Inc cost</th><th>Inc QALY</th><th>ICER</th></tr>
<tr><td>4</td><td>£112</td><td>0.658</td><td colspan="3">Dominated</td></tr>
<tr><td>2</td><td>£92</td><td>0.659</td><td colspan="3">Dominated</td></tr>
<tr><td>1</td><td>£58</td><td>0.662</td><td colspan="3">Baseline</td></tr>
<tr><td>3</td><td>£65</td><td>0.681</td><td>£6</td><td>0.019</td><td>£318</td></tr>
</table>
Probability Intervention 3 cost effective (£20K threshold): 80%

Analysis of uncertainty: This study explored different analytic methods to generate the cost effectiveness results. Presented above is the within-table analysis which they considered their base case. Other methods were at-the-margins and regression approaches. Hand exercised remained the most cost-effective option for the management of hand osteoarthritis regardless of the approach adopted for the economic analysis.</td>
</tr>
</table>

	group setting. Four group sessions held once a week with 4-6 participants. All of these interventions also received a leaflet and advice.			

Data sources
Health outcomes: Quality of life taken from within trial analysis of RCT, associated paper Dziedzic 2015[114]). For each participant included in the study, a QALY score over the 12-month period was estimated using an area under the curve approach. Incremental QALY estimates were adjusted to control for imbalances in baseline utility between the interventions. **Quality-of-life weights:** EQ-5D measured at baseline, 3, 6 and 12 months. UK general population weights applied. Missing EQ-5D scores were imputed using multiple imputation. **Cost sources:** Healthcare resource data obtained from participants responses to self-report questionnaires administered at 6 and 12 months, For the interventions, information collected on number and grades of staff involved and equipment used to deliver each intervention as well as number of sessions each participant attended. Unit costs taken from BNF, PSSRU and NHS reference costs.
Comments
Source of funding: Arthritis Research UK and Support for Science Funding. **Limitations:** Study does not include all exercise treatment options. Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison. **Other:** None.
Overall applicability: Partially applicable(c) **Overall quality:** Potentially serious limitations (d)

Abbreviations: CUA= cost–utility analysis; da= deterministic analysis; EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); ICER= incremental cost-effectiveness ratio; n/a= not applicable; NR= not reported; pa= probabilistic analysis; QALYs= quality-adjusted life years

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Intervention number in order of least to most effective in terms of QALYs. Costs rounded up.

(c) Directly applicable / Partially applicable / Not applicable

(d) Minor limitations / Potentially serious limitations / Very serious limitations

Study	Tan 2016[434]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome: QALYs) **Study design:** Within trial analysis (RCT, associated paper de Teirlinck 2016 [439] **Approach to analysis:** Analysis of individual level data for QALYs calculated using EQ5D measured at baseline 6, 13, 26, 39 and 52 weeks. Dutch tariff applied. Resource use from trial with unit costs applied. **Perspective:** Dutch healthcare perspective **Follow-up:** 1 year **Treatment effect duration:**[(a)] 1 year **Discounting:** Costs: n/a; Outcomes: n/a	**Population:** Adults with hip osteoarthritis in primary care >45yrs **Cohort settings:** Start age: 1: 66.6 years 2: 64.2 years Male: 1: 45% 2: 38% **Intervention 1:** GP care alone (unrestricted access providing education and counselling and prescription of pain medication if applicable) **Intervention 2:** Exercise plus GP care (exercise was supervised by physiotherapists, up to 12 sessions in first 3 months followed by 3 booster sessions at 5,7 and 9 months, GP care as per intervention 1)	**Total costs (mean per patient):** Intervention 1: £1,124 Intervention 2: £1,041 Incremental (2−1): saves £83 (95% CI: -£649, £459; p=NR) **Currency & cost year:** 2011 Euros (presented here as 2011 UK pounds[(b)]) **Cost components incorporated:** Healthcare professional visits in primary and secondary care, medical investigations/interventions and prescribed medications. Interventions - number and grade of staff involved and equipment use to deliver intervention as well as number of sessions attended.	**QALYs (mean per patient):** Intervention 1: NR Intervention 2: NR Incremental (2−1): -0.006 (95% CI: -0.02 to 0.04; p=NR)	**ICER (Intervention 1 versus Intervention 2):** £13,793 per QALY gained **Analysis of uncertainty:** None undertaken from healthcare perspective
Data sources				
Health outcomes: Quality of life taken from within trial analysis of RCT, associated paper Teirlinck 2016 [439]. Incremental QALY estimates were adjusted to control for imbalances in baseline utility between the interventions. **Quality-of-life weights:** EQ5D at baseline 6, 13, 26, 39 and 52 weeks. Dutch tariff. **Cost sources:** Resource use from patient questionnaires, clinical study records (for surgery), from physiotherapist for intervention group. Unit costs from Dutch reference unit prices for healthcare provider visits, inpatient days, lab services and home care. Cost of surgery from micro-costing study across Europe 2008. Medical imaging services used fees issued by Dutch Healthcare Authority. Wholesale prices for medications and appliances.				

Comments
Source of funding: Netherlands Organisation for Health Research and Development **Limitations:** Dutch healthcare perspective may not reflect current UK NHS context. Study does not include all exercise treatment options. Follow-up may not be sufficient to capture all benefits and costs. Within-trial analysis and so does not reflect full body available evidence for this comparison. No analysis of uncertainty. **Other:**
Overall applicability: Partially applicable(c) **Overall quality:** Potentially serious limitations(d)

Abbreviations: 95% CI= 95% confidence interval; CUA= cost–utility analysis; EQ-5D= Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); ICER= incremental cost-effectiveness ratio; n/a= not applicable; NR= not reported; QALYs= quality-adjusted life years

(a) For studies where the time horizon is longer than the treatment duration, an assumption needs to be made about the continuation of the study effect. For example, does a difference in utility between groups during treatment continue beyond the end of treatment and if so for how long.

(b) Converted using 2011 purchasing power parities[334]

(c) Directly applicable / Partially applicable / Not applicable

(d) Minor limitations / Potentially serious limitations / Very serious limitations

Appendix I – Health economic model

No original economic modelling was undertaken.

Appendix J – Excluded studies

Clinical studies

Table 82: Studies excluded from the clinical review

Study	Exclusion reason
Aaboe 2014[1]	Abstract only
Abbott 2019[3]	No usable outcomes (health economic outcomes only)
Aebischer 2016[4]	Systematic review; references checked (inadequate quality assessment)
Ahern 2018[7]	Incorrect interventions (compares multi or unimodal physical therapies to usual care, placebo or sham, which is not an appropriate grouping of interventions for this review)
Alayli 2007[8]	Non-English language studies
Alghadir 2019[9]	Inappropriate comparison (compares retro walking to forward walking to standard care, when standard care includes an exercise intervention. Therefore, the effect of exercise could not be separated).
Alkatan 2016[10]	Incorrect interventions (compares cycling and supervised exercises to swimming and supervised exercises, which would both be classified as mixed modality exercise in this protocol)
Allegrante 1991[11]	Not available
Allen 2016[15]	Incorrect interventions (compares a treatment package to usual care, which is considered in a separate review)
Allen 2016[13]	Inappropriate comparison (compares group therapy to individual therapy)
Allen 2021[14]	Incorrect interventions (includes treatment packages and is included in another review)
Alonso-rodriguez 2021[16]	Non-English language studies
Alrushud 2017[17]	Incorrect interventions (compares a treatment package to usual care, which is considered in a separate review)
An 2013[19]	Incorrect study design (non-randomised study)
Anon 2019[120]	Non-English language studies
Ansanay 2020[21]	No usable outcomes (no relevant outcomes reported)
Anwer 2016[23]	Systematic review; references checked (inadequate quality assessment)
Aoki 2009[24]	No usable outcomes (reports biomechanical outcomes only)
Apparao 2017[25]	Inappropriate comparison (compares two different forms of supervised other exercise)
Armagan 2015[26]	No usable outcomes (reports continuous outcomes as median values rather than means, which could not be used in the analysis)
Arnold 2010[27]	No usable outcomes (did not report any appropriate outcomes)
Ashworth 2005[28]	Systematic review is not relevant to review question or unclear PICO (people without osteoarthritis)
Azizi 2020[30]	No usable outcomes (no relevant outcomes reported)
Baker 2001[31]	Inappropriate comparison (compares a treatment package to a component of the programme, which is considered in a separate review)
Bartels 2016[32]	Cochrane review; references checked (does not included outcomes by the definitions used in this review)

Bartholdy 2017[33]	Systematic review is not relevant to review question or unclear PICO (includes people who are post-operative after knee replacement surgery)
Bearne 2011[36]	Incorrect interventions (included a treatment package, which is considered in a separate review)
Beasley 2019[37]	Systematic review; references checked (inadequate quality assessment)
Benli kucuk 2018[39]	No usable outcomes (reports continuous outcomes as median values rather than means, which could not be used in the analysis)
Bennell 2005[43]	Incorrect interventions (included a treatment package, which is considered in a separate review)
Bennell 2012[42]	Inappropriate comparison (compares a treatment package to a component of the programme, which is considered in a separate review)
Bennell 2015[49]	Inappropriate comparison (compares a treatment package to a component of the programme, which is considered in a separate review)
Bennell 2016[40]	Inappropriate comparison (compares a treatment package to a component of the programme, which is considered in a separate review)
Bennell 2017[41]	Inappropriate comparison (compares a treatment package to a component of the programme, which is considered in a separate review)
Bennell 2018[48]	Incorrect interventions (included a treatment package, which is considered in a separate review)
Bennell 2020[47]	Inappropriate comparison (education plus exercise versus short wave diathermy)
Beydagi 2021[50]	Systematic review; references checked
Bezalel 2010[51]	Inappropriate comparison (compares education and exercise to electrotherapy)
Bilgici 2005[53]	Not available
Boeer 2010[54]	Not guideline condition (includes people with hip prosthesis and people with hip osteoarthritis)
Bokaeian 2016[55]	Inappropriate comparison (compares whole body vibration and strength exercise to strength exercise only)
Bove 2018[60]	Inappropriate comparison (compares to manual therapy, which is considered in a different review)
Braghin 2018[61]	Inappropriate comparison (compares people with symptomatic osteoarthritis to asymptomatic osteoarthritis)
Brandao 2021[62]	Not review population (not osteoarthritis)
Bressel 2014[63]	Crossover study
Bricca 2018[64]	Systematic review is not relevant to review question or unclear PICO (qualitative evaluation only)
Brismee 2007[65]	Inappropriate comparison (compares exercise to an education programme)
Bryk 2016[68]	Inappropriate comparison (compares supervised mixed modality exercise to a different type of supervised mixed modality exercise)
Burrows 2014[69]	Not guideline condition (includes healthy participants). Inappropriate comparison (compares repeated exercise to one episode of exercise). Crossover study.
Cadmus 2010[70]	No usable outcomes (reports beta coefficients only)

Callaghan 1995[71]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Carlson 2011[73]	Abstract only
Carmona-teres 2015[74]	Protocol only
Casilda-lopez 2017[75]	Inappropriate comparison (compares supervised mixed modality exercise to a different type of supervised mixed modality exercise)
Ceballos-laita 2019[76]	Systematic review; references checked (inadequate quality assessment)
Chamberlain 1982[78]	Inappropriate comparison (compares electrotherapy to exercise)
Chen 2021[82]	No usable outcomes (no relevant outcomes reported)
Cheung 2018[86]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Cho 2015[87]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Chopp-hurley 2017[88]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Coleman 2012[92]	Incorrect interventions (compares a treatment package to a component of the programme, which is considered in a separate review)
Cotofana 2010[93]	No usable outcomes (reported imaging outcomes only)
Coudeyre 2016[94]	Systematic review; references checked (inadequate quality assessment)
Cuperus 2015[95]	Incorrect interventions. Inappropriate comparison (compared face-to-face treatment to telephone-based treatment)
Da silva 2015[97]	Inappropriate comparison (compared exercise to no treatment where both arms get education. However, the education is different for each group).
Danazumi 2021[98]	Inappropriate comparison (combined chain exercises plus kinesiotaping versus combined chair exercises)
Davenport 2012[99]	Inappropriate comparison (compares two different forms of sunsupervised strength exercise)
De vos 2017[102]	Incorrect interventions (includes a treatment package, which is considered in a separate review)
Deepeshwar 2018[103]	No usable outcomes (included outcomes that are not specified in the protocol)
Dias 2003[105]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Dias 2017[104]	Incorrect interventions (includes a treatment package, which is considered in a separate review)
Doi 2008[108]	Incorrect interventions (included NSAIDs which were not licensed for use in the United Kingdom)
Dong 2018[109]	Systematic review; references checked (inadequate quality assessment)
Dong 2019[110]	Systematic review is not relevant to review question or unclear PICO (included people with a range of conditions, not just osteoarthritis)
Durmus 2012[113]	Inappropriate comparison (compares to a treatment package, which is considered in a separate review)
Durmus 2013[112]	Inappropriate comparison (compares glucosamine and exercise to exercise alone)

Ettinger 1997[121]	No usable outcomes (reported unvalidated scales for outcomes or biomechanical outcomes)
Farr 2010[123]	Inappropriate comparison (compares to a treatment package, which is considered in a separate review)
Fernandes 2010[124]	Incorrect interventions (compares to a treatment package, which is considered in a separate review)
Fernandopulle 2017[125]	Systematic review; references checked (inadequate quality assessment)
Ferreira 2015[126]	Systematic review; references checked (inadequate quality assessment)
Fisken 2015[127]	Inappropriate comparison (compares supervised other exercise to a different supervised other exercise)
Fitzgerald 2016[128]	Inappropriate comparison (compares to manual therapy, which is considered in a different review)
Focht 2004[132]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Focht 2005[133]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Focht 2014[130]	Incorrect interventions (includes a treatment package, which is considered in a separate review)
Focht 2017[131]	Incorrect interventions (includes a treatment package, which is considered in a separate review)
Foroughi 2011[135]	Inappropriate comparison (compares to sham exercise)
Foroughi 2011[136]	Inappropriate comparison (compares to sham exercise)
Foster 2014[137]	Protocol only
Fransen 2001[138]	No usable outcomes (reports outcomes with groups merged together, which is not appropriate for this review)
Fransen 2003[139]	Cochrane review; references checked (does not include the separation of interventions needed for this review)
Fransen 2006[144]	Abstracts
Fransen 2010[142]	Systematic review; references checked (inadequate quality assessment)
Fransen 2014[143]	Cochrane review; references checked (includes a different definition of outcomes that are included in this review)
Fransen 2015[140]	Cochrane review; references checked (does not include the separation of interventions needed for this review)
Fransen 2015[141]	Cochrane review; references checked (does not include the separation of interventions needed for this review)
French 2015[148]	Protocol only
Fukumoto 2014[149]	Inappropriate comparison (compares high intensity exercise to low intensity exercise)
Fukumoto 2017[150]	Inappropriate comparison (compares high intensity exercise to lower intensity exercise)
Garfinkel 1994[151]	Inadequate randomisation with characteristics of participants making up results being unclear
Ghroubi 2008[152]	Not available
Goh 2019[155]	Systematic review; references checked (inadequate quality assessment)
Goh 2019[154]	Incorrect interventions (Exercise without any additional treatments versus usual care)

Goksen 2021[156]	No usable outcomes (reported only medians and interquartile ranges)
Goonasegaran 2022[159]	Inappropriate comparison (retro walking versus forward walking)
Green 1988[160]	Abstracts
Green 1993[161]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Gudbergsen 2012[162]	Incorrect interventions (includes dietary interventions)
Gur 2002[163]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Halbert 2001[164]	Posthoc analysis (is a substudy of an existing study where it is unclear if randomisation is maintained)
Hale 2012[165]	Inappropriate comparison (compares exercise to a social computer activity)
Hall 2019[166]	Systematic review is not relevant to review question or unclear PICO (compares diet induced weight loss and exercise to diet induced weight loss alone)
Hanada 2018[167]	Less than minimum duration (<1 week)
Handa 2000[168]	Not guideline condition (compares people with low back pain to healthy volunteers)
Harris-hayes 2021[169]	Not review population (hip groin pain, not specified to be osteoarthritis)
Hartman 2000[170]	Spinal osteoarthritis
Hasegawa 2013[171]	Incorrect study design (non-randomised study)
Hay 2006[172]	Incorrect interventions (includes a treatment package, which is considered in a different review)
Henriksen 2015[175]	Systematic review; references checked (inadequate quality assessment)
Henriksen 2016[174]	Systematic review; references checked (inadequate quality assessment)
Henriksen 2016[176]	No usable outcomes (reported radiological outcomes only)
Hiyama 2012[181]	No usable outcomes (Reported biomechanical outcomes only)
Horstmann 2000[186]	Not available
Howe 2016[187]	Systematic review is not relevant to review question or unclear PICO (narrative review only)
Hu 2020[189]	Inappropriate comparison
Hu 2021[188]	Systematic review; references checked
Huang 2018[190]	Inappropriate comparison (compares exercise and pharmacological interventions to exercise)
Hughes 2004[194]	Incorrect interventions (includes a treatment package, which is considered in a different review)
Hughes 2006[195]	Incorrect interventions (includes a treatment package, which is considered in a different review)
Hunt 2013[197]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Hunter 2015[198]	Inappropriate comparison (compares a treatment package to individual components, which is considered in a different review)
Hurley 2007[199]	Incorrect interventions (includes a treatment package, which is considered in a different review)

Husby 2009[201]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise). Not review population (post-operative)
Imoto 2019[202]	Systematic review; references checked (inadequate quality assessment)
Isaramalai 2018[204]	Inappropriate comparison (compares a treatment package to a more intensive treatment package)
Jan 2008[207]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Jan 2009[205]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Jansen 2011[208]	Systematic review; references checked (inadequate quality assurance)
Jegu 2014[209]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Jenkinson 2009[210]	Wrong intervention (included a treatment package, which is considered in a different review)
Jeong 2019[211]	Systematic review; references checked (inadequate quality assurance)
Jigami 2012[212]	Inappropriate comparison (compared weekly exercise to fortnightly exercise)
Jordan 2010[213]	Systematic review is not relevant to review question or unclear PICO (studies interventions to improve exercise adherence, includes people without osteoarthritis)
Juhl 2014[217]	Systematic review; references checked (inadequate quality assurance)
Kabiri 2018[218]	Inappropriate comparison (compares mixed modality exercise to mixed modality exercise)
Kamalakannan 2019[219]	Inappropriate comparison (proprioception training (supervised other exercise) and conventional exercise versus interferential therapy)
Kan 2016[220]	Systematic review: study designs inappropriate (includes non-randomised studies)
Kars 2019[226]	Test paper only
Kelley 2016[229]	Systematic review; references checked (inadequate quality assessment)
Kelley 2018[230]	Systematic review is not relevant to review question or unclear PICO (mixed population including people without osteoarthritis)
Keogh 2018[231]	Inappropriate comparison (compares high intensity to low intensity exercise)
Keshtkaran 2010[232]	Non-English language studies
Kim 2012[235]	Incorrect study design (non-randomised study)
Kloek 2018[236]	Incorrect interventions (includes a treatment package, which will be considered in a different review)
Konishi 2009[240]	Incorrect study design (non-randomised study)
Kovar 1992[241]	Incorrect interventions (includes a treatment package, which will be considered in a different review)
Krauss 2016[243]	Protocol only
Kreindler 1989[245]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)

Kroon 2018[246]	Systematic review; references checked (inadequate quality assessment)
Kudo 2013[247]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Lai 2018[252]	Inappropriate comparison (compares exercise to an education program alone)
Lange 2009[253]	Inappropriate comparison (compares to sham exercise)
Lee 2006[257]	Non-English language studies
Lee 2008[258]	Not available
Lee 2018[254]	Inappropriate comparison(Tai Chi versus Physical Therapy exercise program)
Lee 2019[255]	Not available
Li 2015[260]	Systematic review; references checked (inadequate quality assessment)
Li 2016[261]	Systematic review; references checked (inadequate quality assessment)
Li 2017[259]	Incorrect interventions (includes a treatment package, which will be considered in a different review)
Liao 2013[262]	Not review population (included people post-operative for joint replacement surgery)
Liebs 2012[263]	Not review population (included people post-operative for joint replacement surgery). Inappropriate comparison (compared early aquatic therapy to late aquatic therapy).
Lin 2007[267]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Lin 2020[269]	Inappropriate comparison (both were mixed modality exercise interventions)
Liu 2009[270]	Not guideline condition (included a mixed population)
Loew 2017[271]	Incorrect interventions (included people assigned to a walking group of their choice or a group they didn't choose)
Lorenc 2018[272]	Systematic review is not relevant to review question or unclear PICO (includes a mixed population)
Lu 2015[274]	Not available
Lu 2017[273]	Inappropriate comparison (compares exercise to an education program)
Lue 2017[275]	Systematic review; references checked (inadequate quality assessment)
Lun 2015[276]	Inappropriate comparison (compares a hip strengthening exercise program to a leg strengthening exercise program)
Lund 2008[277]	Inappropriate comparison (compares an aquatic exercise to a land-based exercise of the same type)
Magni 2017[278]	Systematic review; references checked (inadequate quality assessment)
Mangani 2006[279]	Inappropriate comparison (compares exercise to an education programme)
Mangione 1996[280]	Inappropriate comparison (compares treadmill exercise to different degrees of weight unloading while doing treadmill exercise)
Mangione 1999[281]	Inappropriate comparison (compares high intensity ergometry to low intensity ergometry)
Mat 2015[282]	Systematic review; references checked (inadequate quality assessment)

Mattos 2016[283]	Systematic review: study designs inappropriate (included non-randomised trials)
Maurer 1999[284]	Incorrect interventions (includes an education programme)
Mazloum 2018[285]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Mccaffrey 2017[286]	Incorrect interventions (includes an education programme)
Mccarthy 2004[288]	Inappropriate comparison (compares an exercise program to another exercise program of the same type)
Mcknight 2010[290]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Mcveigh 2021[291]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Messier 1997[297]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Messier 2000[296]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Messier 2004[292]	Inappropriate comparison (included a treatment package, which will be considered in a different review)
Messier 2013[294]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Messier 2018[295]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Messier 2021[293]	Inappropriate comparison (high intensity strength training-supervised strength exercise versus low intensity strength training-supervised strength exercise), attention control (workshops - beyond usual care)
Mihalko 2019[298]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Mikesky 2006[299]	Not review population (included a population with greater than 20% not having osteoarthritis)
Mikkelsen 2014[300]	Not review population (included a post-surgical population)
Miller 2003[301]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Minor 1989[303]	People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, diseases of childhood that may predispose to osteoarthritis, medical conditions presenting with joint inflammation and malignancy).
Minor 1993[302]	People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, diseases of childhood that may predispose to osteoarthritis, medical conditions presenting with joint inflammation and malignancy).
Minshull 2017[304]	Systematic review; references checked (inadequate quality assurance)
Monticone 2016[305]	Systematic review; references checked (inadequate quality assurance)
Moonaz 2015[306]	People with conditions that may make them susceptible to osteoarthritis or often occur alongside osteoarthritis (including: crystal arthritis, inflammatory arthritis, septic arthritis, diseases of childhood that may predispose to osteoarthritis, medical conditions presenting with joint inflammation and malignancy).

Moreira 2021[307]	Inappropriate comparison weight-bearing versus nonweight-bearing exercise)
Moseng 2017[308]	Systematic review; references checked (inadequate quality assurance)
Munukka 2020[309]	No usable outcomes (outcomes reported in graphical form only)
Murphy 2010[311]	Inappropriate comparison (compares a tailored pacing activity to a standard pacing activity)
Myers 1998[312]	Commentary only
Na 2000[313]	Not available
Nathani 2020[316]	No usable outcomes (reported aggregate WOMAC score or biomechanical outcomes only)
Nelligan 2021[321]	Conference abstract only
Neelapala 2020[318]	Systematic review; references checked (inadequate quality assurance)
Neelapala 2018[319]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Nery 2015[324]	Abstracts
Nery 2016[323]	Abstracts
Ng 2010[325]	Inappropriate comparison (compares a 3 day walking group to a 5 day walking group)
Ni 2010[326]	Inappropriate comparison (compares supervised other exercise to an education program)
Nicklas 2004[327]	Not guideline condition (included obese adults, not necessarily with osteoarthritis)
Oiestad 2013[329]	Protocol only
Olagbegi 2016[331]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Osborne 2006[335]	Incorrect interventions (included a treatment package, which will be considered in a different review). Protocol only
Osteras 2017[336]	Cochrane review; references checked (included a different definition of exercise to those used in this review)
Østerås 2017[338]	Cochrane review; references checked (included a different definition of exercise to those used in this review)
Osteras 2017[339]	Cochrane review; references checked (included a different definition of exercise to those used in this review)
Osugi 2014[340]	Spinal osteoarthritis
Ozdincler 2005[341]	Not available
Ozturk 2021[342]	Inappropriate comparison (observation plus exercise versus exercise)
Park 2011[343]	Incorrect study design (non-randomised study)
Park 2014[344]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Park 2016[346]	Inappropriate comparison (compared exercise to a health education program)
Park 2017[345]	Inappropriate comparison (compared exercise to a health education program)
Penninx 2001[351]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Perez-huerta 2020[352]	Inappropriate comparison(sitting aerobic exercises versus standing aerobic exercises)

Perez-marmol 2017[353]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Petersen 2011[354]	Inappropriate comparison (compares exercise and pharmacological treatment to exercise)
Pisters 2010[358]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Pisters 2010[359]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Piyakhachornrot 2011[360]	Inappropriate comparison (compares a supervised facility based exercise to a supervised home based exercise)
Praharsitha 2019[361]	Inappropriate comparison(lateral versus medial hamstring exercises)
Qi 2020[362]	Not review population(older adults, including those without OA)
Quicke 2015[363]	Systematic review; references checked (inadequate quality assessment)
Quicke 2020[364]	Systematic review; references checked (inadequate quality assessment, reported subgrouping of populations in studies which did not relate to those stated in the protocol for this review)
Quilty 2003[365]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Raj 2019[366]	No usable outcomes (no relevant outcomes)
Rao 1998[367]	Commentary only
Rashid 2019[368]	No usable outcomes (provides WOMAC as median IQR, doesn't provide subscale values anyway)
Regnaux 2015[370]	Systematic review is not relevant to review question or unclear PICO (compares high intensity and low intensity exercise)
Rejeski 1997[371]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Rejeski 1998[372]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Rejeski 2002[373]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Rewald 2016[375]	Protocol only
Rodriguez-merchan 2016[378]	Systematic review; references checked (inadequate quality assurance)
Rogers 2009[381]	Crossover study. Inappropriate comparison (compares exercise to sham exercise)
Roper 2013[383]	Inappropriate comparison (compares aquatic exercise to land-based exercise of the same type)
Runhaar 2016[385]	Not review population (aiming to prevent osteoarthritis in people without the condition)
Sampath 2016[389]	Systematic review; references checked (inadequate quality assurance)
Sashika 1996[391]	Not review population (post-joint replacement surgery)
Saw 2016[392]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Schencking 2013[394]	Inappropriate comparison (compares hydrotherapy to land-based therapy of the same type)
Schepens 2012[395]	Inappropriate comparison (compares tailored pacing activity to a general pacing activity)

Schilke 1996[396]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Schlenk 2011[398]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Schlenk 2020[397]	Inappropriate comparison (exercise and telephone sessions compared to telephone sessions of a different type)
Schmid 2013[399]	No usable outcomes
Seidler 2018[402]	Systematic review: study designs inappropriate (included non-randomised studies)
Sevick 2000[404]	Protocol only
Sharma 2018[406]	Inappropriate comparison (compared exercise, meditation and exercise to education only)
Shen 2021[407]	Wrong comparison (comparing exercise to a health lecture series, the latter likely being more intense than usual care/no treatment defined in the protocol)
Simao 2012[409]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Sled 2010[410]	Not review population (compared people with osteoarthritis to healthy participants)
Smith 2009[411]	Protocol only
Smith 2012[412]	Systematic review; references checked (inadequate quality assurance)
Somers 2012[413]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Song 2010[416]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Sorour 2014[417]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Srikesavan 2016[418]	Inappropriate comparison (compared unsupervised strength exercise to unsupervised strength exercise)
Stamm 2002[419]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Steinhilber 2017[420]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Stener-victorin 2004[421]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Stensrud 2015[422]	Inappropriate comparison (compares exercise to arthroscopic surgery)
Stoffer-marx 2018[423]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Sullivan 1998[424]	Inappropriate comparison (included a treatment package, which will be considered in a different review)
Suzuki 2019[425]	Inappropriate comparison (compares unsupervised strength exercise to unsupervised strength exercise)
Svege 2015[427]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Svege 2016[426]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Taglietti 2018[428]	Inappropriate comparison (compares exercise to an education programme)

Tak 2005[429]	Inappropriate comparison (compares a treatment package, which will be considered in a different review)
Talbot 2003[431]	Incorrect interventions (included a treatment package, which will be considered in a different review)
Tamin 2018[433]	Inappropriate comparison (compares supervised mixed modality exercise to supervised mixed modality exercise)
Tamin 2018[432]	Abstract only
Tan 2016[434]	Cost-utility analysis only
Tanaka 2013[438]	Inappropriate comparison (systematic review including comparisons of exercise to other exercise of the same type and to psychoeducational interventions)
Tanaka 2014[436]	Systematic review; references checked (inadequate quality assessment)
Tanaka 2015[435]	Systematic review; references checked (inadequate quality assessment)
Tanaka 2016[437]	Systematic review; references checked (inadequate quality assessment)
Thomas 2002[442]	No usable outcomes (includes six different groups that are not well defined and have some overlap when results are reported)
Thomas 2005[441]	Cost-effectiveness analysis only
Thompson 2020[443]	No usable outcomes (outcomes were reported as medians and interquartile ranges only)
Topp 2002[445]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Tossige-gomes 2012[446]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Trans 2009[447]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)
Tsai 2013[448]	Inappropriate comparison (compares supervised other exercise to an education programme)
Tsai 2015[449]	Inappropriate comparison (compares supervised other exercise to an education programme)
Tsauo 2008[450]	Inappropriate comparison (compares supervised mixed modality exercise to supervised mixed modality exercise)
Tuzun 2004[452]	Inappropriate comparison (compares supervised strength exercise and supervised strength exercise)
Unsal 2008[454]	Incorrect interventions (compares exercise and intraarticular injections to exercise and physical therapy modalities)
Uthman 2013[455]	Systematic review; references checked (inadequate quality assurance)
Uzunkulaoglu 2019[457]	Inappropriate comparison (compares supervised other exercise to supervised other exercise)
UzunkulaoGlu 2020[456]	Inappropriate comparison (compares supervised other exercise to supervised other exercise)
Van baar 1998[460]	Systematic review; references checked (inadequate quality assurance)
Van Ginckel 2019[461]	Systematic review is not relevant to review question or unclear PICO (studies imaging changes)
Van Gool 2005[462]	No usable outcomes (outcomes were reported in a form that could not be used in a meta-analysis)

Veenhof 2006[463]	Inappropriate comparison (compares behavioural graded activity to usual care)
Villadsen 2014[464]	Not review population (post-surgical population)
Vincent 2019[465]	No usable outcomes (outcomes reported in graphical form only)
Vincent 2020[466]	No usable outcomes (relevant outcomes reported in graphical format only)
Waller 2014[468]	Systematic review; references checked (inadequate quality assurance)
Waller 2017[467]	Incorrect interventions (intensive aquatic resistance training versus normal physical activity)
Wallis 2017[469]	Incorrect interventions (included a treatment package, which is considered in a different review)
Wang 2008[472]	Inappropriate comparison (compares supervised other exercise to stretching)
Wang 2009[471]	Inappropriate comparison (compares supervised other exercise to stretching)
Wang 2014[470]	Inappropriate comparison (compares whole body vibration and strength exercise to strength exercise)
Wang 2015[477]	Systematic review; references checked (inadequate quality assurance)
Wang 2016[473]	Inappropriate comparison (compares whole body vibration and strength exercise to strength exercise)
Wang 2016[476]	Inappropriate comparison (compares whole body vibration and strength exercise to strength exercise)
Wang 2016[475]	Inappropriate comparison (compares supervised other exercise to manual therapy and mixed modality exercise)
Wang 2018[474]	Inappropriate comparison (compares supervised strength exercise to supervised strength exercise)
Wang 2018[480]	Systematic review is not relevant to review question or unclear PICO (includes people with rheumatoid arthritis)
Watanabe 2013[481]	Inappropriate comparison (compares body-weight supported treadmill training to full body-weight treadmill training)
Weng 2009[482]	Wrong unit of randomisation (knee)
Wetzels 2005[483]	Not available
Wetzels 2008[484]	Incorrect interventions (includes a treatment programme, which is considered in a different review)
Williamson 2015[486]	Systematic review is not relevant to review question or unclear PICO (includes behavioural physical activity interventions compared to sham/no treatment)
Witteveen 2013[487]	Cochrane review; references checked (includes any intervention for use in ankle osteoarthritis)
Wood 2016[488]	Incorrect study design (non-randomised study)
Wyatt 2001[490]	Inappropriate comparison (compares supervised aquatic exercise to supervised land based exercise of the same type)
Xu 2021[493]	Not review population (post total knee arthroplasty)
Yazigi 2013[494]	Protocol only
Ye 2014[495]	Systematic review; references checked (inadequate quality assurance)
You 2021[499]	Systematic review; references checked (inadequate quality assurance)

Zafar 2015[500]	Systematic review; references checked (inadequate quality assurance)
Zammit 2010[501]	Cochrane review; references checked (includes any intervention for use in toe osteoarthritis)
Zgibor 2017[502]	Inappropriate comparison (compares a treatment package to an individual component, which will be considered in another review)
Zhang 2017[503]	Systematic review; references checked (inadequate quality assurance)
Zhang 2020[504]	Inappropriate comparison (Tai Chi (supervised other exercise versus wellness education program (probably more intense than usual care))
Zhu 2016[505]	Inappropriate comparison (compares other supervised exercise to education sessions)

Health Economic studies

Published health economic studies that met the inclusion criteria (relevant population, comparators, economic study design, published 2005 or later and not from non-OECD country or USA) but that were excluded following appraisal of applicability and methodological quality are listed below. See the health economic protocol for more details.

Table 83: Studies excluded from the health economic review

Reference	Reason for exclusion
Jukahowski 2011[216]	This Finnish cost-consequences analysis was selectively excluded as there are UK-based cost utility analyses included.

Appendix K – Research recommendations – full details

K.1.1 Research recommendation

What is the clinical and cost effectiveness of supervised group and individual exercise compared with unsupervised exercise for people with osteoarthritis?

K.1.2 Why this is important

The evidence included in this review showed that exercise was a clinically effective treatment that could be cost effective. However, there was limited evidence to determine if supervised interventions were cost effective compared to unsupervised interventions. The current recommendation states that supervised exercise should be considered. If evidence is found showing that supervised exercise is more effective than unsupervised exercise then this could lead to strengthening of that recommendation and more use of supervised exercise in the future. Additionally, the committee acknowledged the need to find innovative ways to deliver exercise interventions that are widely accessible, inclusive to diverse populations and cost effective during the COVID 19 pandemic and beyond.

K.1.3 Rationale for research recommendation

Importance to 'patients' or the population	Exercise has been shown to be a clinically effective treatment for people with osteoarthritis. However, the optimum way(s) that exercise should be delivered in unclear. While there is some evidence comparing the clinical effectiveness of supervised and unsupervised exercise that showed that supervised exercise was generally more clinically effective, there is no cost-effectiveness evidence to support this . If evidence is available that indicates more benefit from supervised exercise that is also cost effective then this may influence commissioning of services and may result in increased access to therapy for people with osteoarthritis. Currently, including in the context of the COVID 19 pandemic, exercise interventions may not be accessible to all. Therefore, investigating the method of delivering exercise treatment to ensure that it is widely accessible would be important to people with osteoarthritis.
Relevance to NICE guidance	The current recommendation regarding supervised exercise recommends that supervised exercise should be considered for people with osteoarthritis, while any type of exercise should be offered. If additional evidence shows the clinical and cost-effectiveness of supervised exercise compared to unsupervised exercise then this could help to strengthen this recommendation or give more certainty to the current wording.
Relevance to the NHS	Supervised exercise is likely to incur an additional cost compared to unsupervised

	exercise. Given this, stronger evidence investigating the effectiveness of different modes of delivery of supervised exercise when compared to unsupervised exercise is required to show that there is a significant benefit from this.
National priorities	Investigating different modes of delivery is a national priority area discussed in the NHS Long term plan (digitally enabled care and community supported care).
Current evidence base	Current evidence shows that exercise whether supervised or unsupervised can lead to clinically important benefits in quality of life, pain and physical function. There is inconsistent evidence investigating the differences between supervised and unsupervised exercise, but in general this shows clinically important benefits from supervised exercise. Currently there is limited cost-effectiveness evidence investigating exercise compared with usual care for people with osteoarthritis, with no cost effectiveness evidence comparing supervised and unsupervised exercise.
Equality considerations	Some people may benefit more from supervised exercise due to difficulties in completing exercise interventions without additional support (for example: people with comorbidities, people with learning disability). The committee noted that the research identified does not appear to represent the diverse population of people with osteoarthritis. They agreed that any further research should be representative of the population, including people from different family backgrounds, and socioeconomic backgrounds, disabled people, and people of different ages and genders. Future work should be done to consider the different experiences of people from diverse communities to ensure that the approach taken can be made equitable for everyone.

K.1.4 Modified PICO table

Population	Inclusion: • Adults (age ≥16 years) with osteoarthritis affecting any joint
Intervention	• Supervised group-based exercise programme • Supervised individual exercise programme (delivered in person) • Supervised individual exercise programme (delivered with an alternative method of delivery, including telehealth) • Unsupervised exercise programme

Comparator	• Each other
Outcome	Stratify by ≤/>3 months (longest time-point in each): • Health-related quality of life [validated patient-reported outcomes, continuous data prioritised] • Pain [validated patient-reported outcomes, continuous data prioritised] • Physical function [validated patient-reported outcomes, continuous data prioritised] • Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flares [dichotomous] • Serious adverse events [dichotomous]
Study design	Randomised controlled trial
Timeframe	Long term (at least 1 year)
Additional information	Adequately powered high quality randomised controlled trials. Trials with sufficient blinding, adequate randomisation methods and allocation concealment. Subgroups: • Joint site(s) of osteoarthritis • Age (≤/> 75 years) • Multimorbidity (high versus low morbidity score; as defined by study, measured by validated instruments e.g. Charlson Comorbidity Index • People with learning disability