National Institute for Health and Care Excellence

Final

Osteoarthritis in over 16s: diagnosis and management

[F] Evidence review for the clinical and cost-effectiveness of acupuncture for people with osteoarthritis

NICE guideline NG226

Evidence reviews underpinning recommendation 1.3.8 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 All NICE guidance is subject to regular review and may be updated or withdrawn.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 Acupuncture

1.1 Review question

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

1.1.1 Introduction

Acupuncture involves treatment with the careful placement and manipulation (manually or using electricity) of needles to promote pain relief. Acupuncture has been used for the management of pain in a range of conditions including osteoarthritis. Currently, acupuncture continues to be used with people with osteoarthritis in some settings. However, this was not recommended in the NICE Osteoarthritis guideline CG177 as there was insufficient evidence to support its use. Since the guideline was published new evidence has emerged that may demonstrate it is an effective intervention. It is important to consider the effectiveness of acupuncture as a non-pharmacological treatment modalities in the context of a long-term condition such as osteoarthritis, particularly when pharmacological options are limited and potentially harmful.

This review aims to evaluate the effectiveness of acupuncture (including conventional acupuncture, dry needling and electroacupuncture) for the management of 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, diseases of childhood that may predispose to osteoarthritis, medical conditions presenting with joint inflammation and malignancy). • Studies in people with meniscal injury without osteoarthritis • Studies with an unclear population (e,g, type of arthritis, proportion of participants with osteoarthritis) Spinal osteoarthritis
Interventions	• Acupuncture/dry needling • Electroacupuncture
Comparisons	• Compared to each other • Sham acupuncture (in this report, sham acupuncture and sham electroacupuncture are both referred to as sham acupuncture) • No intervention (including either): ○ Acupuncture versus no treatment* ○ Acupuncture 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.*
Outcomes	Stratify by ≤/>3 months (longest time-point in each): Primary outcomes (critical outcomes): • 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] Secondary outcomes (important outcomes); • Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flare-ups [validated patient-reported outcomes, continuous data prioritised] • Serious adverse events [dichotomous data]
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

Thirty-six randomised controlled trial studies (forty-six papers) were included in the review;[9, 10, 16, 17, 19, 26, 30, 34, 37, 47, 54, 61, 63, 64, 67, 73, 82, 92, 95, 96, 110-113, 123, 127, 135, 139, 146, 147, 149, 151, 154-156, 162] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3).

Studies included the following comparisons (some studies reported multiple comparisons):

- Acupuncture compared to sham acupuncture[16, 17, 19, 30, 34, 37, 63, 67, 73, 95, 96, 110, 111, 113, 127, 135, 151, 155]
- Acupuncture compared to no treatment[16, 37, 47, 61, 64, 154-156, 162]
  - Acupuncture compared to no treatment*[47, 154-156]
  - Acupuncture plus additional treatment compared to additional treatment alone**[16, 37, 61, 64, 162]
- Electroacupuncture compared to acupuncture[135, 146, 162]
- Electroacupuncture compared to sham acupuncture[9, 54, 82, 92, 112, 123, 135, 139, 147, 149]
- Electroacupuncture compared to no treatment[10, 26, 92, 123, 162]
  - Electroacupuncture compared to no treatment*[123]
  - Electroacupuncture plus additional treatment compared to additional treatment alone**[10, 26, 92]

* *No treatment defined as either (1) doing nothing or (2) very low intensity intervention such as advice.*

*** Studies where acupuncture/electroacupuncture plus additional treatments was compared to the same additional treatments but without acupuncture/electroacupuncture were included. This includes studies that discuss "usual care". The committee agreed that there is no "usual" care consistently provided by professionals for people with osteoarthritis, therefore agreed to avoid this terminology in this guideline.*

Laser acupuncture is not considered in this review as the committee agreed that it was a form of laser therapy that was considered in the electrotherapy review. For more information see Evidence review G: Electrotherapy.

A network meta-analysis was not conducted for this review. This was decided as sham acupuncture would not be given as a treatment in standard clinical practice making their use for recommendations more limited. Therefore, the committee agreed that the additional benefit of a network meta-analysis would be limited.

See also the study selection flow chart in Appendix C: Exercise, study evidence tables in Appendix D: Weight loss, forest plots in Appendix E: Manual therapy and GRADE tables in Appendix F: Acupuncture.

1.1.4.2 Excluded studies

Cochrane reviews were identified but could not be included due to a different population to that in the protocol (Green 2005[43]), a different comparison to that in the protocol (Manheimer 2018[87]) and different definitions of outcomes (Manheimer 2010[86]). The references were checked any studies that fulfilled the inclusion criteria were included.

See the excluded studies list in Appendix I.

1.1.5 Summary of studies included in the effectiveness evidence

Table 2: Summary of studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Berman 1999[10]	**Electroacupuncture** (n=37) Biweekly for 8 weeks. Acupuncture points based on traditional Chinese medicine theory. 5 local points, 4 distal points. 1 inch, 34 gauge, 0.22mm diameter needles inserted to 0.4-0.6 inches. De qi sensation verified. Electrical stimulation with 2.5-4Hz, square pulses of 1.0ms duration for 20 minutes. Treatment for 8 weeks. **No treatment** (n=36) Conventional therapy **Concomitant therapy:** People were asked to remain on their baseline analgesic/anti-inflammatory regimens as well and not to begin any new physiotherapy or exercise programmes	**Knee osteoarthritis** Mean age (SD): 65.6 (8.6) years N = 73 Definition: Diagnosis of osteoarthritis of the knee (American College of Rheumatology criteria applied) Severity: Kellgren Lawrence grade of 2 or more Duration of symptoms (mean [SD]): 7.2 (6.2) years Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months Serious adverse events at ≤3 months	
Berman 2004[9] Subsidiary paper: Manheimer 2006[88]	**Electroacupuncture** (n=190) 26 weeks of gradually tapering treatment (6 weeks of 2 treatments per week, 2 weeks of 1 treatment per week, 4 weeks of 1 treatment every other week, 12 weeks of 1 treatment per	**Knee osteoarthritis** Mean age (SD): 65.5 (8.6) years N = 570 Definition: A diagnosis of osteoarthritis of the knee with	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
	month). Based on traditional Chinese theory. Uses 5 local points and 4 distal points. 1-1.5 inch, 32 gauge, 0.25mm diameter needles inserted to a conventional depth of approximately 0.3 to 1.0 inch. All people achieved the De qi sensation. Electrical stimulation was applied at knee points Xiyan at low frequency (8Hz), and square biphasic pulses (0.5ms pulse width) for 20 minutes. **Sham acupuncture** (n=191) Needles inserted into sham points in the abdominal are with adhesive tape applied next to the needles. Mock guiding tubes were tapped onto each of the 9 true points used in the intervention group. Electrical stimulation did not take place (although a mock stimulation unit was attached to the sham needles at the knee). A third group was reported (n=189) but was not included as it did not fulfil the inclusion criteria. **Concomitant therapy:** No additional information	radiographic evidence of at least 1 osteophyte at the tibiofemoral joint Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms: ≤5->10 years, median <5 years Presence of multimorbidities: Not stated/Unclear	Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
Ceballos-laita 2019[17] Subsidiary paper: Ceballos-laita 2020[18]	**Dry needling** (n=15) Dry needling was performed by the lead author who had four years of clinical experience. Active MTrPs were located by manual palpation in the hip muscles. They were immobilised between the index and middle finger. Three active MTrPs were treated at most in each session. A standard single-use sterile acupuncture needle (0.25 mm x 50 mm) was inserted perpendicularly through the skin and moved forward until the MTrP was reached. To minimise pain of insertion, a certain pressure was applied to the skin with the insertion tube. Hong's fast-in and fast-out technique was used with the aim of eliciting a local twitch response. After the needle was premoved, pressure with a cotton ball was maintained to prevent bleeding. Patients received three treatment sessions, with one session per week. Duration 3 weeks. **Sham acupuncture** (n=15) Participants received a simulated dry needling technique that has been shown to be valid. The blunted needle was applied to MTrPs to	**Hip osteoarthritis** Mean age (SD): Dry needling group: 55.5 (4.7) years; sham group: 58.6 (6.6) years N = 30 Definition: Unilateral primary hip osteoarthritis according to the clinical criteria of the American College of Rheumatology, a grade II or III Kellgren & Lawrence classification in their most recent hip x-rays, 50-70 years of age, and presence of at least one active MTrP in the hip muscles Severity of symptoms: Grade K-L II: Dry needling group 9/15; sham group 6/15. Grade K-L III: Dry needling group 6/15; sham group 9/15 Duration of symptoms (mean, SD): Dry needling group 64.4 (79.6) months; sham group 72.2 (91.2) months Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Psychological distress at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	provoke a pricking sensation, without penetrating the skin. Sham dry needling was also added in the same regions with the same dose as the dry needling group **Concomitant therapy**: No exercise programme or physical therapy modalities were added to the intervention. Patients were asked not to take any nonsteroidal anti-inflammatory or muscle relaxant drugs.			
Ceballos-laita 2021[16]	**Dry needling** (n=15) Participants received 3 session of dry needling (1 session per week) into active MTrPs in the hip muscles. Iliopsoas, rectus femoris, tensor fasciae latae, and gluteus minimus muscles were examined for the presence of active MTrPs. At most, 3 active MTrPs were treated during each session. **Sham acupuncture** (n=15) Participants received three sessions of a sham needle procedure (one per week). **No intervention** (n=15) Control group participants did not receive any treatment,	**Hip osteoarthritis** Mean age (SD): Dry needling group: 57.53 (3.88) years; sham group 58.20 (5.08) years; control group: 54.67 (4.48) years. N = 45 Definition: Unilateral hip OA according to the American College of Rheumatology criteria, a grade II or III Kellgren & Lawrence classification, age between 50-70 years, and at least 1 active MTrP in the hip muscles. Severity of symptoms (K-L grade II/III): Dry needling	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	education or advice during the study. **Concomitant therapy:** All participants were asked to continue with the same daily routines and not to take any analgesic, anti-inflammatory or muscle relaxant medications 24 hours prior to testing.	group: 7/8; Sham group: 6/9; control group: 6/9 Duration of symptoms (mean, SD): Dry needling group: 66.33 (76.61) months; sham group: 72.20 (53.76) months; control group: 68.13 (56.36) months Presence of multimorbidities: Not stated/Unclear		
Chen 2013[19]	**Acupuncture** (n=105) Acupuncture once or twice a week for 12 weeks (12 treatments in total). 8 gauge, 1.2 inch needles were inserted into 5 local points and 4 distal points (using traditional Chinese medicine theory). The insertion depth was between 0.2-3cm. The needles were left in place for 20 minutes, with a brief manipulation at the beginning and end of the treatment. The de qi sensation was not required and not specifically recorded. **Sham acupuncture** (n=109) Using Streitberger non-penetrating needle (needle that retracts into the handle when it is pressed against the skin). Otherwise same procedure and timing as the acupuncture group.	**Knee osteoarthritis** Mean age (SD): 60.5 (11.4) years N = 214 Definition: People with pain in 1 or both knee joints for more than 6 months with radiological Kellgren-Lawrence grade 2-3 osteoarthritic changes Severity: Kellgren Lawrence grade 2-3, median grade 3 Duration of symptoms (mean [SD]): 9.5 (9.6) years Presence of multimorbidities: Not stated/Unclear	Quality of life at ≤3 months 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
	Concomitant therapy: All people received exercise-based physical therapy once or twice a week for a maximum of 12 total treatments. The standardized program was as vigorous as the person could tolerate with routine encouragement and included range of motion exercises, muscle strengthening and aerobic conditioning.			
Dunning 2018[26]	**Electroacupuncture** (n=105) 9 point protocol for 20 to 30 minutes on each session, for 8-10 sessions. Periosteal dry needling at a frequency of 1-2 times per week over 6 weeks. Needles were of 3 sizes: 0.22mmx30mm, 0.30mmx40mm, and 0.30mmx50mm. The depth ranged from 15 to 45mm. All needles were manipulation to illicit a sensation of aching, tingling, deep pressure, heaviness or warmth. In addition at least 3 needles were trusted and tapped using a “periosteal stimulation” technique. Electrical stimulation used 2Hz, 250microsecond, biphasic continuous waveforms at a maximum tolerable intensity. **No intervention** (n=121)	**Knee osteoarthritis** Mean age (SD): 57.6 (13.2) years N = 242 Definition: People meeting the American College of Rheumatology criteria for the diagnosis of knee osteoarthritis Severity: Not stated Duration of symptoms (mean [SD]): 4.6 (4.9) years Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	No acupuncture. **Concomitant therapy:** Both groups received manual therapy (passive joint mobilizations and muscle stretching) and exercise (riding a stationary bicycle, range of motion, and strengthening exercises to the lower extremity) on each session.			
Farazdaghi 2021[30]	**Dry needling** (n=20) Dry needling involved three repetitive measures at each site of MTrP. At least two hyperalgesic points showing radicular pain, jumping sign or abrupt response were marked for treatment. The dry needling technique consisted of insertion of a disposable 0.25 x 40 mm stainless steel acupuncture needle. Participants received three sessions over one week. **Sham acupuncture** (n=20) Sham acupuncture. A plastic cover of a needle was used. The plastic cover was pushed against the skin with a quick force to mimic sensation of a needle insertion. Patients received three sessions over one week.	**Mixed (hip and knee) osteoarthritis** Mean age (SD): Dry needling group: 61.00 (8.19); sham group: 56.20 (6.03) years N = 40 Definition: Moderate osteoarthritis symptoms (grade 2-3 of Kellgren-Lawrence Classification System) in muscles around the hip and knee joint Severity of symptoms: Grade 2-3 of Kellgren-Lawrence Classification criteria Duration of symptoms: Not reported Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: not reported			
Fink 2001[34]	**Acupuncture** (n=33) Ten treatments over 3 weeks. Using traditional Chinese medicine theory, applied to 3 local points and 3 distal points. Treatment was continued for 20 minutes with twisted to elicit a De qi sensation (manipulation was carried out 2-3 times a session). **Sham acupuncture** (n=34) Acupuncture performed in the same way except selecting puncture sites at least 5cm away from the classical acupuncture points and their interconnecting lines and clear of the painful pressure points.. **Concomitant therapy:** No additional information	**Hip osteoarthritis** Mean age (SD): 62.6 (9.1) years N = 67 Definition: People with pain, discomfort and movement restriction in the hip with radiographic changes of at least 2 on a Kellgren-Lawrence score Severity: Kellgren Lawrence grade 2-4, median grade 3 Duration of symptoms (mean [SD]): 5.2 (3.8) years Presence of multimorbidities: Not stated/Unclear	Serious adverse events at ≤3 months	
Foster 2007[37] Subsidiary papers: Whithurst 2011[153]	**Acupuncture** (n=117) Acupuncture using 6-10 points from 16 local and distal points per session. Treatment was performed with 30x0.3mm needles to a depth of 0.5-2.5cm and were left in for 25-35 minutes. The De qi sensation was achieved. 6 sessions over 3 weeks.	**Knee osteoarthritis** Mean age (SD): 63.2 (8.8) years N = 352 Definition: Clinical diagnosis of knee osteoarthritis Severity: Not stated	Pain at ≤3 months and >3 months Physical function at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Sham acupuncture (n=119) Needles were used that collapsed into the handle creating an illusion of insertion. Otherwise the same approach was used. **No treatment** (n=116) **Concomitant therapy:** Advice and exercise. Advice is given by a leaflet that contains standard advice on the use of analgesia. If people are taking NSAIDs, they were permitted to continue their stable dose. Exercise was conducted as a program with a maximum of 6x 30 minute sessions over 6 weeks including concentric, eccentric, isometric and balance exercises.	Duration of symptoms: <1 - at least 10 years, median time 1 to <5 years. Presence of multimorbidities: Not stated/Unclear		
Hinman 2014[47]	**Acupuncture** (n=70) Twenty minute treatments delivered once or twice weekly for 12 weeks, with 8 to 12 sessions in total permitted. Using a selection of local and distal points (31+ options). Using 0.25x40mm needles. **No treatment** (n=71) Two additional groups (n=70 and 71 respectively) were not	**Knee osteoarthritis** Mean age (SD): 63.6 (8.4) years N = 282 Definition: Knee pain on most days with an average severity of 4 or more out of 10 on a NRS and had morning stiffness lasting less than 30 minutes (consistent with a clinical diagnosis of osteoarthritis)	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
	included as they did not fulfil the inclusion criteria **Concomitant therapy:** No additional information	Severity: Not stated Duration of symptoms: ≤1 to at least 10 years, median 5 to 10 years Presence of multimorbidities: Not stated/Unclear		
Ju 2015[54]	**Electroacupuncture** (n=40) 30mm, 30 gauge needles inserted into 6 local points. De qi sensation was achieved. Stimulators were inserted for three needle pairs and maintained for 30 minutes at the tolerated threshold level (around 5-6mA). People received 16 treatments: five times a week for the first 2 weeks, and three times a week during the following 2 weeks. **Sham acupuncture** (n=40) Same treatment, but the intensity of electrical stimulation was relatively weak so that people couldn't feel the electroacupuncture stimulus and then an additional 1mA was added. **Concomitant therapy:** All people were encouraged to self-exercise, and pay attention to maintaining good posture. Meanwhile, all people received	Knee osteoarthritis Mean age (SD): 61.5 (8.2) years N = 80 Definition: Knee osteoathritis according to the criteria in the American College of Rheumatology and the presence of a severity grade of 2 or 3 according to the radiological Kellgren classification 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
	30mg etoricoxib tablets once a day during the study			
Kong 2018[61]	**Acupuncture** (n=44) Boosted or standard acupuncture (boosting was achieved through expectation management). 9 acupuncture sessions over 13 clinical assessment sessions with acupuncture for 4 weeks (2 times/week for the first 2 weeks, then 1 time/week for the last 2 weeks). Conducted on 6 acupoints using traditional Chinese theory. Needles were stimulated one at a time for 10s with 30s breaks between acupunctures, 120 rotations per minute to achieve a moderate De qi sensation. **No treatment** (n=22) **Concomitant therapy:** All people were told to maintain their baseline medications and other treatments for their knee osteoarthritis during the duration of the study	Knee osteoarthritis Mean age (SD): 61.0 (7.3) years N = 66 Definition: People with knee osteoarthritis meeting the American College of Rheumatology classification with radiographic evidence of grade 2 or 3 knee osteoarthritis using the Kellgren-Lawrence scale 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	
Lam 2021[63]	**Acupuncture** (n=43) Acupuncturists treated 5-8 affected points for each painful knee. The needles were inserted into muscle in a length of 10-20mm at an angle of 0-10	**Knee osteoarthritis** Mean age (SD): Acupuncture group: 62.7 (7.0) years; Sham group: 63.4 (6.7) years N = 86	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	degrees to the skin. The needles were adjusted by extension and flexion of the knee joint to ensure that all the needles would not cause pain during movement. Participants were then advised to walk for 10 minutes, followed by stepping up and down from an 18cm step for 12 rounds per knee and sitting for 5 minutes. **Sham acupuncture** (n=43) Participants in the sham group underwent the same procedures as those in the acupuncture group except that non-insertion sham acupuncture was employed. Briefly, the participants were in the sitting position and blinded by the trolley table. After disinfection of the acupoints, the acupuncturists applied needles (0.30 mm x 40 mm) on acupoints without penetrating the skin, followed by the cover of bandages. The other procedures were identical to those in the acupuncture group. **Concomitant therapy:** not reported	Definition: the American College of Rheumatology clinical classification criteria for osteoarthritis of the knee, had present knee pain, and had less than 30 minutes of morning stiffness or crepitus on active motion and osteophytes, as determined by history and physical examination; had either unilateral knee pain or bilateral knee pain; and knee pain intensity over 40mm on a visual analogue scale Severity of symptoms (VAS, mean, SD): Acupuncture group: 71.2 (16.0); Sham group: 70.1 (19.7) Duration of symptoms (n, %): Acupuncture group: ≤1 year - 1 (2.4), <1 to 5 years - 21 (50), <5 to 10 - 12 (28.6), >10 years - 8 (19.0); Sham acupuncture group: ≤1 year - 2 (4.9), <1 to 5 years - 13 (31.7), <5 to 10 - 18 (43.9), >10 years - 8 (19.5) Presence of multimorbidities: Not stated/Unclear		
Lansdown 2009[64]	**Acupuncture** (n=15)	**Knee osteoarthritis**	Quality of life at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Acupuncture by a flexible approach, with the number of needles inserted, depth of needle insertion, needle responses elicited, needle stimulation used, retention time and needle type varying. Treatments were usually weekly for 10 sessions. **No treatment** (n=15) **Concomitant therapy:** Both groups received usual care, which included any appointments, medications (prescribed or over the counter) and interventions sought by participants from any health practitioner	Mean age (SD): 63.5 (8.2) years N = 30 Definition: People over 50 years old who had consulted their GP in the last 3 years with knee pain (capturing the clinical symptoms of osteoarthritis of the knee, but no radiographically confirmed diagnosis) Severity: Not stated 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 Serious adverse events at >3 months	
Lev-ari 2011[67]	**Acupuncture** (n=28) Acupuncture using traditional Chinese theory using 8 acupoints. Needles of 0.16mm were left in place for 20 minutes and manually manipulated every 5 minutes. This was carried out twice weeks for 8 weeks. **Sham acupuncture** (n=27) Same protocol but empty needle tubes were tapped to the acupoints instead of the needles.	**Knee osteoarthritis** Mean age (SD): 71.2 (8.9) years N = 55 Definition: People diagnosed as having osteoarthritis of the knee of at least 6 months duration with pain Severity: Not stated Duration of symptoms: At least 6 months	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Concomitant therapy: No additional information	Presence of multimorbidities: Not stated/Unclear		
Lin 2018[73]	**Acupuncture** (n=21) Traditional Chinese acupuncture using 10 commonly used local points and 3-4 out of 11 distal points at each session. For local points 0.30mmx40mm needles were used, while 0.30mmx25mm needles were used for distal points. Needles were inserted 10-30mm. People had to achieve a De qi sensation. Needles were manipulated during treatment. Conducted for 8 weeks. **Sham acupuncture** (n=21) Minimal insertion into non-acupoints with no manipulation of the needle. **Concomitant therapy:** Celebrex was given to people with a pain score greater than and equal to 8/10. People were advised to not have any other treatments	**Knee osteoarthritis** Mean age (SD): 59.8 (7.4) years N = 42 Definition: People diagnosed as having osteoarthritis of the knee of at least 6 months duration with pain Severity: Kellgren Lawrence score of 2-3 Duration of symptoms: At least 6 months 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 Serious adverse events at >3 months	
Lv 2019[82]	**Electroacupuncture** (n=225) Strong (2-5mA) and weak (0-0.5mA) electroacupuncture (the two groups were combined due to class effect). Delivered in 10x 30 minute sessions over 2 weeks with needles inserted into	**Knee osteoarthritis** Mean age (SD): 63.7 (9.9) years N = 301 Definition:	Pain at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	four acupoints following traditional Chinese medicine meridian theory. De qi sensation was elicited. **Sham acupuncture** (n=76) Sham using the same number of acupoints, electroacupuncture apparatus and stimulation. However, the needles used were fine and short (35-gauge with an outer diameter of 0.20mm and a length of 25mm) and so were only inserted superficially into non-acupoints. Electrical stimulation was applied in the same manner as the weak electroacupuncture group. **Concomitant therapy:** All people were required not to take analgesic medications and electroacupuncture 48 hours before each treatment session	Clinical criteria for knee osteoarthritis formulated by the American College of Rheumatology Severity: Not stated Duration of symptoms: Between <0.5 years and at least 5 years, median 0.5 to 3 years Presence of multimorbidities: Not stated/Unclear		
Mavrommatis 2012[92]	**Electroacupuncture** (n=40) Acupuncture using 30mm, 30 gauge acupuncture needles into 6 local points and 4 distal points. De qi sensation was confirmed. Treatment was given biweekly for 8 weeks. Electrical stimulation was started from the third session at 6Hz, 150ms for 20 minutes for two needle pairs.	**Knee osteoarthritis** Mean age (SD): 61.8 (10.6) years N = 120 Definition: People had to have met the American College of Rheumatology criteria for diagnosis of knee osteoarthritis with Kellgren	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Sham acupuncture (n=40) Retractable needles were used. Otherwise the same protocol (including electrical stimulation). **No treatment** (n=40) **Concomitant therapy:** Everyone was treated with etoricoxib 60mg only on a daily basis and were examined biweekly. People with risk factors for upper gastrointestinal tract complications received proton pump inhibitors	Lawrence scores of at least 2 and chronic pain in the knee joint for more than 3 months Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms: At least 3 months Presence of multimorbidities: Not stated/Unclear		
Miller 2011[95]	**Acupuncture** (n=28) Acupuncture following Traditional Chinese Medicine treatment methods. Treatment twice weekly for 8 weeks. **Sham acupuncture** (n=27) Same procedure as the acupuncture group, but using tubes that were taped to the skin instead of needle insertion. **Concomitant therapy:** All people received standard therapy, which included treatment with NSAIDs.	**Knee osteoarthritis** Mean age (SD): 71.2 (8.9) years N = 55 Definition: Osteoarthritis of the knee for at least 6 months with moderate to severe pain on most days throughout the past month Severity: Not stated/unclear. Duration of symptoms: At least 6 months. Presence of multimorbidities: Not stated/unclear	Pain at ≤3 months Physical function at ≤3 months	
Min 2009[96]	**Acupuncture** (n=40)	**Knee osteoarthritis**	Quality of life at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Sa-am acupuncture using 0.25 x 40mm needles inserted about 1-5mm for a duration of 20 minutes, being delivered two times per week over 4 weeks. Twirling reinforcement reduction and in six reinforcement reduction methods were used, people reported feeling deqi. Sets of four acupoints were used. **Sham acupuncture** (n=38) Using park sham needles inserted into the same position, by the same frequency and duration **Concomitant therapy:** No additional information	Mean age (SD): 59.4 (5.3) years N = 78 Definition: Osteoarthritis of the knee according to the American College of Rheumatology with documented radiographic changes of osteoarthritis Severity: Kellgren Lawrence grades 1-4, median grade 2 Duration of symptoms: At least 6 months Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	
Sanchez romero 2020[111]	**Acupuncture** (n=31) Dry needling for 6 sessions in 6 weeks. Using a fast-in fast-out technique (15 times of manipulating the needle upwards and downwards). Using 0.3x40mm, 0.4x60mm or 0.3x0.75mm needles inserted into myofascial trigger points located under the index and middle fingers of the nondominant hands. **Sham acupuncture** (n=31)	**Knee osteoarthritis** Mean age (SD): 72.3 (5.7) years N = 62 Definition: Primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnostics Severity: 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	This study was rated as having serious indirectness due to population indirectness (people were required to have myofascial trigger points).

Study	Intervention and comparison	Population	Outcomes	Comments
	Same areas, but with a simulated device that does not penetrate the skin. **Concomitant therapy:** Therapeutic exercise for 1 hour, twice a week for 12 weeks	Duration of symptoms (mean [SD]): 65.6 (36.0) months Presence of multimorbidities: Not stated/Unclear		
Sanchez-romero 2018[110]	**Acupuncture** (n=11) Dry needling for 6 sessions in 6 weeks. Using a fast-in fast-out technique (15 times of manipulating the needle upwards and downwards). Using 0.3x40mm, 0.4x60mm or 0.3x0.75mm needles inserted into myofascial trigger points located under the index and middle fingers of the nondominant hands. **Sham acupuncture** (n=9) Same areas, but with a simulated device that does not penetrate the skin. **Concomitant therapy:** All people received a therapeutic exercise program in 1 hour, group based, supervised sessions twice weekly over 12 weeks. On average, about 10 people attended each training session. A total of 24 sessions were conducted consisting of aerobic exercise (20-25 minutes	**Knee osteoarthritis** Mean age (SD): 71.4 (4.2) years N = 20 Definition: Knee pain and uni- or bilateral dysfunction with primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnosis Severity: Not stated Duration of symptoms (mean [SD]): 28.5 (25.2) months Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months	This study was rated as having serious indirectness due to population indirectness (people were required to have myofascial trigger points).

Study	Intervention and comparison	Population	Outcomes	Comments
	warm up), lower limb muscle strengthening (20-25 minutes), and lower-limb muscle stretching (10-15 minutes)			
Sangdee 2002[112]	**Electroacupuncture** (n=97) Electroacupuncture using four fine stainless steel needles inserted into acupoints in the affected knee (using traditional Chinese theory). Inserted superficially (no more than 0.5 inch in depth). De qi sensation was not required. Electrical stimulation was achieved by biphasic pulses at a frequency of 2Hz and was administered for 20 minutes. People were treated 3 times a week for 4 weeks. People were also given either a) diclofenac 25mg three times a day for 4 weeks or b) placebo three times a day for 4 weeks **Sham acupuncture** (n=95) Same areas but electrodes were connected to a sound producing dummy mode that did not produce a current. People were also given either a) diclofenac 25mg three times a day for 4 weeks or b) placebo three times a day for 4 weeks. **Concomitant therapy:**	**Knee osteoarthritis** Mean age (SD): 62.9 (7.2) years N = 192 Definition: Unilateral or bilateral osteoarthritis of the knee according to the criteria of the American College of Rheumatology for more than 3 months duration Severity: Not stated Duration of symptoms (mean [SD]): 4.9 (3.9) years Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Physical function at ≤3 months Osteoarthritis flares at ≤3 months Serious adverse events at ≤3 months	The osteoarthritis flares and serious adverse events outcomes were rated as having serious indirectness due to the events reported being withdrawal events

Study	Intervention and comparison	Population	Outcomes	Comments
	All additional therapies (e.g. oral or topical NSAIDs, intraarticular corticosteroid injection, other analgesics, chondro-protective agents, surgical procedures on the knee joint etc.) were not allowed. However, all other treatments for concomitant disorders that did not interfere with the study could be continued, but it had to be documented. Paracetamol was prescribed as rescue analgesic			
Scharf 2006[113]	**Acupuncture** (n=330) 10 acupuncture sessions over a 6 week period. Using traditional Chinese theory, needles were inserted into 6 acupoints. In addition, 2 of 16 distal points could be chosen. **Sham acupuncture** (n=367) Minimal depth needles without stimulation into points at defined distanced from traditional Chinese acupoints. A third group (n=342) was reported but not included as it did not fulfil the inclusion criteria. **Concomitant therapy:** All people received conservative therapy of 150mg of diclofenac per day during the first 2	**Knee osteoarthritis** Mean age (SD): 62.81 (10.04) years N = 1039 Definition: Chronic pain in the knee joint for the last 6 months, according to the American College of Rheumatology criteria with radiologic confirmation of osteoarthritis in 1 or both knees (Kellgren Lawrence score 2-3) Severity: Kellgren-Lawrence score 0-4, median score 2 Duration of symptoms (mean [SD]): 65.08 (71.07) months 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 Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	treatment weeks up to a total of 1g until week 23			
Suarez-almazor 2010[123]	**Electroacupuncture** (n=153) Traditional Chinese theory used to insert needles. Electroacupuncture achieved using TENS equipment to emit a dense disperse wave impulse at 50Hz, dispersing at 15Hz, 20 cycles/minute with a voltage increased slowly from 5V to 60v until maximal tolerance was achieved. People rested for 20 minutes of treatment. **Sham acupuncture** (n=302) Shallow needle insertion with the device set to deliver an adjustable wave with a voltage increased to where the person could feel it, and then switched off. **No treatment** (n=72) **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 62.81 (10.04) years N = 1039 Definition: Chronic pain in the knee joint for the last 6 months, according to the American College of Rheumatology criteria with radiologic confirmation of osteoarthritis in 1 or both knees (Kellgren Lawrence score 2-3) Severity: Kellgren-Lawrence score 0-4, median score 2 Duration of symptoms (mean [SD]): 65.08 (71.07) months Presence of multimorbidities: Not stated/Unclear	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	This study also examined the effect of treatment expectations on outcome.
Takeda 1994[127]	**Acupuncture** (n=20) Acupuncture three times a week for 3 weeks. 30mm needles with 0.23mm diameter were inserted into 5 acupoints, then inserted, rotated and inserted deeper until the person experienced Te chi. Needles were left in place for 30	**Knee osteoarthritis** Mean age (SD): 61.6 (9.4) years N = 40	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	minutes and rotated every 5 minutes **Sham acupuncture** (n=20) Superficial needle insertion and only touched periodically, no specific manipulation technique. **Concomitant therapy:** No additional information	Definition: Grade 1-4 radiographic osteoarthritis with pain in one or both knees Severity: Grade 1-4, median grade 2 Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear		
Tu 2021[135] Subsidiary studies: Tu 2019[134] Wang 2021[145]	**Electroacupuncture** (n=156) Electroacupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. Disposable sterile needles (0.25mm x 25-40mm) and HANS-200 electroacupuncture devices were used. The prescription was based on traditional Chinese medicine. Five obligatory acupoints and three adjunct acupoints were used. **Manual acupuncture** (n=155) Manual acupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. . The prescription was based on traditional Chinese medicine. Five obligatory acupoints and three adjunct acupoints were used.	**Knee osteoarthritis** Mean age (SD): 62.8 (7.1). years N = 480 Definition: People reporting knee pain for longer than 6 months with a radiological confirmation of osteoarthritis (Kellgren Lawrence score 2-3) Severity: Radiological grade 2-3, median grade 2 Duration of symptoms: 6.6 (5.7) years Presence of multimorbidities: Low comorbidity score (0 concomitant diseases: 226, 1 concomitant disease: 145, 2 concomitant diseases: 59, 3 or more concomitant diseases: 12). 4	Quality of life at ≤3 months and >3 months 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
	Sham acupuncture (n=157) Sham acupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. The prescription was based on traditional Chinese medicine. Eight non-acupoints were used. **Concomitant therapy**: not reported.			
Vas 2004[139] Subsidiary paper: Vas 2006[138]	**Electroacupuncture** (n=48) 30 gauge, 45mm length needles inserted into 4 local points and 4 distal points. Sensation of De qi was necessary. Electrical stimulation was conducted with a WQ-10D1 electrostimulator. Treatments were for 11 weeks. **Sham acupuncture** (n=49) Retractable needles were used. Otherwise the same frequency, duration and electrical stimulation. **Concomitant therapy:** All people were given a bag with 21 tablets of 50mg diclofenac for the week (50mg every 8 hours) with instruction to reduce the dose if symptoms improved.	**Knee osteoarthritis** Mean age (SD): 67.1 (10.2) years N = 97 Definition: Outpatients who had been clinically and radiologically diagnosed according to the criteria of the American College of Rheumatology Severity: Ahlback grade 1-4, median grade 2 Duration of symptoms (mean [SD]): 7.5 (8.6) 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
	People with risk factors received gastroprotective drugs			
Wang 2020[146]	**Electroacupuncture** (n=30) All participants underwent acupuncture needling at a selection of local and distant traditional acupuncture points or ah shi points chosen by the acupuncturists according to the principles of traditional Chinese medicine. Needles were inserted at 6-7 local points. Individual syndrome differentiation was used. If pain occurred on the outside of the affected knee joint, GB points were mainly selected. If pain occurred in front of the affected knee joint, ST points were selected. If pain occurred in the interior of the affected knee joint, SP, LR and KI points were chosen. If pain occurred in the rear of the affected knee, BL joints were used. Needles were stimulated manually for 1-seconds to achieve di qi sensation. There was a total of 24 sessions lasting 30 minutes each, over a period of 8 weeks **Acupuncture/dry needling** (n=30) Participants in the manual acupuncture group had the same schedule as the	**Knee osteoarthritis** Mean age (SD): Electroacupuncture group: 58.89 (6.75); manual acupuncture group: 59.70 (7.36). years N = 60 Definition: radiographically confirmed KOA affecting one or both knees with a duration of more than 6 months and pain intensity of 40 or more on a 100-point visual analogue scale Severity: All participants had to have a Kellgren-Lawrence grade II or III and score 40 or more on a pain intensity VAS Duration of symptoms (mean [SD]): Electroacupuncture group: 69.93 (56.69) months; manual acupuncture group: 73.20 (56.71) months 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 Serious adverse events at >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	electroacupuncture except that the electrical apparatus featured a working power indicator and sound without actual current output. The middle wire was cut, although the appearance of the unit was identical. Thus, the EA instrument appeared to be "on", but the actual power was not energised. After elicitation of de qi sensation by MA, needles were retained for 30 minutes. Although no manual manipulation of the needles was performed after initially achieving de qi sensation, the stimulation associated with needle retention was still expected to induce therapeutic effects. Therefore the only difference between the two groups was the electrical current in the electroacupuncture group. Duration 8 weeks. **Concomitant therapy:** Participants were advised not to take any NSAIDs or analgesics except for a 'rescue medication'. The use of NSAIDs was recorded.			
Wang 2021[147]	**Electroacupuncture** (n=30) Acupoints were chosen according to traditional Chinese medicine and were localised according to the WHO Standard	**Knee osteoarthritis** Mean age (SD): Electroacupuncture group: 64.73 (5.39) years; Sham	Quality of life at ≤3 months and >3 months Pain at ≤3 months and >3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Participants Acupuncture Locations. Needles were stimulated manually for 10 seconds to achieve de qi sensation. Treatment consisted of 24 sessions lasting 30 minutes each over 8 weeks (usually three times per week). **Sham acupuncture** (n=30) Eight non-acupoints that were separate from conventional acupoints or meridians were used for the sham group. The schedule, electrode placements and other treatment settings were the same as for the electroacupuncture group but with superficial skin penetration (2-3 mm in depth) and no electricity output or needle manipulation for de qi. Duration 8 weeks **Concomitant therapy:** All participants were advised not to take any NSAIDs or analgesics except for a "rescue analgesic" (1 tablet of 200mg Acetaminophen orally as needed, once per day). The use of NSAIDs was recorded.	acupuncture group 66.10 (7.42) years N = 60 Definition: radiographically confirmed KOA affecting one or both knees with a duration of more than six months and pain intensity ≥4 on a 10 point numerical rating scale Severity: All participants had Kellgren-Lawrence grade II or III. Pain score on numerical rating scale (mean, SD): Electroacupuncture group 6 (1.34); Sham group 6.13 (1.33) Duration of symptoms (mean [SD]): Electroacupuncture group: 85.73 (74.15) months; Sham acupuncture group: 104.37 (96.45) months Presence of multimorbidities: Not stated / Unclear	Physical function at ≤3 months and >3 months Serious adverse events at >3 months	
Weiner 2007[149]	**Electroacupuncture** (n=44) Periosteal stimulation therapy using four 30 gauge acupuncture needles being	**Knee osteoarthritis** Mean age (SD): 71.5 (5.4) years	Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	inserted into the medial femoral condyle, lateral femoral condyle, flare of tibia and head of fibula. The needles were stimulated with 100Hz for 30 minutes. The intensity was adjusted so it was clearly felt but not uncomfortable. Two needles were inserted into the soft tissue on the upper third of the tibial shaft and were stimulated with 100Hz for 1 minute. **Sham acupuncture** (n=44) Same needle insertion, but no stimulation of the needles in bone. Two needles were inserted into the soft tissue on the upper third of the tibial shaft and were stimulated with 100Hz for 1 minute. **Concomitant therapy:** No additional information	N = 88 Definition: Chronic knee pain and radiographic knee osteoarthritis Severity: Kellgren Lawrence grade 2-4, median grade 4 Duration of symptoms (mean [SD]): 8.0 (7.4) years Presence of multimorbidities: High comorbidity score (Comorbidity mean (SD): 1.95 (1.30))	Psychological distress at ≤3 months Serious adverse events at ≤3 months	
White 2012[151]	**Acupuncture** (n=74) Western acupuncture using a flexible approach with a prescribed range of points. A mean of 6 points were used, with deep needling, which lasted 20 minutes twice a week for 4 weeks. De qi was elicited for each needle through rotation. **Sham acupuncture** (n=73)	**Mixed osteoarthritis (hip or knee)** Mean age (SD): 66.75 (8.29) years N = 221 Definition: Chronic osteoarthritis pain from a single joint (hip or knee)	Pain at ≤3 months Serious adverse events at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	Sham needling using streitberger needles. A third group (n=74) was not included as it did not fulfil the inclusion criteria. **Concomitant therapy:** No additional information	Severity: Not stated Duration of symptoms: Not stated Presence of multimorbidities: Not stated/Unclear		
Williamson 2007[154]	**Acupuncture** (n=60) Group acupuncture (groups of 6-10 people) once a week for 6 weeks using 1 inch, 0.25 gauge needles. De chi sensation was achieved were possible, and needles were left in situ for 20 minutes. **No treatment** (n=60) Exercise and advice leaflet only A third group (n=60) was not included as it did not fulfil the inclusion criteria. **Concomitant therapy:** No additional information	**Knee osteoarthritis** Mean age (SD): 70.7 (9.0) years N = 180 Definition: People listed for knee arthroplasty due to osteoarthritis with unilateral or bilateral knee pain for more than 3 months Severity: Not stated Duration of symptoms: At least 3 months Presence of multimorbidities: Not stated/Unclear	Pain at ≤3 months Psychological distress at ≤3 months Serious adverse events at ≤3 months	
Witt 2005[155] Subsidiary paper: Brinkhaus 200712	**Acupuncture** (n=150) Acupuncture with 12 sessions of 40 minute treatment over 8 weeks (usually 2 sessions per week for the first 4 weeks, then 1 session per week for the remaining 4). Needles were	**Knee osteoarthritis** Mean age (SD): 64.0 (6.5) years N = 300	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
	inserted in at least 6 local points and at least 2 distal points. Physicians were able to choose the needle length and diameter. People were instructed to achieve De qi if possible, with the needles being stimulated at least once during each session **Sham acupuncture** (n=76) Minimal acupuncture using the same methods, but only superficial insertion of fine needles (20-40mm in length) at predefined, distant non-acupuncture points **No treatment** (n=74) Waiting list control **Concomitant therapy:** No additional information	Definition: Diagnosis with osteoarthritis to the American College of Rheumatology criteria with documented radiological alterations in the knee joint of grade 2 or more according to Kellgren Lawrence criteria Severity: Kellgren Lawrence grade 0-4, median grade 3 Duration of symptoms (mean [SD]): 9.2 (7.9) years Presence of multimorbidities: Not stated/Unclear	Psychological distress at ≤3 months and >3 months Serious adverse events at >3 months	
Witt 2006[156] Subsidiary papers: Martins 2014[91] Reinhold 2008[107]	**Acupuncture** (n=357) 15 sessions during the first 3 months of the study. The number of needles and points were chosen at the physician's discretion. Only needle acupuncture was allowed. In addition, only manual needle stimulation was allowed. **No treatment** (n=355)	**Mixed osteoarthritis (hip or knee)** Mean age (SD): 61.2 (10.4) years N = 712 Definition: Clinical diagnosis of osteoarthritis-associated pain in the knee or hip with disease duration of >6 months with radiologic	Quality of life at ≤3 months Pain at ≤3 months Physical function at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	At 3 months people received the same acupuncture treatment (therefore, only results from ≤3 months are included in this review) **Concomitant therapy:** No additional information	evidence of osteoarthritis (osteophyte formation) Severity: Not stated Duration of symptoms (mean [SD]): 5.3 (6.2) years Presence of multimorbidities: Not stated/Unclear		
Zhang 2019[162]	**Electroacupuncture** (n=30) direct current and dilatational wave was delivered with Electronic acupuncture Treatment Instrument at 2/100 Hz frequency and 0.2 ms pulse width for 30 minutes.. The intensity was not prescribed equally, but patients in the acupuncture and electroacupuncture groups were encouraged to increase them along with the physical fitness. Participants had 10 sessions over a period of two weeks, each lasting 30 minutes **Acupuncture** (n=30) Patients in this group were treated with acupuncture on eight ipsilateral acupoints once a day. After the local area had been disinfected, the needles (30-gauge with an outer diameter of 0.3mm and a length of 40 mm) would be inserted at	**Mixed osteoarthritis (hip or knee)** Mean age (SD): Acupuncture group: 55.9 (5.8) years; usual care group: 55.2 (6.0) years; Electroacupuncture group: 54.9 (6.2) years N = 90 Definition: diagnoses according to the American College of Rheumatology criteria Severity of symptoms (number with Kellgren Lawrence Grade II or III): Acupuncture group 15/15; Usual care group 13/17; Electroacupuncture group 11/19 Duration of pain (mean, SD): Acupuncture group 6.6 (2.0) years; usual care group 6.2 (1.8) years; electroacupuncture group 6.5 (1.8) years	Quality of life at ≤3 months	

Study	Intervention and comparison	Population	Outcomes	Comments
	a depth of 25 to 40 mm vertically through lifting and thrusting combined with twirling and rotating the needles, De Qi sensation were achieved. Participants had 10 sessions over a period of two weeks, each lasting 30 minutes. **No treatment** (n=30) **Concomitant therapy:** Patients who had previously taken drugs for activating blood circulation (Ds-ABC) and non-steroidal anti-inflammatory drugs (NSAIDs) or COX2-inhibitors were allowed to continue to take these medications during the study period, however they were acted to avoid physical therapy as much as possible to ensure that the results reflected the role of acupuncture or electroacupuncture as much as possible rather than other forms of treatment.	Presence of multimorbidities: Not stated/Unclear		

See Appendix D for full evidence tables.

1.1.6 Summary of the effectiveness evidence

Table 3: Clinical evidence summary: Acupuncture/dry needling compared to sham acupuncture

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with sham acupuncture	Risk difference with acupuncture	
Quality of life (EQ-5D, 5-15, high is good, final value) at ≤3 months	62 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	MD **0.13 higher** (0.56 lower to 0.82 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	1541 (6 RCTs) follow up: mean 9 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 22.7	MD **1.26 higher** (0.21 higher 1.44 higher)	MID = 4.3 (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	1541 (6 RCTs) follow up: mean 9 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 28.5	MD **0.56 higher** (0.48 lower to 1.6 higher)	MID = 2.9 (0.5 x median baseline SD)
Quality of life (EQ-5D, 5-15, high is good, final value) at >3 months	62 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,b,c]	-	-	MD **0.15 higher** (0.58 lower to 0.88 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score and final values) at >3 months	1250 (4 RCTs) follow up: mean 33 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 28.6	MD **1.31 higher** (0.13 higher to 2.49 higher)	MID = 4.1 (0.5 x median baseline SD)
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change score and final values) at >3 months	1250 (4 RCTs) follow up: mean 33 weeks	⊕⊕◯◯ LOW [d]	-	The mean quality of life was 35.5	MD **0.92 higher** (2.11 lower to 3.95 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 sham acupuncture	Risk difference with acupuncture	
Pain (WOMAC [different scale ranges], high is poor, change scores) at ≤3 months	1252 (5 RCTs) follow up: mean 9 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.08 SD lower** (0.19 lower to 0.03 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, KSS [different scale ranges], high is poor, final values) at ≤3 months	880 (8 RCTs) follow up: mean 9 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.30 SD lower** (0.44 lower to 0.17 lower)	MID = 0.5 SD (SMD)
Pain (WOMAC [different scale ranges], high is poor, change scores) at >3 months	1108 (3 RCTs) follow up: mean 35 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.06 SD lower** (0.18 lower to 0.06 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC [different scale ranges], high is poor, final values) at >3 months	621 (4 RCTs) follow up: mean 29 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.21 SD lower** (0.38 lower to 0.05 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC [different scale ranges], high is poor, change scores) at ≤3 months	1247 (5 RCTs) follow up: mean 9 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.06 SD lower** (0.17 lower to 0.05 higher)	MID = 0.5 SD (SMD)
Physical function (WOMAC, KSS [different scale ranges], high is poor, final values) at ≤3 months	735 (7 RCTs) follow up: mean 10 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.28 SD lower** (0.43 lower to 0.13 lower)	MID = 0.5 SD (SMD)
Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months	1108 (3 RCTs) follow up:	⊕⊕⊕⊕ HIGH	-	-	SMD **0.03 SD lower** (0.15 lower to 0.09 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 sham acupuncture	Risk difference with acupuncture	
	mean 35 weeks					
Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months	621 (4 RCTs) follow up: mean 29 weeks	⊕⊕⊕⊕ HIGH	-	-	SMD **0.22 SD lower** (0.38 lower to 0.06 lower)	MID = 0.5 SD (SMD)
Psychological distress (Depression ADS, 0-100, high is poor, final value) at ≤3 months	226 (1 RCT) follow up: 8 weeks	⊕⊕⊕⊕ HIGH	-	The mean psychological distress was 48.3	MD **0.4 lower** (3.07 lower to 2.27 higher)	MID = 0.5 SD (SMD)
Psychological distress (Depression ADS, 0-100, high is poor, final value) at >3 months	226 (1 RCT) follow up: 52 weeks	⊕⊕⊕⊕ HIGH	-	The mean psychological distress was 49.8	MD **1.2 lower** (4 lower to 1.6 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	212 (2 RCTs) follow up: mean 4 weeks	⊕⊕◯◯ LOW [e]	RD 0.12 (-0.26, 0.50)	38 per 1,000	**120 more per 1,000** (260 fewer to 500 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	1485 (5 RCTs) follow up: mean 26 weeks	⊕◯◯◯ VERY LOW [e,g]	RD 0.04 (-0.02, 0.10)	312 per 1,000	**40 more per 1,000** (20 fewer to 100 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).

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 of population indirectness

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 by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

e. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with sham acupuncture	Risk difference with acupuncture	
f. Absolute effect calculated by risk difference due to zero events in at least one arm of one study g. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size						

Table 4: Clinical evidence summary: Acupuncture/dry needling 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 acupuncture	
Quality of life (EQ-5D, KOOS [different scale ranges], high is good, final values) at ≤3 months	309 (3 RCTs) follow up: mean 7 weeks	⊕⊕◯◯ LOW [a]	-	-	SMD **0.11 SD higher** (0.11 lower to 0.34 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	989 (3 RCTs) follow up: mean 11 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 34.2	MD **4.69 higher** (1.27 higher to 8.11 higher)	MID = 4.4 (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	1034 (3 RCTs) follow up: mean 11 weeks	⊕◯◯◯ VERY LOW [a]	-	The mean quality of life was 52.0	MD **0.41 higher** (2.86 lower to 3.69 higher)	MID = 5.9 (0.5 x median baseline SD)
Quality of life (AQoL-SF-36 physical functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 640	MD **53 higher** (0.88 lower to 106.88 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 acupuncture	
Quality of life (AQoL-SF-36 bodily pain, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 207	MD **10 higher** (13.28 lower to 33.28 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 role physical, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 233	MD **10 higher** (55.64 lower to 75.64 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 vitality, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 307	MD **24 higher** (2.32 lower to 50.32 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 general health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 312	MD 7**5 higher** (34.31 higher to 115.69 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 mental health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 331	MD **24 higher** (10.92 lower to 58.92 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 role emotional, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 207	MD **36 higher** (10.55 lower to 82.55 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF-36 social functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 127	MD **27 higher** (13.3 higher to 40.7 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 acupuncture	
Quality of life (EQ-5D, -0.11-1, high is good, final values) at >3 months	263 (2 RCTs)	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 0.62	MD **0.01 higher** (0.06 lower to 0.08 higher)	MID = 0.03 (established value)
Quality of life (SF-12 physical component, 0-100, high is good, final value) at >3 months	121 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 38.9	MD **2.8 higher** (1.12 lower to 6.72 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12 mental component, 0-100, high is good, final value) at >3 months	121 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 54.4	MD **2.9 lower** (6.68 lower to 0.88 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score and final values) at ≤3 months	381 (3 RCTs) follow up: mean 10 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean pain was 5.3	MD **0.86 lower** (1.62 lower to 0.1 lower)	MID = 1.8 (0.5 x median baseline SD)
Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months	1022 (4 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.81 SD lower** (1.18 lower to 0.45 lower)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score and final values) at >3 months	348 (3 RCTs) follow up: mean 12 months	⊕⊕◯◯ LOW [a]	-	The mean pain was 5.1	MD **0.22 lower** (1.07 lower to 0.63 higher)	MID = 1.8 (0.5 x median baseline SD)
Physical function (WOMAC, 0-68, high is poor, change score and final values) at ≤3 months	381 (3 RCTs) follow up: mean 10 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean physical function was 17	MD **2.05 lower** (4.46 lower to 0.36 higher)	MID = 6.4 (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 acupuncture	
Physical function (KOOS, WOMAC, 0-100, high is poor, final values) at ≤3 months	902 (3 RCTs) follow up: mean 8 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	The mean physical function was 55.8	MD **15.58 lower** (23.58 lower to 7.57 lower)	MID = 10.6 (0.5 x median baseline SD)
Physical function (WOMAC, 0-68, high is poor, change score and final values) at >3 months	348 (3 RCTs) follow up: 12 months	⊕⊕◯◯ LOW [a]	-	The mean physical function was 15.5	MD **1.14 lower** (3.92 lower to 1.63 higher)	MID = 6.4 (0.5 x median baseline SD)
Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months	120 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean psychological distress was 6.54	MD **0.34 higher** (1.11 lower to 1.79 higher)	MID = 0.5 SD (SMD)
Psychological distress (HADS depression, depression ADS [different scale ranges], high is poor, final values) at ≤3 months	344 (2 RCTs) follow up: mean 10 weeks	⊕⊕⊕◯ MODERATE [a]	-	-	SMD **0.14 SD lower** (0.36 lower to 0.08 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	120 (1 RCT) follow up: 12 weeks	⊕◯◯◯ 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.
Serious adverse events at >3 months	30 (1 RCT) follow up: 12 months	⊕◯◯◯ VERY LOW [a,d]	RD 0.00 (-0.12 to 0.12)	0 per 1,000	**0 fewer per 1,000** (120 fewer to 120 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 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 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 5: Clinical evidence summary: Electroacupuncture compared to acupuncture/dry needling

Outcomes	№ of participants (studies) Follow-up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with acupuncture	Risk difference with electroacupuncture	
Quality of life (SF-12, 0-100, high is good, final value) at ≤3 months	58 (1 RCT) follow-up: 12 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 60.87	MD **1.34 higher** (7.75 lower to 10.43 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12 physical health, 0-100, high is good, final value) at ≤3 months	296 (1 RCT) follow-up: 8 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 39.22	MD **0.25 lower** (2.14 lower to 1.64 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12 mental health, 0-100, high is good, final value) at ≤3 months	296 (1 RCT) follow-up: 8 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 54.67	MD **0.94 lower** (2.96 lower to 1.08 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 physical functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 693	MD **25 higher** (35.76 lower to 85.76 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 bodily pain, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 217	MD **6 higher** (16.53 lower to 28.53 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 role physical, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life at <3 months was 243	MD **4 higher** (56.49 lower to 64.49 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 vitality, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a,b]	-	The mean quality of life was 331	MD **6 higher** (21.59 lower to 33.59 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 acupuncture	Risk difference with electroacupuncture	
Quality of life (AQoL-SF 36 general health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 387	MD **20 higher** (24.4 lower to 64.4 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 mental health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 355	MD **16 higher** (18.92 lower to 50.92 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 role emotional, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 243	MD **6 lower** (45.05 lower to 33.05 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 social functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 154	MD **18 higher** (3.07 higher to 32.93 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12, 0-100, high is good, final value) at >3 months	58 (1 RCT) follow-up: 16 weeks	⨁⨁◯◯ LOW [a,b]	-	The mean quality of life was 61.87	MD **1.77 higher** (7.32 lower to 10.86 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12 physical health, 0-100, high is good, final value) at >3 months	296 (1 RCT) follow-up: 26 weeks	⨁⨁⨁⨁ HIGH	-	The mean quality of life was 39.2	MD **0.02 lower** (2 lower to 1.96 higher)	MID = 0.5 SD (SMD)
Quality of life (SF-12 mental health, 0-100, high is good, final value) at >3 months	296 (1 RCT) follow-up: 26 weeks	⨁⨁⨁⨁ HIGH	-	The mean quality of life was 54.92	MD **0.81 lower** (2.66 lower to 1.04 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months	354 (2 RCTs)	⨁⨁⨁⨁ HIGH	-	The mean pain was 3.11	MD **0.37 lower** (0.78 lower to 0.04 higher)	MID = 1.42 (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 acupuncture	Risk difference with electroacupuncture	
	follow-up: mean 10 weeks					
Pain (WOMAC, 0-20, high is poor, final value) at >3 months	354 (2 RCTs) follow-up: mean 21 weeks	⊕⊕⊕⊕ HIGH	-	The mean pain was 3.38	MD **0.43 lower** (0.9 lower to 0.04 higher)	MID = 1.42 (0.5 x median baseline SD)
Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months	354 (2 RCTs) follow-up: mean 10 weeks	⊕⊕⊕◯ MODERATE [b]	-	The mean physical function was 10.95	MD **1.47 lower** (2.96 lower to 0.02 higher)	MID = 2.96 (0.5 x median baseline SD)
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months	354 (2 RCTs) follow-up: mean 21 weeks	⊕⊕⊕◯ MODERATE [b]	-	The mean physical function was 11.81	MD **1.63 lower** (3.19 lower to 0.06 lower)	MID = 2.96 (0.5 x median baseline SD)
Serious adverse events at >3 months	369 (2 RCTs) follow-up: mean 21 weeks	⊕⊕◯◯ LOW [b]	RR 0.83 (0.53 to 1.31)	178 per 1,000	**30 fewer per 1,000** (84 fewer to 55 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 6: Clinical evidence summary: Electroacupuncture compared to sham acupuncture

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with sham acupuncture	Risk difference with electroacupuncture	
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score and final values) at ≤3 months	1230 (5 RCTs) follow up: mean 9 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	MD **3.07 higher** (0.55 lower to 6.68 higher)	MID = 3.41 (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	892 (4 RCTs) follow up: mean 9 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean quality of life was 46.7	MD **0.71 higher** (0.4 lower to 1.83 higher)	MID = 4,68 (0.5 x median baseline SD)
Quality of life (PLQC physical capability, 0-4, high is good, final value) at ≤3 months	97 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 2.5	MD **0.3 higher** (0 to 0.6 higher)	MID = 0.5 SD (SMD)
Quality of life (PLQC psychological functioning, 0-4, high is good, final value) at ≤3 months	97 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 2.5	MD **0.2 higher** (0 to 0.4 higher)	MID = 0.5 SD (SMD)
Quality of life (PLQC negative mood, 0-4, high is good, final value) at ≤3 months	97 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 3.1	MD **0.1 higher** (0.18 lower to 0.38 higher)	MID = 0.5 SD (SMD)
Quality of life (PLQC social functioning, 0-4, high is good, final value) at ≤3 months	97 (1 RCT) follow up: 12 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 2.7	MD **0.1 higher** (0.14 lower to 0.34 higher)	MID = 0.5 SD (SMD)
Quality of life (PLQC social wellbeing, 0-4, high is good, final value) at ≤3 months	97 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 3.2	MD **0** (0.2 lower to 0.2 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 sham acupuncture	Risk difference with electroacupuncture	
Quality of life (SF-36 physical component, SF-12 physical health, 0-100, high is good, change score) at >3 months	640 (3 RCTs) follow up: mean 26 weeks	⊕⊕⊕◯ MODERATE [c]	-	The mean quality of life was 27.2	MD **1.36 higher** (0.4 lower to 3.11 higher)	MID = 3.12 (0.5 x median baseline SD)
Quality of life (SF-36 mental component, SF-12 mental health, 0-100, high is good, change score) at >3 months	357 (2 RCTs) follow up: 26 weeks	⊕⊕⊕⊕ HIGH	-	The mean quality of life was 47.42	MD **2.21 higher** (0.47 higher to 3.96 higher)	MID = 4.79 (0.5 x median baseline SD)
Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at ≤3 months	798 (3 RCTs) follow up: mean 7 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **1.32 SD lower** (3 lower to 0.36 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC [different scale ranges], high is poor, final values) at ≤3 months	1154 (7 RCTs) follow up: mean 10 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.59 SD lower** (1.02 lower to 0.17 lower)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score and final value) at >3 months	283 (1 RCT) follow up: mean 26 weeks	⊕⊕⊕⊕ HIGH	-	The mean pain was 3.34	MD **1.06 lower** (1.55 lower to 0.58 lower)	MID = 1.58 (0.5 x median baseline SD)
Physical function (WOMAC, 0-68, high is poor, change scores) at ≤3 months	501 (2 RCTs) follow up: mean 9 weeks	⊕⊕⊕◯ MODERATE [a]	-	The mean physical function was -11.4	MD **3.74 lower** (5.89 lower to 1.59 lower)	MID = 6.1 (0.5 x median baseline SD)
Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months	1154 (7 RCTs) follow up:	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **0.64 SD lower** (1.06 lower to 0.22 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 sham acupuncture	Risk difference with electroacupuncture	
	mean 10 weeks					
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months	580 (2 RCTs) follow up: mean 26 weeks	⊕⊕⊕⊕ HIGH	-	The mean physical function was -11.55	MD **3.1 lower** (4.66 lower to 1.55 lower)	MID = 5.25 (0.5 x median baseline SD)
Psychological distress (Geriatric depression scale, 0-20, high is poor, final value) at ≤3 months	88 (1 RCT) follow up: 12 weeks	⊕⊕◯◯ LOW [a,c]	-	The mean psychological distress was 3.36	MD **0.42 higher** (1.32 lower to 2.16 higher)	MID = 0.5 SD (SMD)
Osteoarthritis flares at ≤3 months	192 (1 RCT) follow up: 4 weeks	⊕◯◯◯ VERY LOW [a,c,d]	RR 0.98 (0.14 to 6.81)	21 per 1,000	**0 fewer per 1,000** (18 fewer to 122 more)	MID (precision) = RR 0.8-1.25.
Serious adverse events at ≤3 months	572 (3 RCTs) follow up: mean 6 weeks	⊕◯◯◯ VERY LOW [a,e,f]	RD 0.01 (-0.03, 0.06)	51 per 1,000	**10 more per 1,000** (30 fewer to 60 more) [g]	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	754 (3 RCTs) follow up: mean 26 weeks	⊕◯◯◯ VERY LOW [a,b,c]	RR 1.27 (0.61 to 2.64)	74 per 1,000	**20 more per 1,000** (29 fewer to 121 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 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 by 1 or 2 increments because of outcome indirectness

e. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

Outcomes	№ of participants (studies) Follow up	Certainty of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects		Comments
				Risk with sham acupuncture	Risk difference with electroacupuncture	
f. Downgraded by 1 to 2 increments for imprecision due to zero events and small sample size g. Absolute effect calculated by risk difference due to zero events in at least one arm of one study						

Table 7: Clinical evidence summary: Electroacupuncture 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 electroacupuncture	
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months	305 (2 RCTs) follow up: 10 weeks	⨁◯◯◯ VERY LOW [a,b,c]	-	The mean quality of life was 35.6	MD **7.1 higher** (0.44 higher to 13.77 higher)	MID = 3.5 (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	305 (2 RCTs) follow up: mean 10 weeks	⨁⨁◯◯ LOW [a]	-	The mean quality of life was 51.2	MD **2.13 higher** (0.06 higher to 4.19 higher)	MID = 4.5 (0.5 x median baseline SD)
Quality of life (AQoL-SF 36 physical functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 640	MD **78 higher** (21.88 higher to 134.12 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 bodily pain, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⨁◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 207	MD 1**6 higher** (50.37 lower to 78.47 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 role physical, scale range unclear,	60 (1 RCT)	⨁◯◯◯ VERY LOW [a,c]	-	The mean quality of life at	MD **14 higher** (50.47 lower to 78.47 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 electroacupuncture	
high is good, final value) at ≤3 months	follow-up: 2 weeks			<3 months was 233		
Quality of life (AQoL-SF 36 vitality, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 307	MD **30 higher** (2.9 higher to 57.1 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 general health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 312	MD **95 higher** (58.06 higher to 131.94 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 mental health, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 331	MD **40 higher** (5.08 higher to 74.92 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 role emotional, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕◯◯◯ VERY LOW [a,c]	-	The mean quality of life was 207	MD **30 higher** (14.38 lower to 74.38 higher)	MID = 0.5 SD (SMD)
Quality of life (AQoL-SF 36 social functioning, scale range unclear, high is good, final value) at ≤3 months	60 (1 RCT) follow-up: 2 weeks	⊕⊕◯◯ LOW [a]	-	The mean quality of life was 127	MD **45 higher** (31.03 higher to 58.97 higher)	MID = 0.5 SD (SMD)
Pain (WOMAC, 0-20, high is poor, change score and final value) at ≤3 months	295 (2 RCTs) follow up: mean 12 weeks	⊕⊕⊕◯ MODERATE [a]	-	-	MD **3.31 lower** (4.05 lower to 2.57 lower)	MID = 1.6 (0.5 x median baseline SD)
Pain (WOMAC [different scale ranges], high as poor, final values) at ≤3 months	305 (2 RCTs) follow up: mean 12 weeks	⊕◯◯◯ VERY LOW [a,b,c]	-	-	SMD **1.32 SD lower** (2.65 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 electroacupuncture	
Physical function (WOMAC, 0-68, high is poor, change score and final value) at ≤3 months	295 (2 RCTs) follow up: mean 12 weeks	⨁⨁◯◯ LOW [a,c]	-	-	MD **10.17 lower** (12.6 lower to 7.74 lower)	MID = 5.4 (0.5 x median baseline SD)
Physical function (WOMAC [different scale ranges], high as poor, final values) at ≤3 months	305 (2 RCTs) follow up: mean 12 weeks	⨁◯◯◯ VERY LOW [a,b,c]	-	-	SMD **1.37 SD lower** (2.96 lower to 0.21 higher)	MID = 0.5 SD (SMD)
Serious adverse events at ≤3 months	58 (1 RCT) follow up: 12 weeks	⨁◯◯◯ VERY LOW [a,d]	RD 0.00 (-0.06, 0.06)	0 per 1,000	**0 fewer per 1,000** (60 fewer to 60 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 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 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

See Appendix F for full GRADE tables.

1.1.7 Economic evidence

Four health economic studies with the relevant comparison were included in this review.[65, 85, 107, 153] These are summarised in the health economic evidence profile below (**Table 8**) and the health economic evidence tables in Appendix H.

1.1.7.2 Excluded studies

No relevant health economic studies were excluded due to assessment of limited applicability or methodological limitations.

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

1.1.8 Summary of included economic evidence

Table 8: Health economic evidence profile: Acupuncture

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
Latimer 2012[65] (UK)	Directly applicable	Potentially serious limitations [(a)]	• Reanalysis of the NICE osteoarthritis guideline CG59 model wherein separate analyses of three trials were conducted. • Cost-utility analysis (QALYs) • Population: Patients aged 16 year and over with a diagnosis of OA of the knee. • Acupuncture versus usual care • Time horizon was the same as treatment duration in all three studies: 6 months in Berman 2004, 6 weeks in Scharf 2006, and 8 weeks in Witt 2005.	Berman 2004: £414 Scharf 2006: £225 Witt 2005: £216 [(b)]	Berman 2004: 0.024 Scharf 2006: 0.033 Witt 2005: 0.014	Berman, 2004: £17,381 Scharf 2006: £6,911 Witt 2005: £15,621	Using comparison with Sham for effects and comparison with usual care for costs gave the following ICERs: Berman 2004: £40,039 Scharf 2006: £68,284 Witt 2005: £70,519 No further sensitivity analyses conducted.
MacPherson 2017[85] (UK)	Directly applicable	Potentially serious limitations[(c)]	• Probabilistic model based on three separate network meta-analyses of RCTs[(d)] • Cost-utility analysis (QALYs) • Population: Patients reporting pain resulting from OA of the knee. • Comparators:[(e)] • Acupuncture versus usual care	All trials: £179[(f)] Trials with adequate allocation concealment: £192[(f)] Trials with adequate allocation concealment and an end	All trials: 0.014 Trials with adequate allocation concealment: 0.017 Trials with adequate allocation concealment	All trials: £12,786 Trials with adequate allocation concealment: £11,294 Trials with adequate allocation concealment	This study analysed a variety of different intervention classes and so all reports of uncertainty were based on an analysis of all interventions and not any intervention(s) in isolation.

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
			• Time horizon was 8 weeks	point reported at 3-13 weeks: £192[f]	and an end point reported at 3-13 weeks: 0.017	and an end point reported at 3-13 weeks: £11,294 [g]	For a summary of the analysis of uncertainty involving all interventions, see Appendix H.
Reinhold 2008[107] (Germany)	Partially applicable[h]	Potentially serious limitations[i]	• Within-trial analysis. Study is part of the RCT by Witt 2006.[156] • Population: Patients over 40 years of age with chronic pain (defined as more than 6 months) due to osteoarthritis of the knee or hip. • Acupuncture versus delayed acupuncture • Time horizon was 12 months – the treatment duration and follow up were 3 months with treatment effect extrapolated beyond follow up by assuming a linear decline back to baseline at 12 months. • Not stated whether traditional Chinese acupuncture points are used.	£353[j]	0.0241	£13,944	85% probability of being cost effective at a threshold of £20,000 per QALY gained. Sensitivity analysis showed that the parameters which had the largest effect were the cost of acupuncture, and the effect duration.
Whitehurst 2011[153] (UK)	Directly applicable	Potentially serious limitations[k]	• Within-trial analysis based on the RCT by Foster 2007.[37] • Population: Patients aged 50 years or older who had	£85[l,m]	0.022	£3,889	77% probability of being cost effective at a threshold of £20,000 per QALY gained.

Study	Applicability	Limitations	Other comments	Incremental cost	Incremental effects	Cost effectiveness	Uncertainty
			been referred to NHS physical therapy centres with a clinical diagnosis of knee OA. • Acupuncture + advice and exercise versus advice and exercise • Treatment duration was 6 weeks, but patients were followed up for 12 months. • Traditional Chinese acupuncture points were used.				The following sensitivity analyses were carried out (using data from the sample of participants who had complete resource use and EQ-5D data): 1. A complete case analysis explored the implication of missing data. This had a cost per QALY of £2,278. 2. A cost perspective that incorporated non-NHS health care resource use (had little effect on the results). 3. An analysis of the AE+NPA group within the base case. AE+NPA was associated with an additional cost of £36 and an incremental QALY of 0.001 compared with AE+A. Thus ICER (Intervention 3 vs. Intervention 2): £36/ 0.001 = £36,000.

Abbreviations: CG59= clinical guideline 59; ICER= incremental cost-effectiveness ratio; NHS= National Health Service; OA= osteoarthritis; QALY= quality-adjusted life years; RCT= randomised controlled trial

(a) 2010 resource use and unit costs may not reflect current UK NHS practice. The time horizon varies across studies and the results did not include any extrapolations outside the trial settings. The costs of sham acupuncture were assumed to be zero, which is not reflective of 'real-world' costs. Adverse events and their downstream consequences were not considered. Sensitivity analysis of the results were not conducted.

(b) 2010 UK pounds. Cost components: Intervention costs only (physiotherapists time and the cost of acupuncture needles).

(c) Unit costs taken from 2011/12 may not reflect current UK NHS practice. The time horizon was only 8 weeks. Adverse events and their downstream consequences were not considered.

(d) The three network meta-analyses were: 1) an analysis involving all trials; 2) an analysis including only trials with adequate allocation concealment and 3) an analysis including only trials with adequate allocation concealment and a reported end-point between 3-13 weeks. See Appendix H for all model results.

(e) See Appendix H for all model results.

(f) The original report listed 13 interventions in total. Only those interventions that fit the protocol for acupuncture were included here. Please note intervention numbers in this profile do not match to intervention numbers in evidence table (Appendix H).

(g) 2011/12 UK pounds. Cost components incorporated: Physiotherapist's time to conduct sessions. Changes in non-treatment-related visits to GPs and specialists arising from changes to EQ-5D score

(h) In a full incremental analysis of all interventions, TENS was the most cost-effective option in the network meta-analysis all trials with a cost per QALY of £2,690. In the other two network meta-analyses (1. only those trials with adequate allocation concealment and 2. only those trials with adequate allocation concealment and an endpoint between 3-13 weeks), acupuncture was the most cost-effective option with costs per QALYs of £13,502 and £14,275, respectively.

(i) Study took a German perspective; therefore, costs may not be applicable to UK. SF-36 scores were mapped to the SF-6D, rather than EQ5D.

(j) Short time horizon (3 months). Cost of lost workdays were included in the cost of intervention.

(k) 2006 Euros converted to UK pounds.[101]. Cost components incorporated: Acupuncture costs, physician visits, medication, hospital stays (and indirect costs of lost workdays).

(l) 2004/05 resource use and unit costs may not reflect current UK NHS practice. Time horizon could be longer to capture any longer-term health effects. Study relies on patients to recall healthcare usage.

(m) 2004-2005 UK pounds. Costs components: consultations with primary care-based practitioners, hospital consultants (outpatient attendance) or any other health care provider. Participants were also asked to report any prescribed medications and over the counter purchases) and this commentary could be added as discussion below the table rather than as a footnote.

(n) The incremental cost in Whitehurst is lower (despite the fact that it has a longer time horizon) than the Reinhold analysis because of the comparator. Since it does not cost much more to incorporate acupuncture into the advice and exercise sessions, the cost difference between the two groups is small (as the cost of sessions is the main driver of total costs). Also, the length of treatment in Whitehurst 2011 was only 6 weeks (Reinhold 2008 was 3 months), and the difference in resource use between the two groups over the 12 months.

1.1.9 Economic model

1.1.9.1 Population and strategies evaluated

The modelled population were adults aged 16 years and over with osteoarthritis of the knee. The strategies compared were:

- Electroacupuncture
- Usual care

Acupuncture was not included as the clinical evidence showed no clinically important benefit along with a clinically important harm in serious adverse events versus sham, and an unclear benefit in the short term versus no treatment.

1.1.9.2 Methods and data sources (Summary)

- Each treatment was assumed to have an immediate impact on quality of life (direct EQ-5D valuations were not available so were mapped from SF-12/36 and WOMAC scores). These were estimated from four randomised trials comparing electroacupuncture with some form of usual care.
- Due to heterogeneity observed between trials in the type of electroacupuncture device used, the voltage given as well as the pressure points selected for needling, it was decided that base case results would consist of individual trials as well as pooled estimates of effect using both a weighted average and an unweighted average.
- For the base case, the improvement in EQ-5D was 0.171, 0.107 and 0.0.098 after 12 weeks in Berman 1999[10], Dunning 2018[26], and Suarez Almazor 2010[124], respectively. The improvement in EQ-5D was 0.163 after 8 weeks in Mavrommatis 2012.[92]
- It was assumed that treatment effect with electroacupuncture would persist up to 26 weeks. Since data from trials were available to either of 8 or 12 weeks, the treatment effect was extrapolated by assuming a linear decline back to baseline from weeks 12 to 26. In the study that reported up to week 8, the treatment effect at week 8 was assumed to remain constant up to week 12.
- The incremental change in quality adjusted life years (QALYs) were 0.057, 0.043, 0.048 and 0.031 from Berman 1999[10], Dunning 2018[26], Mavrommatis 2012[92], and Suarez Almazor 2010[124], respectively.
- In a pooled analysis, the weighted incremental change in QALY was calculated 0.041 and the unweighted incremental change in QALY was calculated as 0.045.
- Mortality was not impacted by treatment and there were no adverse effects modelled.
- A model time horizon of one year was chosen as this was sufficiently long to capture the costs and treatment effects. Mortality is not affected by treatment.

1.1.9.3 Costs

- It was assumed that electroacupuncture would be carried out in 1-1 sessions with a band 6 community physiotherapist. The length of sessions as well as their frequency were based on the data from the four trials. A weighted average of the trials calculated that each session lasted for 25 minutes with a frequency of 1.92 sessions per week over 7 weeks.
- The cost of the electroacupuncture device and consumables were taken from online sources. An ES-160 electroacupuncture device was used in the model base case - £395. The device costs were annuitized using a discount rate of 3.5% and assuming the equipment is replaced after 5 years.
- In addition to the device the following costs were included:
    - Crocodile clips (£39.50).
    - Lead cables (£141.20)
    - Four batteries to power the ES-160 device with a lifespan of 18 hours (£1.38)

- Costs specific to the sessions included:
    - 10 copper-plated needles (£0.80
    - 1 disinfectant swab (£0.30)
    - 1 pair of examination gloves (£0.12)

The key outcomes were mean NHS cost per patient and mean QALYs per patient. These were calculated using a simple area under the curve model approach. Only incremental costs and QALYs were calculated. The results were calculated both:

- Deterministically, based on the point estimates of each input parameter
- Probabilistically (for QALYs only), based on a distribution for each input parameter (estimated using its standard error) and sampling the results 5,000 times before calculating a mean (Monte Carlo simulation).

1.1.9.4 Results

The base case results can be found in Table 9 and show that showed that both the probabilistic and deterministic costs per QALY for electroacupuncture versus usual care were below the NICE cost effectiveness threshold of £20,000 per QALY gained (£7,504 and £7,209).

When the individual trials were scrutinised (Table 12, Table 13, Table 14 & Table 15), all showed that electroacupuncture was cost effective versus usual care. This trend was also observed in the results of the sensitivity analyses where electroacupuncture was cost effective versus usual care, except when it was delivered by a GP.

Table 9. Base case results– incremental EA vs. UC (probabilistic and deterministic)

Base case	Analysis	Incremental cost	Incremental QALYs	Cost per QALY gained	Probability cost effective at £20k	Probability cost effective at £30k
Pooled trials (weighted average)	Probabilistic	£296	0.039	£7,504	97%	99%
	Deterministic	£296	0.041	£7,209	NA	NA
Berman 1999	Probabilistic	£295	0.039	£7,641	62%	71%
	Deterministic	£295	0.057	£5,163	NA	NA
Dunning 2018	Probabilistic	£268	0.027	£10,098	59%	68%
	Deterministic	£268	0.043	£6,217	NA	NA
Mavrommatis 2012	Probabilistic	£295	0.098	£3,010	99%	99%
	Deterministic	£295	0.048	£6,204	NA	NA
Suarez Almazor 2010	Probabilistic	£323	0.031	£10,267	99%	99%
	Deterministic	£323	0.031	£10,314	NA	NA

Abbreviations: EA=electroacupuncture; NA=not applicable; NT=Usual care; QALYs=quality adjusted life years

Table 10. Pooled trials (costs and QALYs calculated based on a weighted average)

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£296	0.039	£7,504	97%	99%
Time horizon					
SA1 3-month time horizon	£296	0.024	£12,581	86%	97%
SA2 Booster sessions	£497	0.094	£5,259	99%	99%
Costs					
SA3 Group sessions	£207	0.039	£5,258	99%	99%
SA4 Using AS-super 4 device	£289	0.039	£7,353	98%	99%
SA4 Band 5 physiotherapist	£230	0.040	£5,801	98%	99%
SA6 Band 7 physiotherapist	£350	0.040	£8,830	97%	99%
SA7 GP	£858	0.039	£21,833	34%	75%
Utilities					
SA8 Alternative utilities (Barton)	£296	0.041	£7,298	79%	84%
SA9 Alternative utilities (Lawrence)	£296	0.050	£5,857	100%	100%
SA10 Alternative utilities (Price)	£296	0.031	£9,604	86%	94%
SA11 Alternative utilities (Maund)	£296	0.031	£9,420	56%	58%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

Table 11. Pooled trials (costs and QALYs calculated based on an unweighted average)

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£296	0.049	£6,068	79%	85%
Time horizon					
SA1 3-month time horizon	£296	0.029	£10,222	67%	79%
SA2 Booster sessions	£492	0.116	£4,245	85%	88%
Costs					
SA3 Group sessions	£213	0.049	£4,383	83%	87%
SA4 Using AS-super 4 device	£288	0.048	£5,946	79%	85%
SA4 Band 5 physiotherapist	£231	0.049	£4,714	83%	87%
SA6 Band 7 physiotherapist	£350	0.049	£7,148	76%	83%
SA7 GP	£853	0.048	£17,612	41%	59%
Utilities					
SA8 Alternative utilities (Barton)	£296	0.052	£5,723	79%	81%
SA9 Alternative utilities (Lawrence)	£296	0.060	£4,934	99%	100%
SA10 Alternative utilities (Price)	£296	0.033	£8,847	75%	82%
SA11 Alternative utilities (Maund)	£296	0.036	£8,189	57%	62%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

Table 12. Berman 1999 trial

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£295	0.039	£7,641	62%	72%
Time horizon					
SA1 3-month time horizon	£295	0.025	£11,989	52%	63%
SA2 Booster sessions	£483	0.098	£4,934	72%	77%
Costs					
SA3 Group sessions	£227	0.039	£5,851	68%	75%
SA4 Using AS-super 4 device	£287	0.038	£7,598	63%	71%
SA4 Band 5 physiotherapist	£231	0.039	£5,940	68%	74%
SA6 Band 7 physiotherapist	£349	0.039	£8,901	58%	68%
SA7 GP	£845	0.038	£22,292	34%	46%
Utilities					
SA8 Alternative utilities (Barton)	£295	0.052	£5,716	65%	67%
SA9 Alternative utilities (Lawrence)	£295	0.061	£4,861	100%	100%
SA10 Alternative utilities (Price)	£295	0.038	£7,858	62%	70%
SA11 Alternative utilities (Maund)	£295	0.039	£7,664	63%	71%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

Table 13. Dunning 2018 trial

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£268	0.027	£10,098	59%	68%
Time horizon					
SA1 3-month time horizon	£268	0.016	£16,824	44%	56%
SA2 Booster sessions	£468	0.064	£7,289	67%	75%
Costs					
SA3 Group sessions	£194	0.027	£7,320	66%	74%
SA4 Using AS-super 4 device	£263	0.026	£9,936	58%	68%
SA4 Band 5 physiotherapist	£208	0.027	£7,612	65%	73%
SA6 Band 7 physiotherapist	£318	0.027	£11,681	54%	65%
SA7 GP	£783	0.027	£29,184	27%	40%
Utilities					
SA8 Alternative utilities (Barton)	£268	0.025	£10,656	54%	57%
SA9 Alternative utilities (Lawrence)	£268	0.049	£5,430	100%	100%
SA10 Alternative utilities (Price)	£268	0.026	£10,230	58%	67%
SA11 Alternative utilities (Maund)	£268	0.027	£9,898	59%	69%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

Table 14. Mavrommatis 2012 trial

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£295	0.098	£3,010	100%	100%
Time horizon					
SA1 3-month time horizon	£295	0.056	£5,272	100%	100%
SA2 Booster sessions	£483	0.225	£2,149	100%	100%
Costs					
SA3 Group sessions	£227	0.098	£2,308	100%	100%
SA4 Using AS-super 4 device	£287	0.098	£2,926	100%	100%
SA4 Band 5 physiotherapist	£231	0.098	£2,364	100%	100%
SA6 Band 7 physiotherapist	£349	0.098	£3,563	100%	100%
SA7 GP	£845	0.098	£8,632	100%	100%
Utilities					
SA8 Alternative utilities (Barton)	£295	0.098	£3,011	100%	100%
SA9 Alternative utilities (Lawrence)	£295	0.098	£3,010	100%	100%
SA10 Alternative utilities (Price)	£295	0.040	£7,392	95%	98%
SA11 Alternative utilities (Maund)	£295	0.053	£5,563	59%	60%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

Table 15. Suarez Almazor 2010 trial

Analysis	Mean difference (EA-UC)		ICER (Cost per QALY gained)	Probability cost effective at £20k	Probability cost effective at £30k
	Inc. cost	Inc. QALY			
Base case results	£323	0.031	£10,267	96%	99%
Time horizon					
SA1 3-month time horizon	£323	0.019	£16,970	72%	96%
SA2 Booster sessions	£534	0.077	£6,959	100%	100%
Costs					
SA3 Group sessions	£206	0.031	£6,585	99%	100%
SA4 Using AS-super 4 device	£316	0.032	£10,019	96%	99%
SA4 Band 5 physiotherapist	£251	0.031	£7,998	98%	100%
SA6 Band 7 physiotherapist	£383	0.031	£12,194	92%	98%
SA7 GP	£941	0.031	£30,117	3%	49%
Utilities					
SA8 Alternative utilities (Barton)	£323	0.032	£10,231	96%	99%
SA9 Alternative utilities (Lawrence)	£323	0.031	£10,324	95%	99%
SA10 Alternative utilities (Price)	£323	0.030	£10,827	83%	91%
SA11 Alternative utilities (Maund)	£323	0.026	£12,619	48%	49%

Abbreviations: EA=electroacupuncture; UC=Usual care; QALYs=quality adjusted life years

1.1.11 Economic evidence statements

- One cost–utility analysis found that acupuncture plus advice and exercise was cost effective compared to just advice and exercise for treating knee osteoarthritis (ICER: £3,889 per QALY gained). This analysis was assessed as directly applicable with minor limitations.
- One cost–utility analysis found that acupuncture was cost effective for treating knee osteoarthritis compared to
    - education sessions, (£17,381 per QALY gained)
    - conservative pharmacological therapy (£6,911 per QALY gained) and
    - delayed acupuncture (£15,621 per QALY gained).

This analysis was assessed as directly applicable with potentially serious limitations.

- One cost-utility analysis reported that acupuncture was cost effective compared with usual care (ICER: £12,786. This analysis was assessed as partially applicable with potentially serious limitations. A full incremental analysis of various non-pharmacological interventions (acupuncture, braces, heat treatment, insoles, interferential therapy, laser/light therapy, manual therapy, neuromuscular electrical stimulation, pulsed electromagnetic field, pulsed electrical stimulation, static magnets and transcutaneous electrical nerve stimulation) also reported that acupuncture was the most cost-effective strategy in two of the three network meta-analyses (£13,502 and 14,275), with transcutaneous electrical nerve stimulation the most cost-effective option in the other (£2,690).
- One original-cost utility analysis reported that electroacupuncture was cost effective compared to usual care. (ICER: £7,504 in an analysis of pooled trials and between £3,010 and £10,267 in an analysis of individual trials). This analysis was assessed as directly 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 flares 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 a treatment adverse event 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. Serious adverse events were defined by the studies included, and if no other evidence was reported any adverse events reported by the study was included in the outcome.

There was evidence available for all outcomes. However, while some was available, there was only limited evidence available for osteoarthritis flares, psychological distress and serious adverse events throughout the literature.

1.1.12.2 The quality of the evidence

Twenty-six randomised controlled trial studies were included in this review. There was limited evidence comparing electroacupuncture to acupuncture (1 study). However, evidence was available for all other comparisons.

The evidence varied from high to very low quality. The majority of evidence when compared to sham procedures was of moderate quality, while the majority of evidence when compared to no treatment was of very low quality. Outcomes were commonly downgraded for inconsistency and imprecision. When present, inconsistent results were not explained by subgroup analysis because the majority of studies fell into the same subgroup (studies with people of an average age less than or equal to 75 with knee osteoarthritis, diagnosed with imaging, and unclear information about the presence of multimorbidities). Where examined, there was no difference seen between conventional acupuncture and dry needling techniques. When compared to no treatment, outcomes were commonly downgraded for risk of bias (in particular for risk of performance bias).

The type of sham acupuncture used varied across studies. The variations in technique included changes in: the type of needle used (including non-penetrating needles), depth of needle insertion, position of needle, in situ stimulation of needle and the application of electrical stimulation. Some studies would vary one parameter, while others would change multiple of these factors at the same time. This could be a potential source of heterogeneity across all studies and contributes to the uncertainty seen in the results.

The committee agreed that there was sufficient evidence to compare acupuncture and electroacupuncture to sham procedures and no treatment. There was limited evidence for the comparison of electroacupuncture and acupuncture. However, the committee agreed they could consolidate and draw conclusions about the respective techniques through their efficacy against sham procedures and no treatment.

Acupuncture/dry needling compared to sham acupuncture

The evidence for acupuncture compared to sham acupuncture varied from high to very low quality, with the majority of the outcomes being of high quality. Outcomes of low and very low quality were downgraded due to risk of bias, population indirectness (as participants were required to have myofascial trigger points), imprecision and inconsistency, with heterogeneity that could not be resolved by subgroup analysis.

Acupuncture/dry needling compared to no treatment

The evidence for acupuncture compared to no treatment varied from moderate to very low quality, with the majority of the outcomes being of very low quality. Outcomes were commonly downgraded due to risk of bias, imprecision and inconsistency, with heterogeneity that could not be resolved by subgroup analysis.

Electroacupuncture compared to acupuncture

The evidence for electroacupuncture compared to acupuncture varied from high to low quality, with the majority of outcomes being of low quality. Outcomes were commonly downgraded due to risk of bias and imprecision.

Electroacupuncture compared to sham acupuncture

The evidence for electroacupuncture compared to sham acupuncture varied from high to very low quality, with the majority of outcomes being of very low quality. Outcomes were commonly downgraded due to risk of bias, imprecision and inconsistency, with heterogeneity that could not be resolved by subgroup analysis. One outcome was downgraded for outcome indirectness as the outcome reported osteoarthritis flares but only events that cause withdrawal from the trial.

Electroacupuncture compared to no treatment

The evidence of electroacupuncture compared to no treatment varied from moderate to very low quality, with the majority of outcomes being of very low quality. Outcomes were commonly downgraded due to risk of bias, imprecision and inconsistency, with heterogeneity that could not be resolved by subgroup analysis.

1.1.12.3 Benefits and harms

Key uncertainties

Comparing acupuncture to sham acupuncture is challenging due to the potential for sham acupuncture to have an active effect on the outcomes investigated beyond a placebo effect. In this review, the types of sham acupuncture used were not consistent between studies, with variations including: using different needle positions, different sizes of needles, different depths of needle insertion, using retractable needles, the absence of electrical stimulation or the presence of mock electrical stimulation (for example: using alternating currents and lower voltages). This adds to the uncertainty of the results when compared to sham acupuncture. This is further complicated by the complex nature of acupuncture treatment, including the effect of the relationship between the health care professional and the person with osteoarthritis. The committee weighed up the benefits and limitations of comparing to sham and no treatment throughout this intervention. They balanced this against their consideration of other interventions in the guideline and noted that the involvement of the participant in the procedure was important. While exercise was an active treatment, requiring direction and participation from the person with osteoarthritis, acupuncture (and other treatments such as manual therapy and pharmacotherapy) was passive requiring the person to have treatment applied to them by another person. The committee noted that this made comparing the different treatments more difficult and considered this an additional uncertainty in their analysis.

The committee acknowledged the limitations in the adverse event data identified in this review. 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. The outcome included adverse events that could be considered minor in nature and only limited evidence was available investigating serious adverse events. From experience, the committee agreed that the likely adverse events were bruising, bleeding and brief pain worsening after stimulation. Given the importance of these events, if they were identified this would change the ultimate decision making. The committee agreed that future research in this area should ensure to investigate and transparently report adverse events data that may provide more clarification around this.

The populations included in the studies were, in general, younger people with knee or hip osteoarthritis who were not noted to have significant multimorbidity. Therefore, from the evidence included, it is unclear whether acupuncture has an effect for osteoarthritis affecting other joints (for example: hand, foot) or whether there are subgroups of the population that may respond differently to acupuncture. It was noted by clinicians with experience in the technique, that acupuncture may be useful for a subset of people with osteoarthritis (for

example: people who have not responded well to pharmacological management who are trying to exercise but have too much pain so that they feel that this is not possible).

Acupuncture compared to sham acupuncture

The results showed that, when compared to sham acupuncture, acupuncture led to no clinically important difference in quality of life, pain, physical function and psychological distress at less than and equal to 3 months and more than 3 months and serious adverse events at more than 3 months only, with a clinically important harm in serious adverse events at less than and equal to 3 months.

The committee noted the complexity in examining the efficacy of acupuncture. They agreed that acupuncture was a complex intervention with multiple factors that could influence the outcome, including: acupuncture point specific effects, physiological effects due to skin penetration and patient-practitioner interaction and beliefs/expectations. The variety of sham confounds the interpretation of the results due to this, as different techniques could influence these results (for example: non-penetrating sham acupuncture may not provide skin penetration and so reduce the physiological effects in the way that a penetrating sham would). In this way, the committee agreed that the effect of acupuncture is more complicated to examine using sham then for other conditions (for example: pharmacological management and placebo tablets).

On examining the evidence, the committee agreed that there was insufficient evidence of benefit from acupuncture with the potential for harm. The adverse events varied, but more commonly included increase in pain, bleeding and bruising and fatigue. The committee noted that while there were clinically important harms for serious adverse events, this data was based on a small number of studies and was heterogenous with some studies including zero events in both study arms. The committee agreed that, in general, observational trials may provide higher quality of evidence for adverse events than that in randomised controlled trials. On reflecting on their experience, the type of adverse events reported and the uncertainty in the evidence, they agreed that acupuncture is unlikely to cause significant long-term adverse events. Given the lack of benefit seen from acupuncture when compared to sham acupuncture, the health economic model did not include acupuncture.

Acupuncture compared to no treatment

The results showed that, when compared to no treatment, acupuncture led to an unclear effect in the short term for quality of life (with 1 outcome showing a clinically important benefit, and 2 outcomes showing no clinically important difference), pain (with 1 outcome showing a clinically important benefit, and 1 outcome showing no clinically important difference) and physical function (with 1 outcome showing a clinically important benefit, and 1 outcome showing no clinically important difference). Additionally, there was no clinically important difference in psychological function and serious adverse events in the short term (≤3 months). However, in the long term (>3 months) each of these outcomes showed no clinically important difference with acupuncture when compared to no treatment.

As with sham acupuncture, the committee noted the limitations of examining acupuncture against no treatment. They agreed that given the complex nature of the intervention, there was likely to be a more pronounced benefit from acupuncture when compared to no treatment in comparison to sham, and that this benefit will include factors beyond the insertion of needles. This was further complicated by the variety of definitions of no treatment included for this comparison, ranging from supervised exercise therapy available to both study arms to waiting list controls. The committee noted that the presence of health care professional input, in particular the relationship between the practitioner and the participant, may lead to a substantial effect and so studies that do not include this in the control arm are likely to see larger benefits from acupuncture.

The committee noted that where clinically important benefits were seen in quality of life, pain and physical function, the was significant heterogeneity with the quality of the outcomes being very low. They noted that the studies driving the positive effect were larger studies. However, given the conflicting information from a larger number of other outcomes of higher quality (generally moderate quality) showing no clinically important effect they were unconvinced by this benefit.

Electroacupuncture compared to acupuncture

The results showed that, when compared to acupuncture, electroacupuncture led to no clinically important differences in quality of life, pain and physical function at less than and equal to 3 months and more than 3 months, and at serious adverse events at more than 3 months only.

The committee noted the limited evidence for this comparison, being reported in only one study. On conclusion, weighing up the limited evidence and the evidence seen with sham and no treatment comparisons, the committee did not make any additional recommendations based purely on this data.

Electroacupuncture compared to sham acupuncture

The results showed that, when compared to sham acupuncture, electroacupuncture led to a clinically important benefit in the short term for pain (with 2 outcomes of very low quality including 9 studies showing a clinically important benefit) and an unclear possible benefit for physical function (with 1 outcome of very low quality including 6 studies showing a clinically important benefit, and 1 outcome of moderate quality including 2 studies showing no clinically important difference). Additionally, there was no clinically important difference in quality of life, psychological distress, osteoarthritis flares and serious adverse events in the short term. However, in the long term there was no clinically important difference in quality of life, pain, physical function, and serious adverse events.

The committee again noted the complexity of examining electroacupuncture against sham due to the complex nature of the intervention (as with acupuncture in the sections above). In addition, for electroacupuncture, some studies only varied the presence of electrical stimulation which effectively compared electroacupuncture to acupuncture. As the study identified this as sham acupuncture, the study was included under this comparison, rather than a head-to-head comparison of electroacupuncture and acupuncture. The committee considered this while examining the evidence.

The committee noted that the evidence for benefit for electroacupuncture was generally heterogenous with the quality of the outcomes being very low. Where benefit was seen in pain at less than 3 months, the effect was driven by two studies with around 80 participants in each study. This was compared to three studies, which included between 80 and 450 participants between the studies. However, while the quality of the evidence was the same for the physical function outcome, the results were driven by three studies instead. The committee further noted that there was no evidence of harm for the procedure. The adverse events reported included heart disease, cancer, non-study related injuries, non-arthritis related surgery, stroke and pneumonia. The committee agreed that, in general, randomised controlled trials do not provide the highest quality of evidence for adverse events and that long-term observational studies may provide more information for this area. On reflecting on their experience, they agreed that acupuncture is unlikely to cause significant long-term adverse events.

Electroacupuncture compared to no treatment

The results showed that, when compared to no treatment, electroacupuncture led to a short term clinically important benefit in pain (based on 2 outcomes of moderate-very low quality including 4 studies) and physical function (based on 2 outcomes of low-very low quality

including 4 studies), an unclear potential clinically important benefit in quality of life (with 1 outcome of very low quality including 2 studies showing a clinically important benefit, and 1 outcome of moderate quality including 2 studies showing no clinically important difference), and no clinically important difference in serious adverse events. There was no information in the long-term effects.

The committee noted that there was less evidence for this comparison than others (including only 4 studies). As with acupuncture, the committee noted the limitations of examining electroacupuncture against no treatment for the same reasons. In particular for this review, the committee noted that the definition of no treatment was fairly rigorous for some studies, including manual therapy and exercise available to all participants in one, and etoricoxib and proton pump inhibitors available to all participants in another.

The committee noted the absence of long term data for this comparison. They agreed the benefits were seen for multiple outcomes with the quality being variable (in some outcomes there was significant heterogeneity, while in others there was no heterogeneity and the clinically important benefit was retained). Where heterogeneity was present, the values of both studies still lay in the range of a clinically important benefit according to the minimally important difference values used in the analysis. There were no adverse events present in the one study that reported this.

Weighing up the clinical benefits and harms

On considering the evidence of benefit, the committee discussed the quality of the evidence. The committee agreed that due to a lack of clinical efficacy and potential for harm, that acupuncture was not likely to be clinically effective for use for people with osteoarthritis. Therefore, they recommended against the use of acupuncture or dry needling for the management of osteoarthritis.

The committee considered electroacupuncture. While there were signs of benefit, this was mostly based on very low quality data. This was conflicted by the evidence comparing electroacupuncture to acupuncture where no clinically important difference was seen. These studies also included a small sample size. Due to this, the committee agreed that there was insufficient evidence to show the benefits of electroacupuncture. However, there did not appear to be any significant harms from electroacupuncture. Therefore, they agreed recommendation 1.3.8 would be most appropriate for providing information to people with osteoarthritis and practitioner. However, they acknowledged that the absence of consistent evidence may be because there are specific groups of people with osteoarthritis who may respond more to electroacupuncture than others. Due to this the committee agreed a research recommendation to further investigate populations where electroacupuncture may be more appropriate.

1.1.12.4 Cost effectiveness and resource use

The committee discussed the correct comparator for evaluations of acupuncture. In economic evaluations, comparisons with usual care or no treatment or an alternative active comparator are usually considered most relevant for assessing the real-world impact of an intervention on resource use and QALYs. It was decided that for acupuncture to be recommended there should be:

- a clinical benefit compared with both sham and usual care, and
- cost effectiveness compared with usual care.

Comparing acupuncture to usual care is the most common approach to assessing its cost-effectiveness and this approach has been taken on the NICE guidelines on low back pain (NG59) and management of primary chronic pain (NG193).

Four economic studies were included in the review. One was based on a network meta-analysis of randomised controlled trials (RCTs); one was based on three separate RCTs and the other two were based on individual RCTs. All were in an osteoarthritis population.

One study took a German perspective and therefore included lost workdays resulting from illness within costs. QALYs were captured using the SF-6D. The trial period was three months with data extrapolated from three months onwards to 12 months. It was the only study to extrapolate costs and outcomes beyond the trial period. This study was graded as being partially applicable with minor limitations.

The other three studies took a UK perspective and calculated QALYs using the EQ-5D measure; the first UK study collected this data directly using the EQ-5D questionnaire, the second calculated QALYs separately for three different trials by mapping from the WOMAC index to EQ-5D, while the third calculated the QALYs by mapping from various measures to EQ-5D and then pooling the results to give an overall estimate. All three studies were deemed directly applicable. The first study was based on a single trial and was judged to have minor limitations. The second study was based on three trials, with a separate analysis for each. The time horizon in all three trials was short (2 to 6 months) and sensitivity analysis was not conducted. For these reasons, it was deemed to have potentially serious limitations. The final study was based on a network meta-analysis where the model time horizon was relatively short at 8 weeks. The unit costs were taken from 2011/12 and were therefore unlikely to be representative of current NHS practice. For these reasons, it was graded as having potentially serious limitations.

Although only one of the studies estimated cost-effectiveness from a pooled estimate of effect, the results were consistent both from a pooled estimate and from across five separate randomised trials that acupuncture was cost effectiveness compared with usual care (£12,786 per QALY gained). In the study conducted with a pooled estimate of effect, a full incremental analysis of various non-pharmacological interventions (acupuncture, braces, heat treatment, insoles, interferential therapy, laser/light therapy, manual therapy, neuromuscular electrical stimulation, pulsed electromagnetic field, pulsed electrical stimulation, static magnets and transcutaneous electrical nerve stimulation) showed that acupuncture was the most cost effective strategy in analyses of trials limited to those with adequate allocation concealment and also trials with adequate allocation concealment with an endpoint between 3-13 weeks with costs per QALYs of £13,502 and 14,275, respectively. A third full incremental analysis that included all trials reported that transcutaneous electrical nerve stimulation was the most cost-effective strategy (£2,690 per QALY gained).

An original cost-utility analysis was developed specifically for electroacupuncture. It was based on 4 trials from the clinical review comparing electroacupuncture with usual care. This model also showed the cost per QALY gained to be £7,504 in an analysis of pooled trials and ranging between £3,010 and £10,267 in an analysis of individual trials.

Although, the cost-effectiveness evidence appeared to support the use of acupuncture, the guideline's clinical review did not show a convincing clinical effect for acupuncture over sham acupuncture. For electroacupuncture versus sham, there was a significant improvement in one outcome, but the committee decided that the evidence was not robust enough to be confident that there was an effect beyond the placebo effect. Given the uncertainty in the results compared with sham and potential for a large resource impact, the committee decided not to recommend standard acupuncture or electroacupuncture.

1.1.12.5 Other factors the committee took into account

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 sub-grouped 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 recommendation 1.3.8 and the research recommendation on acupuncture and electroacupuncture. Other evidence supporting these recommendations can be found in evidence review F.

1.1.14 References

1. Ahsin S, Aslam M, Saleem S, Bhatti A. Role of electroacupuncture in treatment of osteoarthritis knee: a single blind sham controlled study. Rawal Medical Journal. 2010; 35(1):41-44

2. Ahsin S, Saleem S, Bhatti AM, Iles RK, Aslam M. Clinical and endocrinological changes after electro-acupuncture treatment in patients with osteoarthritis of the knee. Pain. 2009; 147(1-3):60-66

3. Al Rashoud AS, Abboud RJ, Wang W, Wigderowitz C. Efficacy of low-level laser therapy applied at acupuncture points in knee osteoarthritis: a randomised double-blind comparative trial. Physiotherapy. 2014; 100(3):242-248

4. Ammer K, Petschnig R. Comparison of the effectiveness of acupuncture and physical therapy in ambulatory patients with gonarthrosis. Wiener medizinische wochenschrift (1946). 1988; 138(22):566-569

5. Appleyard I, Crichton N, Robinson N. Warm needle acupuncture vs. needle acupuncture for osteoarthritis of the knee: a pilot study protocol. European journal of integrative medicine. 2016; 8(4):407-413

6. Ashraf A, Zarei F, Hadianfard MJ, Kazemi B, Mohammadi S, Naseri M et al. Comparison the effect of lateral wedge insole and acupuncture in medial compartment knee osteoarthritis: a randomized controlled trial. Knee. 2014; 21(2):439-444

7. Atalay SG, Durmus A, Gezginaslan O. The effect of acupuncture and physiotherapy on patients with knee osteoarthritis: A randomized controlled study. Pain Physician. 2021; 24(3):E269-E278

8. Bao F, Wu Z. Observation on therapeutic effect of knee osteoarthritis treated by electroacupuncture. International journal of clinical acupuncture. 2007; 16(3):191-195

9. Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. Annals of Internal Medicine. 2004; 141(12):901-910

10. Berman BM, Singh BB, Lao L, Langenberg P, Li H, Hadhazy V et al. A randomized trial of acupuncture as an adjunctive therapy in osteoarthritis of the knee. Rheumatology. 1999; 38(4):346-354

11. Boswell DJ, Ostergaard K, Philipson RS, Hodge RA, Blum D, Brown JC et al. Evaluation of GW406381 for treatment of osteoarthritis of the knee: two randomized, controlled studies. Medscape Journal of Medicine. 2008; 10(11):259

12. Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Hummelsberger J et al. Physician and treatment characteristics in a randomised multicentre trial of acupuncture in patients with osteoarthritis of the knee. Complementary Therapies in Medicine. 2007; 15(3):180-189

13. Brinkhaus B, Witt CM, Linde K, Streng A, Melchart D, Willich SN. Efficacy of acupuncture in patients with osteoarthritis of the knee: a ramdomized controlled trial. Chirurgische praxis. 2006; 66(2):311-321

14. Byun H, Kim SW, Ahn JH, Kim YS, Seo JC, Choi SM. Individualized acupuncture versus standardized acupuncture in symptomatic treatment of osteoarthritis of the knee-a randomized controlled trial (isrctn 40706107). The journal of korean acupuncture & moxibustion society= taehan chimgu hakhoe chi. 2007; 24(4):183-195

15. Cao L, Zhang XL, Gao YS, Jiang Y. Needle acupuncture for osteoarthritis of the knee. A systematic review and updated meta-analysis. Saudi Medical Journal. 2012; 33(5):526-532

16. Ceballos-Laita L, Jimenez-Del-Barrio S, Marin-Zurdo J, Moreno-Calvo A, Marin-Bone J, Albarova-Corral MI et al. Effectiveness of dry needling therapy on pain, hip muscle strength, and physical function in patients with hip osteoarthritis: A randomized controlled trial. Archives of Physical Medicine and Rehabilitation. 2021; 102(5):959-966

17. Ceballos-Laita L, Jimenez-Del-Barrio S, Marin-Zurdo J, Moreno-Calvo A, Marin-Bone J, Albarova-Corral MI et al. Effects of dry needling in HIP muscles in patients with HIP osteoarthritis: A randomized controlled trial. Musculoskeletal Science & Practice. 2019; 43:76-82

18. Ceballos-Laita L, Jimenez-Del-Barrio S, Marin-Zurdo J, Moreno-Calvo A, Marin-Bone J, Albarova-Corral MI et al. Effects of dry needling on pain, pressure pain threshold and psychological distress in patients with mild to moderate hip osteoarthritis: Secondary analysis of a randomized controlled trial. Complementary Therapies in Medicine. 2020; 51:102443

19. Chen LX, Mao JJ, Fernandes S, Galantino ML, Guo W, Lariccia P et al. Integrating acupuncture with exercise-based physical therapy for knee osteoarthritis: a randomized controlled trial. JCR: Journal of Clinical Rheumatology. 2013; 19(6):308-316

20. Chen N, Wang J, Mucelli A, Zhang X, Wang C. Electro-acupuncture is beneficial for knee osteoarthritis: The evidence from meta-analysis of randomized controlled trials. American Journal of Chinese Medicine. 2017; 45(5):965-985

21. Chen RL, Deng KF, Wei XC, Wang HS, Feng JQ, Yan JX et al. Effect of electro-acupuncture on pain relief and joint function in patients with knee osteoarthritis: A meta-analysis. [Chinese]. Chinese journal of tissue engineering research. 2020; 24(21):3438-3444

22. Christensen BV, Iuhl IU, Vilbek H, Bulow HH, Dreijer NC, Rasmussen HF. Acupuncture treatment of severe knee osteoarthrosis. A long-term study. Acta Anaesthesiologica Scandinavica. 1992; 36(6):519-525

23. Corbett MS, Rice SJ, Madurasinghe V, Slack R, Fayter DA, Harden M et al. Acupuncture and other physical treatments for the relief of pain due to osteoarthritis of the knee: network meta-analysis. Osteoarthritis and Cartilage. 2013; 21(9):1290-1298

24. Dickens W, Lewith GT. A single-blind controlled and randomised clinical trial to evaluate the effect of acupuncture in the treatment of trapezio-metacarpal osteoarthritis. Complementary medical research. 1989; 3(2):5-8

25. Ding Y, Wang Y, Shi X, Luo Y, Gao Y, Pan J. Effect of ultrasound-guided acupotomy vs electro-acupuncture on knee osteoarthritis: a randomized controlled study. Journal of Traditional Chinese Medicine. 2016; 36(4):450-455

26. Dunning J, Butts R, Young I, Mourad F, Galante V, Bliton P et al. Periosteal electrical dry needling as an adjunct to exercise and manual therapy for knee osteoarthritis: A multicenter randomized clinical trial. Clinical Journal of Pain. 2018; 34(12):1149-1158

27. Elbadawy MA. Effectiveness of periosteal stimulation therapy and home exercise program in the rehabilitation of patients with advanced knee osteoarthritis. Clinical Journal of Pain. 2017; 33(3):254-263

28. Endres HG, Victor N, Haake M, Witte S, Streitberger K, Zenz M. Acupuncture for the treatment of chronic knee and back pain. 574. Deutsches arzteblatt ausg a. 2007; 104(3):123-130

29. Ernst E. Acupuncture as a symptomatic treatment of osteoarthritis. A systematic review. Scandinavian Journal of Rheumatology. 1997; 26(6):444-447

30. Farazdaghi M, Kordi Yoosefinejad A, Abdollahian N, Rahimi M, Motealleh A. Dry needling trigger points around knee and hip joints improves function in patients with mild to moderate knee osteoarthritis. Journal of Bodywork and Movement Therapies. 2021; 27:597-604

31. Fargas-Babjak AM, Pomeranz B, Rooney PJ. Acupuncture-like stimulation with Codetron for rehabilitation of patients with chronic pain syndrome and osteoarthritis. Acupuncture and Electro-Therapeutics Research. 1992; 17:95-105

32. Fink M, Karst M, Wippermann B, Gehrke A. Acupuncture as a complementary treatment in physical medicine: its use in osteoarthritis of the hip. Physikalische medizin rehabilitationsmedizin kurortmedizin. 2001; 11(4):123-128

33. Fink MG, de Groot M, Künsebeck HW, Gutenbrunner CH, Lamprecht F, Gehrke A. Standardized acupuncture versus placebo-acupuncture in coxarthrosis patients - lack of efficacy. Forschende Komplementarmedizin. 1996; 3(6):316

34. Fink MG, Kunsebeck H, Wipperman B, Gehrke A. Non-specific effects of traditional Chinese acupuncture in osteoarthritis of the hip. Complementary Therapies in Medicine. 2001; 9(2):82-89

35. Fink MG, Kunsebeck HW, Wippermann B. Effects of acupuncture on pain amd impairement in patients with osteoarthritis of the hip. Zeitschrift für Rheumatologie. 2000; 59(3):191-199

36. Fink MG, Künsebeck HW, Wippermann B. Effect of needle acupuncture on pain perception and functional impairment of patients with coxarthrosis. Zeitschrift für Rheumatologie. 2000; 59(3):191-199

37. Foster NE, Thomas E, Barlas P, Hill JC, Young J, Mason E et al. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. BMJ. 2007; 335(7617):436

38. Foster NE, Vertosick EA, Lewith G, Linde K, MacPherson H, Sherman KJ et al. Identifying patients with chronic pain who respond to acupuncture: results from an individual patient data meta-analysis. Acupuncture in Medicine. 2021; 39(2):83-90

39. Fu MY, Zhang ZL. Knee osteoarthritis treated with acupuncture based on syndrome differentiation: a randomized controlled trial. World Journal of Acupuncture - Moxibustion. 2012; 22(3):11-17

40. Gaw AC, Chang LW, Shaw LC. Efficacy of acupuncture on osteoarthritic pain. A controlled, double-blind study. New England Journal of Medicine. 1975; 293(8):375-378

41. Gollub RL, Kirsch I, Maleki N, Wasan AD, Edwards RR, Tu Y et al. A functional neuroimaging study of expectancy effects on pain response in patients with knee osteoarthritis. Journal of Pain. 2018; 19(5):515-527

42. Gong Z, Liu R, Yu W, Wong TK, Guo Y, Sun Y. Acutherapy for knee osteoarthritis relief in the elderly: A systematic review and meta-analysis. Evidence-Based Complementary & Alternative Medicine: eCAM. 2019; 2019:1868107

43. Green S, Buchbinder R, Hetrick S. Acupuncture for shoulder pain. Cochrane Database of Systematic Reviews 2005, Issue 2. Art. No.: CD005319. DOI: 10.1002/14651858.CD005319.

44. Grotle M. Traditional Chinese acupuncture was not superior to sham acupuncture for knee osteoarthritis but delivering treatment with high expectations of improvement was superior to delivering treatment with neutral expectations. Journal of Physiotherapy. 2011; 57(1):56

45. Haslam R. A comparison of acupuncture with advice and exercises on the symptomatic treatment of osteoarthritis of the hip--a randomised controlled trial. Acupuncture in Medicine. 2001; 19(1):19-26

46. Helianthi DR, Simadibrata C, Srilestari A, Wahyudi ER, Hidayat R. Pain reduction after laser acupuncture treatment in geriatric patients with knee osteoarthritis: A randomized controlled trial. Acta Medica Indonesiana. 2016; 48(2):114-121

47. Hinman RS, McCrory P, Pirotta M, Relf I, Forbes A, Crossley KM et al. Acupuncture for chronic knee pain: a randomized clinical trial. JAMA. 2014; 312(13):1313-1322

48. Hou M, Wang X, Yu J, Fu S, Yang F, Li Z et al. The effect of electroacupuncture on dynamic balance during stair climbing for elderly patients with knee osteoarthritis. Evidence-Based Complementary & Alternative Medicine: eCAM. 2020; 2020:3563584

49. Huang LJ, Luo WJ, Zhang KB, Geng JH. Efficacy observation on cervical spondylosis of vertebral artery type treated with warm needling and rehabilitation physiotherapy therapy. World Journal of Acupuncture - Moxibustion. 2012; 22(2):12-16

50. Itoh K, Hirota S, Katsumi Y, Ochi H, Kitakoji H. A pilot study on using acupuncture and transcutaneous electrical nerve stimulation (TENS) to treat knee osteoarthritis (OA). Chinesische Medizin. 2008; 3:2

51. Itoh K, Hirota S, Katsumi Y, Ochi H, Kitakoji H. Trigger point acupuncture for treatment of knee osteoarthritis--a preliminary RCT for a pragmatic trial. Acupuncture in Medicine. 2008; 26(1):17-26

52. Jia J, Mao GL, Hu SH, Dong XC. Acupuncture combined with function exercise for the elder patients with knee osteoarthritis. Chinese journal of clinical rehabilitation. 2005; 9(10):18-19

53. Jia P, Liu J, Li L, Luo Y, Li Y, Zhao L et al. Acupuncture for knee osteoarthritis with sensitized acupoints: results from a pilot, feasibility randomized controlled trial. Pilot & Feasibility Studies. 2020; 6:144

54. Ju Z, Guo X, Jiang X, Wang X, Liu S, He J et al. Electroacupuncture with different current intensities to treat knee osteoarthritis: a single-blinded controlled study. International Journal of Clinical and Experimental Medicine. 2015; 8(10):18981-18989

55. Jubb RW, Tukmachi ES, Jones PW, Dempsey E, Waterhouse L, Brailsford S. A blinded randomised trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupuncture in Medicine. 2008; 26(2):69-78

56. Jun S, Lee JH, Gong HM, Choi SH, Bo MH, Kang MS et al. Efficacy and safety of miniscalpel acupuncture on knee osteoarthritis: - a randomized controlled pilot trial. Journal of Pharmacopuncture. 2018; 21(3):151-158

57. Karner M, Brazkiewicz F, Remppis A, Fischer J, Gerlach O, Stremmel W et al. Objectifying specific and nonspecific effects of acupuncture: a double-blinded

randomised trial in osteoarthritis of the knee. Evidence-Based Complementary & Alternative Medicine: eCAM. 2013; 2013:427265

58. Kim EJ, Jang MK, Yoon EH, Jung CY, Nam DW, Lee SD et al. Efficacy of pharmacopuncture using root bark of Ulmus davidiana Planch in patients with knee osteoarthritis: a double-blind randomized controlled trial. Journal of Acupuncture & Meridian Studies. 2010; 3(1):16-23

59. Kim EJ, Lim CY, Lee EY, Lee SD, Kim KS. Comparing the effects of individualized, standard, sham and no acupuncture in the treatment of knee osteoarthritis: a multicenter randomized controlled trial. Trials [Electronic Resource]. 2013; 14:129

60. Kim SY, Lee H, Chae Y, Park HJ, Lee H. A systematic review of cost-effectiveness analyses alongside randomised controlled trials of acupuncture. Acupuncture in Medicine. 2012; 30(4):273-285

61. Kong J, Wang Z, Leiser J, Minicucci D, Edwards R, Kirsch I et al. Enhancing treatment of osteoarthritis knee pain by boosting expectancy: A functional neuroimaging study. NeuroImage Clinical. 2018; 18:325-334

62. Kwon YB, Kim JH, Yoon JH, Lee JD, Han HJ, Mar WC et al. The analgesic efficacy of bee venom acupuncture for knee osteoarthritis: a comparative study with needle acupuncture. American Journal of Chinese Medicine. 2001; 29(2):187-199

63. Lam WC, Au KY, Qin Z, Wu FM, Chong CO, Jiang F et al. Superficial needling acupuncture vs sham acupuncture for knee osteoarthritis: A randomized controlled trial. American Journal of Medicine. 2021; 134(10):1286-1294 e1282

64. Lansdown H, Howard K, Brealey S, MacPherson H. Acupuncture for pain and osteoarthritis of the knee: a pilot study for an open parallel-arm randomised controlled trial. BMC Musculoskeletal Disorders. 2009; 10:130

65. Latimer NR, Bhanu AC, Whitehurst DG. Inconsistencies in NICE guidance for acupuncture: reanalysis and discussion. Acupuncture in Medicine. 2012; 30(3):182-186

66. Lee MS, Shin BC, Choi SM, Kim JY. Randomized clinical trials of constitutional acupuncture: a systematic review. Evidence-Based Complementary & Alternative Medicine: eCAM. 2009; 6 (Suppl 1):59-64

67. Lev-Ari S, Miller E, Maimon Y, Rosenblatt Y, Mendler A, Hasner A et al. Delayed effect of acupuncture treatment in OA of the knee: A blinded, randomized, controlled trial. Evidence-Based Complementary and Alternative Medicine. 2011; 2011:792975

68. Li LW, Harris RE, Tsodikov A, Struble L, Murphy SL. Self-acupressure for older adults with symptomatic knee osteoarthritis: A randomized controlled trial. Arthritis Care and Research. 2018; 70(2):221-229

69. Li S, Xie P, Liang Z, Huang W, Huang Z, Ou J et al. Efficacy comparison of five different acupuncture methods on pain, stiffness, and function in osteoarthritis of the knee: A network meta-analysis. Evidence-Based Complementary & Alternative Medicine: eCAM. 2018; 2018:1638904

70. Li Z, Li DP. Electro-acupuncture for treating 38 cases with knee osteoarthritis. Henan traditional chinese medicine [henan zhong yi]. 2015; 35(2):300-302

71. Liang YF, Li WX, Ma Y, Zhang BY, Huang YR. Comparison of therapeutic effect of soft-tissue relaxing needling and electroacupuncture for knee osteoarthritis. Zhen CI yan jiu = acupuncture research. 2019; 44(6):439-442

72. Lim JA, Lee JD, Lee SK, Lee SY, Moon HC, Choi SM. Pilot study of acupuncture treatment for the osteoarthritis of the knee joint on the ebm(evidence basement medicine). The journal of korean acupuncture & moxibustion society= taehan chimgu hakhoe chi. 2006; 23(1)

73. Lin LL, Li YT, Tu JF, Yang JW, Sun N, Zhang S et al. Effectiveness and feasibility of acupuncture for knee osteoarthritis: a pilot randomized controlled trial. Clinical Rehabilitation. 2018; 32(12):1666-1675

74. Lin LL, Tu JF, Wang LQ, Yang JW, Shi GX, Li JL et al. Acupuncture of different treatment frequencies in knee osteoarthritis: a pilot randomised controlled trial. Pain. 2020; 161(11):2532-2538

75. Lin X, Huang K, Zhu G, Huang Z, Qin A, Fan S. The effects of acupuncture on chronic knee pain due to osteoarthritis: A meta-analysis. Journal of Bone & Joint Surgery - American Volume. 2016; 98(18):1578-1585

76. Linde K, Witt CM, Streng A, Weidenhammer W, Wagenpfeil S, Brinkhaus B et al. The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. Pain. 2007; 128(3):264-271

77. 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

78. Lu TW, Wei IP, Liu YH, Hsu WC, Wang TM, Chang CF et al. Immediate effects of acupuncture on gait patterns in patients with knee osteoarthritis. Chinese Medical Journal. 2010; 123(2):165-172

79. 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

80. Lundeberg T, Eriksson SV, Lundeberg S, Thomas M. Effect of acupuncture and naloxone in patients with osteoarthritis pain. A sham acupuncture controlled study. Pain clinic. 1991; 4(3):155-161

81. Luo X, Hou XS, Tian ZY, Meng X, Li SM, Bai P. Randomized controlled clinical trial of acupuncture treatment for knee osteoarthritis in the early stage. Zhen CI yan jiu = acupuncture research. 2019; 44(3):211-215

82. Lv ZT, Shen LL, Zhu B, Zhang ZQ, Ma CY, Huang GF et al. Effects of intensity of electroacupuncture on chronic pain in patients with knee osteoarthritis: a randomized controlled trial. Arthritis Research & Therapy. 2019; 21(1):120

83. Lv ZT, Shen LL, Zhu B, Zhang ZQ, Ma CY, Huang GF et al. Effects of intensity of electroacupuncture on chronic pain in patients with knee osteoarthritis: a randomized controlled trial. Arthritis Research & Therapy. 2019; 21(1):120

84. Maa SH, Sun MF, Wu CC. The effectiveness of acupuncture on pain and mobility in patients with osteoarthritis of the knee: a pilot study. Journal of Nursing Research. 2008; 16(2):140-148

85. MacPherson H, Vickers A, Bland M, Torgerson D, Corbett M, Spackman E et al. Acupuncture for chronic pain and depression in primary care: a programme of research. Acupuncture for chronic pain and depression in primary care: a programme of research. Programme Grants for Applied Research. Southampton (UK). 2017.

86. Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S et al. Acupuncture for peripheral joint osteoarthritis. Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD001977. DOI: 10.1002/14651858.CD001977.pub2.

87. Manheimer E, Cheng K, Wieland L, Shen X, Lao L, Guo M et al. Acupuncture for hip osteoarthritis. Cochrane Database of Systematic Reviews 2018, Issue 5. Art. No.: CD013010. DOI: 10.1002/14651858.CD013010.

88. Manheimer E, Lim B, Lao L, Berman B. Acupuncture for knee osteoarthritis--a randomised trial using a novel sham. Acupuncture in Medicine. 2006; 24 (Suppl):S7-14

89. Manheimer E, Linde K, Lao L, Bouter LM, Berman BM. Meta-analysis: acupuncture for osteoarthritis of the knee. Annals of Internal Medicine. 2007; 146(12):868-877

90. Manyanga T, Froese M, Zarychanski R, Abou-Setta A, Friesen C, Tennenhouse M et al. Pain management with acupuncture in osteoarthritis: a systematic review and meta-analysis. BMC Complementary and Alternative Medicine. 2014; 14:312

91. Martins F, Kaster T, Schutzler L, Witt CM. Factors influencing further acupuncture usage and a more positive outcome in patients with osteoarthritis of the knee and the hip: a 3-year follow-up of a randomized pragmatic trial. Clinical Journal of Pain. 2014; 30(11):953-959

92. Mavrommatis CI, Argyra E, Vadalouka A, Vasilakos DG. Acupuncture as an adjunctive therapy to pharmacological treatment in patients with chronic pain due to osteoarthritis of the knee: a 3-armed, randomized, placebo-controlled trial. Pain. 2012; 153(8):1720-1726

93. McIndoe AK, Young K, Bone ME. A comparison of acupuncture with intra-articular steroid injection as analgesia for osteoarthritis of the hip. Acupuncture in Medicine. 1995; 13(2):67-70

94. Meng CR, Fan L, Fu WB, Li Y. Clinical research on abdominal acupuncture plus conventional acupuncture for knee osteoarthritis. Journal of Traditional Chinese Medicine. 2009; 29(4):249-252

95. Miller E, Maimon Y, Rosenblatt Y, Mendler A, Hasner A, Barad A et al. Delayed effect of acupuncture treatment in OA of the knee: A blinded, randomized, controlled trial. Evidence-Based Complementary and Alternative Medicine. 2011; 2011:792975

96. Min MH, Choi YG, Kim YJ, Park HJ, Lee SH, Joo HN. The Effect of Sa-am acupuncture on Knee Osteoarthritis. The korean journal of meridian & acupoint = kyunglak kyunghyul hakhoe chi. 2009; 26(4):55-66

97. Molsberger A, wG B. Acupuncture treatment in the relief of gonarthrosis pain - a controlled clinical trial. The third world conference on acupuncture (kyoto) I P-62 (unpublished). 1993;

98. 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

99. Ng MM, Leung MC, Poon DM. The effects of electro-acupuncture and transcutaneous electrical nerve stimulation on patients with painful osteoarthritic knees: a randomized controlled trial with follow-up evaluation. Journal of Alternative and Complementary Medicine. 2003; 9(5):641-649

100. Nie B, Yan YP, Zhang XY, Yang D, Wei ZL. Clinical research of pain symptom and efficacy on knee osteoarthritis treated by multi-needling in muscle meridian focus. Journal of clinical acupuncture and moxibustion [zhen jiu lin chuang za zhi]. 2015; 31(2):43-45

101. Organisation for Economic Co-operation and Development (OECD). Purchasing power parities (PPP). 2012. Available from: http://www.oecd.org/std/ppp Last accessed: 10/02/2022.

102. Penagos-Martinez AA, Patino-Palma BE, Rodriguez-Puerto OL. Acupuncture and therapeutic exercise in women over 50 with osteoarthritis of the knee. Randomized clinical trial. Revista Internacional de Acupuntura. 2021; 15(1):15-23

103. Petrou P, Winkler V, Genti G, Balint G. Double blind trial to evaluate the effect of acupuncture treatment on knee osteoarthrosis. Scandinavian journal of acupuncture and electrotherapy. 1988; 3(3):112-115

104. Plaster R, Vieira WB, Alencar FA, Nakano EY, Liebano RE. Immediate effects of electroacupuncture and manual acupuncture on pain, mobility and muscle strength in patients with knee osteoarthritis: a randomised controlled trial. Acupuncture in Medicine. 2014; 32(3):236-241

105. Rahou-El-Bachiri Y, Navarro-Santana MJ, Gomez-Chiguano GF, Cleland JA, Lopez-de-Uralde-Villanueva I, Fernandez-de-Las-Penas C et al. Effects of trigger point dry needling for the management of knee pain syndromes: A systematic review and meta-analysis. Journal of Clinical Medicine. 2020; 9(7):29

106. Rauschkolb C, Lange B, Kuperwasser B, Kelly K, Okamoto A, Van Hove I. Tapentadol extended release for the relief of chronic osteoarthritis knee pain: results from the EuroQol-5 dimension (EQ-5D) and Western Ontario and McMaster Universities osteoarthritis index (WOMAC) questionnaires. Osteoarthritis and Cartilage. 2009; 17:S179

107. Reinhold T, Witt CM, Jena S, Brinkhaus B, Willich SN. Quality of life and cost-effectiveness of acupuncture treatment in patients with osteoarthritis pain. European Journal of Health Economics. 2008; 9:209-219

108. Rodriguez-Merchan EC. Conservative treatment of acute knee osteoarthritis: A review of the Cochrane Library. Journal of Acute Disease. 2016; 5(3):190-193

109. Saleki M, Ahadi T, Razi M, Raeisi GR, Forough B, Ali MK. Comparison of the effects of acupuncture and isometric exercises on symptom of knee osteoarthritis. International Journal of Preventive Medicine. 2013; 4(Suppl 1):S73-77

110. Sanchez-Romero EA, Pecos-Martin D, Calvo-Lobo C, Ochoa-Saez V, Burgos-Caballero V, Fernandez-Carnero J. Effects of dry needling in an exercise program for older adults with knee osteoarthritis: A pilot clinical trial. Medicine. 2018; 97(26):e11255

111. Sanchez Romero EA, Fernandez-Carnero J, Calvo-Lobo C, Ochoa Saez V, Burgos Caballero V, Pecos-Martin D. Is a combination of exercise and dry needling effective for knee OA? Pain Medicine. 2020; 21(2):349-363

112. Sangdee C, Teekachunhatean S, Sananpanich K, Sugandhavesa N, Chiewchantanakit S, Pojchamarnwiputh S et al. Electroacupuncture versus diclofenac in symptomatic treatment of osteoarthritis of the knee: a randomized controlled trial. BMC Complementary and Alternative Medicine. 2002; 2:3

113. Scharf HP, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. Annals of Internal Medicine. 2006; 145(1):12-20

114. Selfe TK, Taylor AG. Acupuncture and osteoarthritis of the knee: a review of randomized, controlled trials. Family and Community Health. 2008; 31(3):247-254

115. Shafshak TS. Electroacupuncture and exercise in body weight reduction and their application in rehabilitating patients with knee osteoarthritis. American Journal of Chinese Medicine. 1995; 23(1):15-25

116. Shen X, Zhao L, Ding G, Tan M, Gao J, Wang L et al. Effect of combined laser acupuncture on knee osteoarthritis: a pilot study. Lasers in Medical Science. 2009; 24(2):129-136

117. Sheng YX, Jia Q, Shu GJ, Wang GX, Cui JF, Zhao N et al. Clinical study on treatment of knee osteoarthritis by acupuncture combined with ultrasonic. China journal of chinese medicine [zhong yi xue bao]. 2015; 30(8):1224-1226

118. Shim JW, Jung JY, Kim SS. Effects of electroacupuncture for knee osteoarthritis: A systematic review and meta-analysis. Evidence-Based Complementary & Alternative Medicine: eCAM. 2016; 2016:3485875

119. Singh BB, Berman BM, Hadhazy V, Bareta J, Lao L, Zarow FM et al. Clinical decisions in the use of acupuncture as an adjunctive therapy for osteoarthritis of the knee. Alternative Therapies in Health and Medicine. 2001; 7(4):58-65

120. Song W-G, Wu T, Liu M. Clinical analysis of knee osteoarthritis treated mainly by acupuncture. Journal of Acupuncture and Tuina Science. 2004; 2(3):26

121. Soni A, Joshi A, Mudge N, Wyatt M, Williamson L. Supervised exercise plus acupuncture for moderate to severe knee osteoarthritis: a small randomised controlled trial. Acupuncture in Medicine. 2012; 30(3):176-181

122. 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

123. Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Marcus DM et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. Arthritis Care and Research. 2010; 62(9):1229-1236

124. Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Marcus DM et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. Arthritis Care and Research. 2010; 62(9):1229-1236

125. Suen LK, Yeh CH, Yeung SK. Erratum to: Using auriculotherapy for osteoarthritic knee among elders: a double-blinded randomised feasibility study. BMC Complementary and Alternative Medicine. 2016; 16(1):391

126. Taechaarpornkul W, Suvapan D, Theppanom C, Chanthipwaree C, Chirawatkul A. Comparison of the effectiveness of six and two acupuncture point regimens in osteoarthritis of the knee: a randomised trial. Acupuncture in Medicine. 2009; 27(1):3-8

127. Takeda W, Wessel J. Acupuncture for the treatment of pain of osteoarthritic knees. Arthritis Care and Research. 1994; 7(3):118-122

128. Tang L, Jia P, Zhao L, Kang D, Luo Y, Liu J et al. Acupuncture treatment for knee osteoarthritis with sensitive points: protocol for a multicentre randomised controlled trial. BMJ Open. 2018; 8(10):e023838

129. Thomas M, Eriksson SV, Lundeberg T. A comparative study of diazepam and acupuncture in patients with osteoarthritis pain: a placebo controlled study. American Journal of Chinese Medicine. 1991; 19(2):95-100

130. Tillu A, Tillu S, Vowler S. Effect of acupuncture on knee function in advanced osteoarthritis of the knee: a prospective, non-randomised controlled study. Acupuncture in Medicine. 2002; 20(1):19-21

131. Trinh KV. Blinding in acupuncture research: A systematic review of randomized controlled trials for pain using a sham acupuncture control. Clinical Acupuncture and Oriental Medicine. 2003; 4(2-3):71-77

132. Tu JF, Wang LQ, Shi GX, Yang JW, Li JL, Li YT et al. Effect of acupuncture on knee injury and osteoarthritis outcome score in patients with knee osteoarthritis. Zhongguo zhen jiu [Chinese acupuncture & moxibustion]. 2021; 41(1):27-30

133. Tu JF, Yang JW, Lin LL, Wang TQ, Du YZ, Liu ZS et al. Correction to: Efficacy of electro-acupuncture and manual acupuncture versus sham acupuncture for knee osteoarthritis: study protocol for a randomised controlled trial. Trials [Electronic Resource]. 2019; 20(1):204

134. Tu JF, Yang JW, Lin LL, Wang TQ, Du YZ, Liu ZS et al. Efficacy of electro-acupuncture and manual acupuncture versus sham acupuncture for knee osteoarthritis: study protocol for a randomised controlled trial. Trials [Electronic Resource]. 2019; 20(1):79

135. Tu JF, Yang JW, Shi GX, Yu ZS, Li JL, Lin LL et al. Efficacy of intensive acupuncture versus sham acupuncture in knee osteoarthritis: A randomized controlled trial. Arthritis & Rheumatology. 2021; 73(3):448-458

136. Tukmachi E, Jubb R, Dempsey E, Jones P. The effect of acupuncture on the symptoms of knee osteoarthritis--an open randomised controlled study. Acupuncture in Medicine. 2004; 22(1):14-22

137. Ughreja RA, Prem V. Effectiveness of dry needling techniques in patients with knee osteoarthritis: A systematic review and meta-analysis. Journal of Bodywork and Movement Therapies. 2021; 27:328-338

138. Vas J, Mendez C, Perea-Milla E. Acupuncture versus Streitberger needle in knee osteoarthritis -- an RCT. Acupuncture medicine. 2006; 24(Suppl):S15-S24

139. Vas J, Mendez C, Perea-Milla E, Vega E, Panadero MD, Leon JM et al. Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: randomised controlled trial. BMJ. 2004; 329(7476):1216

140. Vickers AJ, Cronin AM, Maschino AC, Lewith G, MacPherson H, Foster NE et al. Acupuncture for chronic pain: individual patient data meta-analysis. Archives of Internal Medicine. 2012; 172(19):1444-1453

141. Vickers AJ, Vertosick EA, Lewith G, MacPherson H, Foster NE, Sherman KJ et al. Acupuncture for chronic pain: Update of an individual patient data meta-analysis. Journal of Pain. 2018; 19(5):455-474

142. Wang GX, Deng KF, Liao ZL, Zhu Y, Chen RL. Effect of hysteretic acupuncture on joint function and related cytokines in articular fluid in patients with knee osteoarthritis. Zhen CI yan jiu = acupuncture research. 2020; 45(7):564-568

143. Wang Q, Lv H, Sun ZT, Tu JF, Feng YW, Wang TQ et al. Effect of electroacupuncture versus sham electroacupuncture in patients with knee osteoarthritis: A pilot randomized controlled trial. Evidence-Based Complementary & Alternative Medicine: eCAM. 2020; 2020:1686952

144. Wang QL, Lu TX, Cai GH, Xu RX, Zhu LJ, Lin ZG et al. Therapeutic effect observation on thumb-tack needling for knee osteoarthritis in elderly patients: A radom control trial. World Journal of Acupuncture - Moxibustion. 2020; 30(4):268-271

145. Wang TQ, Li LR, Tan CX, Yang JW, Shi GX, Wang LQ et al. Effect of electroacupuncture on gut microbiota in participants with knee osteoarthritis. Frontiers in Cellular and Infection Microbiology. 2021; 11 597431

146. Wang TQ, Li YT, Wang LQ, Shi GX, Tu JF, Yang JW et al. Electroacupuncture versus manual acupuncture for knee osteoarthritis: a randomized controlled pilot trial. Acupuncture in Medicine. 2020; 38(5):291-300

147. Wang TQ, Lu WT, Yang J. Electroacupuncture for knee osteoarthritis based on different meridian syndrome: A randomized controlled pilot trial. Acupuncture and Electro-Therapeutics Research. 2021; 46(3):247-258

148. Weiner DK, Moore CG, Morone NE, Lee ES, Kent Kwoh C. Efficacy of periosteal stimulation for chronic pain associated with advanced knee osteoarthritis: a randomized, controlled clinical trial. Clinical Therapeutics. 2013; 35(11):1703-1720.e1705

149. Weiner DK, Rudy TE, Morone N, Glick R, Kwoh CK. Efficacy of periosteal stimulation therapy for the treatment of osteoarthritis-associated chronic knee pain: an initial controlled clinical trial. Journal of the American Geriatrics Society. 2007; 55(10):1541-1547

150. White A, Foster NE, Cummings M, Barlas P. Acupuncture treatment for chronic knee pain: a systematic review. Rheumatology. 2007; 46(3):384-390

151. White P, Bishop FL, Prescott P, Scott C, Little P, Lewith G. Practice, practitioner, or placebo? A multifactorial, mixed-methods randomized controlled trial of acupuncture. Pain. 2012; 153(2):455-462

152. White P, Prescott P, Lewith G. Does needling sensation (de qi) affect treatment outcome in pain? Analysis of data from a larger single-blind, randomised controlled trial. Acupuncture in Medicine. 2010; 28(3):120-125

153. Whitehurst DG, Bryan S, Hay EM, Thomas E, Young J, Foster NE. Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee. Physical Therapy. 2011; 91(5):630-641

154. 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

155. Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S et al. Acupuncture in patients with osteoarthritis of the knee: a randomised trial. Lancet. 2005; 366(9480):136-143

156. Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN. Acupuncture in patients with osteoarthritis of the knee or hip: a randomized, controlled trial with an additional nonrandomized arm. Arthritis and Rheumatism. 2006; 54(11):3485-3493

157. Witt CM, Vertosick EA, Foster NE, Lewith G, Linde K, MacPherson H et al. The effect of patient characteristics on acupuncture treatment outcomes: An individual patient data meta-analysis of 20,827 chronic pain patients in randomized controlled trials. Clinical Journal of Pain. 2019; 35(5):428-434

158. Wu ZH, Bao F. Observation on therapeutic effects of electroacupuncture for the treatment of knee osteoarthritis. Zhongguo gu shang [China journal of orthopaedics and traumatology]. 2008; 21(3):170-172

159. Xi MJ, Zhao L, Shen XY, Ding GH, Gao JP, Wang LZ et al. Efficacy and specificity of laser acupuncture on the treatment of knee osteoarthritis. Journal of clinical rehabilitative tissue engineering research. 2008; 12(26):5075-5078

160. Xu H, Wu HG. Clinical observation of electroacupuncture combined with low-dose diclofenac in treating osteoarthritis of the knee. Zhong xi yi jie he xue bao [Journal of Chinese integrative medicine]. 2007; 5(4):457-459

161. Yurtkuran M, Alp A, Konur S, Ozcakir S, Bingol U. Laser acupuncture in knee osteoarthritis: a double-blind, randomized controlled study. Photomedicine and Laser Surgery. 2007; 25(1):14-20

162. Zhang L, Yuan H, Zhang L, Li J, Li H. Effect of acupuncture therapies combined with usual medical care on knee osteoarthritis. Journal of Traditional Chinese Medicine. 2019; 39(1):103-110

163. Zhang Q, Fang J, Chen L, Wu J, Ni J, Liu F et al. Different kinds of acupuncture treatments for knee osteoarthritis: a multicentre, randomized controlled trial. Trials [Electronic Resource]. 2020; 21(1):264

164. Zhang Q, Yue J, Golianu B, Sun Z, Lu Y. Updated systematic review and meta-analysis of acupuncture for chronic knee pain. Acupuncture in Medicine. 2017; 35(6):392-403

165. Zhang Y, Bao F, Wang Y, Wu Z. Influence of acupuncture in treatment of knee osteoarthritis and cartilage repairing. American Journal Of Translational Research. 2016; 8(9):3995-4002

166. Zhen X, Che T. Clinical observations on electroacupuncture treatment of hip osteoarthritis. Shanghai journal of acupuncture and moxibustion. 2004; 23(8):16

167. Zhou SF, Xue CC. Acupuncture as an adjunct to exercise-based physiotherapy does not improve the pain of knee osteoarthritis. Australian journal of acupuncture and chinese medicine. 2008; 3(1):53-55

Appendices

Appendix A – Review protocols

Review protocol for the clinical and cost-effectiveness of acupuncture in the management of osteoarthritis

ID	Field	Content
0.	PROSPERO registration number	CRD42020217866
1.	Review title	What is the clinical and cost-effectiveness of acupuncture for the management of osteoarthritis?
2.	Review question	3.4 What is the clinical and cost-effectiveness of acupuncture for the management of osteoarthritis?
3.	Objective	To assess the clinical and cost-effectiveness of acupuncture (including conventional acupuncture, dry needling, electroacupuncture) 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
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, diseases of childhood that may predispose to osteoarthritis, medical conditions presenting with joint inflammation and malignancy). • Studies in people with meniscal injury without osteoarthritis • Studies with an unclear population (e,g, type of arthritis, proportion of participants with osteoarthritis) • Spinal osteoarthritis

7.	Intervention/Exposure/Test	• Acupuncture/dry needling • Electroacupuncture

8.	Comparator/Reference standard/Confounding factors	• Compared to each other
		• Sham acupuncture
		• No intervention (including either):
		  ○ Acupuncture versus no treatment*
		  ○ Acupuncture 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.*

9.	Types of study to be included	• Systematic reviews of RCTs • Parallel RCTs • Crossover RCTs will be considered if insufficient evidence is available from parallel RCTs* Non-randomised studies will be excluded. *Insufficient evidence defined as evidence that is insufficient to inform recommendations (either quality or quantity).
10.	Other exclusion criteria	• Non-English language studies • Non-randomised/observational studies

		• 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
13.	Secondary outcomes (important outcomes)	• Psychological distress [validated patient-reported outcomes, continuous data prioritised] • Osteoarthritis flare-ups [validated patient-reported outcomes, continuous data prioritised] • Serious adverse events [dichotomous data]
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: • 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) • Site of osteoarthritis • Acupuncture/dry needling

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.	Stage of review at time of this submission	Review stage	Started	Completed
		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] Emma Cowles [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	Acupuncture; Adults; Electroacupuncture; Intervention; Laser acupuncture; Non-Pharmacological; Osteoarthritis

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 16: 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).[98] **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 acupuncture 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.[98]

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 17: 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 18: 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.

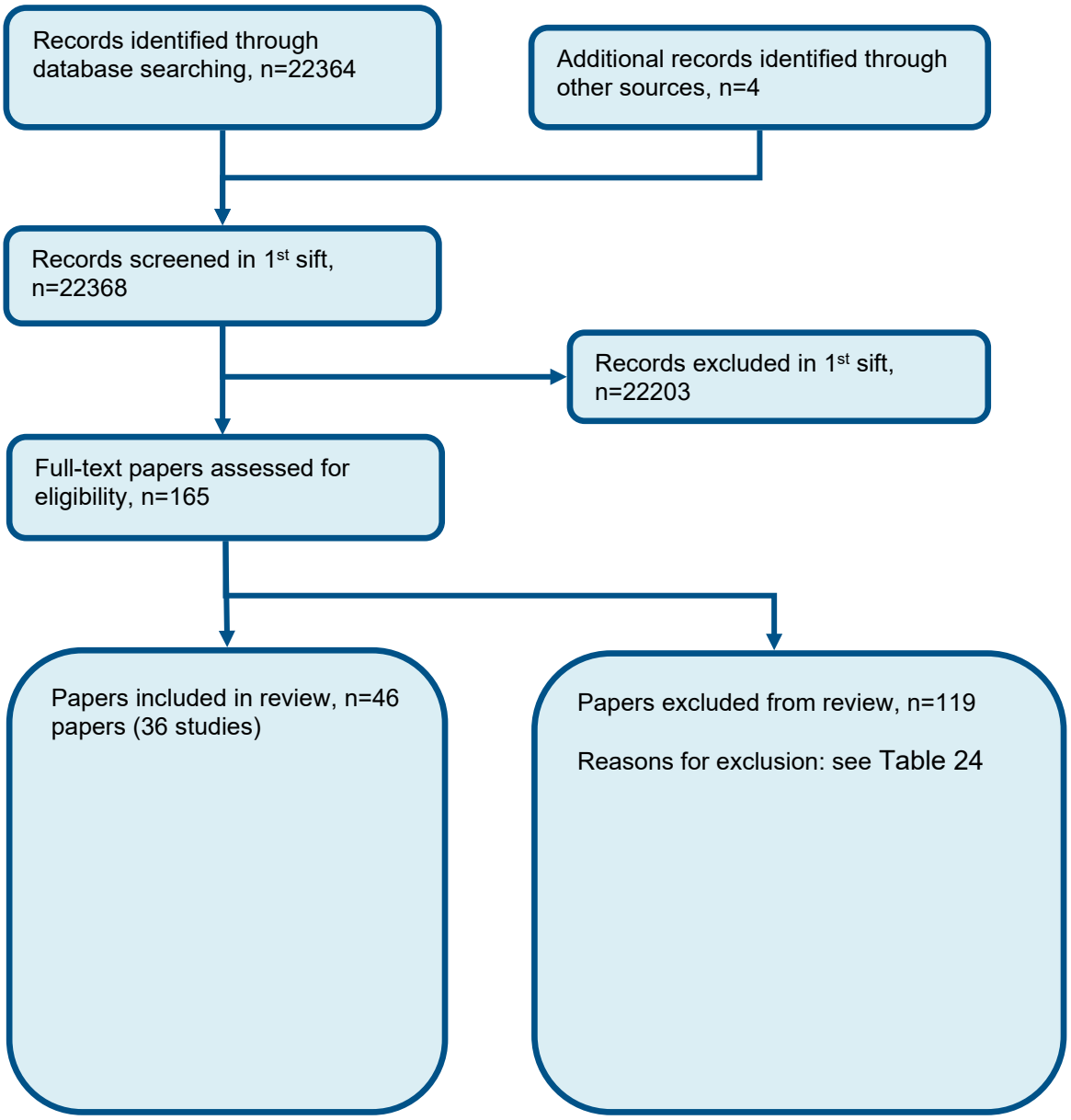
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	quality adjusted 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 acupuncture for the management of osteoarthritis



Appendix D – Effectiveness evidence

Study	Berman 1999[10]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=73)
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: Diagnosis of osteoarthritis of the knee (American College of Rheumatology criteria applied)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Diagnosis of osteoarthritis of the knee (American College of Rheumatology criteria applied) of at least 6 months duration; at least moderate pain in the knee for most days in the last month; aged 50 years or above; taking analgesic or anti-inflammatory agents for control of pain for at least 1 month; documented radiographic changes of osteoarthritis (Kellgren-Lawrence grade of 2 or more); signed informed consent.
Exclusion criteria	Intra-articular corticosteroid injection into the knee(s) within 4 weeks immediately preceding entry into the study; severe chronic or uncontrolled concomitant illness (e.g. coronary artery disease); history or clinical indications of bleeding diathesis, including current use of anticoagulants.
Recruitment/selection of patients	People were recruited from the Faculty Practice of the Division of Rheumatology at the University of Maryland, and through public service advertisements in radio and print media in the greater Baltimore area
Age, gender and ethnicity	Age - Mean (SD): 65.6 (8.6). Gender (M:F): 29:44. Ethnicity: White = 83%, non-white = 17%
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grade of 2 or more Duration of symptoms (mean [SD]): 7.2 (6.2) years
Indirectness of population	No indirectness

Interventions	(n=37) Intervention 1: Electroacupuncture. Acupuncture biweekly for 8 weeks with electrical stimulation. Selection of acupuncture points was based on the TCM theory for treating Bi syndrome, which uses local and distal points on channels that traverse the area of pain. The following local acupuncture points were used: Yanglinquan (GB 34), Yinlinquan (Sp 9), Zusanli (St 36), Dubi (St 35) and the extra point Xiyan. The distal points used were Kunlun (UB 60), Xuanzhong (GB 39), Sanyinjiao (Sp 6) and Taixi (Kid 3). The skin was sterilized with alcohol and acupuncture needles (1 inch, 34 gauge, 0.22mm diameter needle) were inserted to standard depths (0.4-0.6 inches). The De Qi sensation was verified by the patient. Two electrodes were attached to the needles at local point Dubi and Xiyan. Electrical stimulation with 2.5-4Hz, square pulses of 1.0ms duration was used for 20 minutes. Duration 8 weeks with an additional 6 weeks of follow up. Concurrent medication/care: People were asked to remain on their baseline analgesc/anti-inflammatory regimens as well and not to begin any new physiotherapy or exercise programmes. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=36) Intervention 2: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. Conventional therapy. Duration 12 weeks. Concurrent medication/care: People were asked to remain on their baseline analgesc/anti-inflammatory regimens as well and not to begin any new physiotherapy or exercise programmes. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (This work was supported by the Maurice Laing Foundation and National Institutes of Health-National Center of Complementary and Alternative medicine and the National Institutes of Arthritis/Musculoskeletal/Skin Diseases (Grant no. 1 R21-RR09327-01).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain (longitudinal linear regression analysis) at 12 weeks; Group 1: mean 5.56 (SD 3.44); n=37, Group 2: mean 9.51 (SD 3.01); n=36; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline electroacupuncture: 9.58 (3.26). Baseline no treatment: 9.78 (2.83).
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 sex and race. Baseline values of outcomes and age were similar.; Group 1 Number missing: 8, Reason: 8 people dropped out of treatment; Group 2 Number missing: 7, Reason: 7 people dropped out of treatment

Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: WOMAC disability (longitudinal linear regression analysis) at 12 weeks; Group 1: mean 23.17 (SD 13.92); n=37, Group 2: mean 36.78 (SD 10.71); n=36; WOMAC disability 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 34.56 (12.20). Baseline no treatment: 36.19 (9.22). 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 sex and race. Baseline values of outcomes and age were similar.; Group 1 Number missing: 8, Reason: 8 people dropped out of treatment; Group 2 Number missing: 7, Reason: 7 people dropped out of treatment Protocol outcome 3: Serious adverse events at ≤ 3 months - Actual outcome: Side effects at 12 weeks; Group 1: 0/29, Group 2: 0/29 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 sex and race. Baseline values of outcomes and age were similar.; Group 1 Number missing: 8, Reason: 8 people dropped out of treatment; Group 2 Number missing: 7, Reason: 7 people dropped out of treatment	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Berman 2004[9] (Manheimer 2006[88])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=570)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 26 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: A diagnosis of osteoarthritis of the knee with radiographic evidence of at least 1 osteophyte at the tibiofemoral joint
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 50 years or older, a diagnosis of osteoarthritis of the knee, radiographic evidence of at least 1 osteophyte at the tibiofemoral joint (Kellgren Lawrence at least grade 2), moderate or greater clinically significant knee pain on most days during the past month, and willingness to be randomly assigned
Exclusion criteria	The presence of serious medical conditions that precluded participation in study; bleeding disorders that might contraindicate acupuncture; intra-articular corticosteroid or hyaluronate injections (as well as any knee surgeries or concomitant use of topical capsaicin cream) during the past 6 months; previous experience with acupuncture; or any planned events (including total knee replacement) that would interfere with participation in the study during the following 26 weeks
Recruitment/selection of patients	People were recruited primarily through print and radio advertisements from the areas of 3 sites in Baltimore and Towson, Maryland, and New York City, New York
Age, gender and ethnicity	Age - Mean (SD): 65.5 (8.6). Gender (M:F): 205:365. Ethnicity: White = 394, African American = 163, Other = 13
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence at least grade 2 Duration of symptoms: <5->10 years, median <5 years
Indirectness of population	No indirectness
Interventions	(n=190) Intervention 1: Electroacupuncture. True acupuncture. 26 weeks of gradually tapering treatment according to the following schedule; 6 weeks of 2 treatments per week followed by 2 weeks of 1 treatment per week, 4 treatments of 1 treatment every other week, and 12 weeks of 1 treatment per month. The acupuncture point selections

	was based on Traditional Chinese Medicine meridian theory to treat knee joint pain, known as "Bi" syndrome. These points consisted of 5 local points (Yanglinquan [gallbladder meridian point 34], Yinlinquan [spleen meridian point 9], Zhusanli [stomach meridian point 36], Dubi [stomach meridian point 35], and extra point Xiyan) and 4 distal points (Kunlun [urinary-bladder, meridian point 60], Xuanzhong [gall bladder meridian point 39], Sanyinjiao [spleen meridian point 6[, and Taixi [kidney meridian point 3]) on meridians that traverse the area of pain. The same points were treated for each affected leg. If both knees were affected, 9 needles were inserted in each leg. The acupuncturists inserted 1.5-inch (for local points) and 1-inch (for distal points) 32-guage (0.25mm diameter) acupuncture needles to a conventional depth of approximately 0.3 to 1.0 inch, depending on point location. All participants in the treatment group achieved the "De-Qi" sensation, a local sensation of heaviness, numbness, soreness or paresthesia that accompanies the insertion and manipulation of needles during acupuncture at these 9 months. Acupuncturists applied electrical stimulation at knee points Xiyan, at low frequency (8Hz) and square biphasic pulses (0.5ms pulse width) for 20 minutes. To be similar to the control group, they tapped 2 guiding tubes at 2 sham points in the abdominal area, approximately 3cm lateral to and slightly above the umbilicus bilaterally, and immediately affixed a pair of needles to the surface of the same points, without needle insertion, with adhesive tape.. Duration 26 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=191) Intervention 2: Sham acupuncture. Sham acupuncture, modifying the combined insertion and noninsertion procedure from a previous method. Acupuncturists inserted 2 needles into sham points in the abdominal area, approximately 3cm lateral to and slightly above the umbilicus bilaterally, and then immediately applied 2 pieces of adhesive tape next to the needles. In addition, they tapped a mock plastic needle guiding tube on the surface of each of the 9 true points in the leg to produce some discernible sensation and then immediately applied a needle with a piece of adhesive tape to the dermal surface, without needle insertion, of each point for a total of 20 minutes. The sham procedure was given on the same schedule as the experimental group and used the same active needle placements, except actual insertion did not take place. Although electrical stimulation did not occur, a mock transelectrical stimulation unit was attached to the sham needles at the knee.. Duration 26 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture

	(n=189) Intervention 3: Other. Education control - 6 two hour group sessions and educational materials mailed out to participants. Duration 26 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: This group was not included in the final analysis as they did not fulfill the inclusion criteria in the protocol
Funding	Academic or government funding (Grant support by the National Center for Complementary and Alternative Medicine (National Institutes of Health Cooperative Agreement U01 AT-00171), with advice and encouragement by the National Institute of Arthritis and musculoskeletal and Skin Diseases)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical health at 8 weeks; Group 1: mean 9.2 (SD 18.2); n=169, Group 2: mean 7.6 (SD 15.6); n=169; SF-36 physical health 0-100 Top=High is good outcome; Comments: Reported change scores and standard error. Reported acupuncture: 9.2 (1.4). Reported sham: 7.6 (1.2). Baseline acupuncture: 48.69 (20.44). Baseline sham: 49.65 (19.92).
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 46, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks.; Group 2 Number missing: 68, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks.

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-36 physical health at 26 weeks; Group 1: mean 10.7 (SD 19.1); n=142, Group 2: mean 8.2 (SD 17.8); n=141; SF-36 physical health 0-100 Top=High is good outcome; Comments: Reported change scores and standard error. Reported acupuncture: 10.7 (1.6). Reported sham: 8.2 (1.5). Baseline acupuncture: 48.69 (20.44). Baseline sham: 49.65 (19.92).
Risk of bias: All domain - Very high, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 71, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 18 dropped out at 15-26 weeks.; Group 2 Number missing: 87, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 12 dropped out at 15-36 weeks.

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 14 weeks; Group 1: mean -3.63 (SD 3.9); n=158, Group 2: mean -2.68 (SD 4.14); n=157; WOMAC pain 0-20 Top=High is

poor outcome; Comments: Reported change scores and standard error. Reported acupuncture: -3.63 (0.31). Reported sham: -2.68 (0.33). Baseline acupuncture: 8.92 (3.42). Baseline sham: 8.90 (3.39).
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, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 53, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks.; Group 2 Number missing: 75, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out a t9-14 weeks.

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean -3.79 (SD 3.93); n=142, Group 2: mean -2.92 (SD 3.56); n=141; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reported change scores and standard error. Reported acupuncture: -3.79 (0.33). Reported sham: -2.92 (0.3). Baseline acupuncture: 8.92 (3.42). Baseline sham: 8.90 (3.39).
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, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 71, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 18 dropped out at 15-26 weeks.; Group 2 Number missing: 87, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 12 dropped out at 15-36 weeks.

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 14 weeks; Group 1: mean -12.18 (SD 12.07); n=158, Group 2: mean -9.4 (SD 11.78); n=157; WOMAC function 0-68 Top=High is poor outcome; Comments: Reported change scores and standard error. Reported acupuncture: -12.18 (0.96). Reported sham: -9.4 (0.94). Baseline acupuncture: 31.31 (12.06). Baseline sham: 31.29 (12.00).
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, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 53, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks.; Group 2 Number missing: 75, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out a t9-14 weeks.

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 26 weeks; Group 1: mean -12.42 (SD 13.4); n=142, Group 2: mean -9.88 (SD 11); n=141; WOMAC function 0-68 Top=High is poor outcome; Comments: Reported change scores and standard error. Reported acupuncture: -12.42 (1.12). Reported sham: -9.88 (0.93). Baseline acupuncture: 31.31 (12.06). Baseline sham: 31.29 (12.00).
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, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 71, Reason: 4

declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 18 dropped out at 15-26 weeks.; Group 2 Number missing: 87, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 12 dropped out at 15-36 weeks. Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Serious adverse events at 26 weeks; Group 1: 14/190, Group 2: 5/191; Comments: Acupuncture: Heart disease = 1, cancer, = 2, non-study related injuries = 3, non-arthritis related surgery = 6, stroke = 1, pneumonia = 1. Sham: Non-study related injuries = 1, exacerbation of knee pain = 1, non-arthritis related surgery = 3. 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, education, race, target knees, length of diagnosis of osteoarthritis, walking pain on flat surface, concurrent medications, and baseline values of outcomes; Group 1 Number missing: 71, Reason: 4 declined baseline assessment. 17 were disqualified for medical reasons. 19 dropped out at 0-4 weeks. 6 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 18 dropped out at 15-26 weeks.; Group 2 Number missing: 87, Reason: 8 declined baseline assessment. 27 were disqualified for medical reasons. 28 dropped out at 0-4 weeks. 5 dropped out at 5-8 weeks. 7 dropped out at 9-14 weeks. 12 dropped out at 15-36 weeks.	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study (subsidiary papers)	**Ceballos-laita 2019[17] (Ceballos-laita 2020[18])**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in Spain; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 3 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Unilateral primary hip osteoarthritis according to the clinical criteria of the American College of Rheumatology, a grade II or III Kellgren & Lawrence classification in their most recent hip x-rays, 50-70 years of age, and presence of at least one active MTrP in the hip muscles

Exclusion criteria	Previous lower limb replacement surgery, neurological, vascular or other lower extremity musculoskeletal conditions that affected sensation, gait or functional performance, previous physiotherapy treatment to the hip in the last three months, DN contraindications (local infection, bleeding disorders, immune suppression, or significant fear of needles), previous experience of DN technique to maintain blinding of patients or inability to understand the instructions and complete the study assessments.
Recruitment/selection of patients	Participants were recruited from private practice physiotherapy clinics or referred by general practitioners and orthopedic surgeons
Age, gender and ethnicity	Age - Mean (SD): Dry needling group: 55.5 (4.7); sham group: 58.6 (6.6). Gender (M:F): 17/13. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging (Inclusion criteria specifies K-L grade/classification on hip x-rays). 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hip
Extra comments	Severity of symptoms: Grade K-L II: Dry needling group 9/15; sham group 6/15. Grade K-L III: Dry needling group 6/15; sham group 9/15 Duration of symptoms (mean, SD): Dry needling group 64.4 (79.6) months; sham group 72.2 (91.2) months
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Acupuncture/dry needling - Dry needling. Dry needling was performed by the lead author who had four years of clinical experience. Active MTrPs were located by manual palpation in the hip muscles. They were immobilised between the index and middle finger. Three active MTrPs were treated at most in each session. A standard single-use sterile acupuncture needle (0.25 mm x 50 mm) was inserted perpendicularly through the skin and moved forward until the MTrP was reached. To minimise pain of insertion, a certain pressure was applied to the skin with the insertion tube. Hong's fast-in and fast-out technique was used with the aim of eliciting a local twitch response. After the needle was premoved, pressure with a cotton ball was maintained to prevent bleeding. Patients received three treatment sessions, with one session per week. Duration 3 weeks. Concurrent medication/care: No exercise programme or physical therapy modalities were added to the intervention. Patients were asked not to take any nonsteroidal anti-inflammatory or muscle relaxant drugs. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=15) Intervention 2: Sham acupuncture. Participants received a simulated dry

	needling technique that has been shown to be valid. The blunted needle was applied to MTrPs to provoke a pricking sensation, without penetrating the skin. Sham dry needling was also added in the same regions with the same dose as the dry needling group. Duration 3 weeks. Concurrent medication/care: No exercise programme or physical therapy modalities were added to the intervention. Patients were asked not to take any nonsteroidal anti-inflammatory or muscle relaxant drugs. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DRY NEEDLING versus SHAM ACUPUNCTURE Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: Pain at 3 weeks (end of intervention); Group 1: mean 0.4 (SD 0.8); n=15, Group 2: mean 2.6 (SD 2.5); n=15; VAS 0-10 Top=High is poor outcome; Comments: Baseline scores: Dry needling group: 2.1 (1.8); Sham group 1.3 (1.6) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Psychological distress at ≤ 3 months - Actual outcome: Anxiety and depression at 3 weeks (end of intervention); Group 1: mean 5.4 (SD 4.2); n=15, Group 2: mean 10.4 (SD 3.9); n=15; Hospital Anxiety and Depression Scale 0-21 Top=High is poor outcome; Comments: Baseline values: Dry needling group: 10.1 (5.8); Sham group 10.4 (4.3) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; 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; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	**Ceballos-laita 2021[16]**
Study type	RCT (Patient randomised; Parallel)

Number of studies (number of participants)	1 (n=45)
Countries and setting	Conducted in Spain; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 3 weeks
Method of assessment of guideline condition	Unclear method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Unilateral hip OA according to the American College of Rheumatology criteria, a grade II or III Kellgren & Lawrence classification, age between 50-70 years, and at least 1 active MTrP in the hip muscles. Presence of MTrP was confirmed based on the criteria described by Travell and Simons: (1) presence of a palpable taut band, (2) local pain upon pressure applied to the nodule of the taut band; and (3) reproduction of the patient's pain by palpatation
Exclusion criteria	Neurologic, vascular, or other lower extremity musculo-skeletal conditions that affected sensation, gait, or functional performance, previous surgery in the lower limbs, previous physiotherapy treatment for hip OA in the previous 3 months, MTrP therapy experience, and dry needling contraindications
Recruitment/selection of patients	Participants were recruited from private practice physiotherapy clinics or by general practitioners and orthopedic surgeons
Age, gender and ethnicity	Age - Mean (SD): Dry needling group: 57.53 (3.88); sham group 58.20 (5.08); control group: 54.67 (4.48). Gender (M:F): 20/25. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hip
Extra comments	Severity of symptoms (K-L grade II/III): Dry needling group: 7/8; Sham group: 6/9; control group: 6/9 Duration of symptoms (mean, SD): Dry needling group: 66.33 (76.61) months; sham group: 72.20 (53.76) months; control group: 68.13 (56.36) months
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Acupuncture/dry needling - Dry needling. Interventions were carried out by a physiotherapist with more than 5 years of clinical experience in dry needling therapy. Participants received 3 session of dry needling (1 session per week) into active MTrPs in the hip muscles. Iliopsoas, rectus femoris, tensor fasciae latae, and gluteus minimus muscles were examined for the presence of active MTrPs. At most, 3 active MTrPs were treated during each session. Patients were

	placed in a supine position to treat iliopsoas and rectus femoris muscles or in a contralateral side lying position for tensor fasciae latae and gluteus minimus muscles. The MTrP taut band was held between the physiotherapist's index and middle fingers while a 0.25 x 50 mm needles was inserted using the fast-in fast-out technique. This technique consists of rapid multiple introductions of the needle into the MTrP. When the needle mechanically stimulates the MTrP, a brisk contraction of the taut band, called local twitch response, can be elicited. The needle was repeatedly inserted until the local twitch responses became extinct. After the needle was removed, the injected area was compressed firmly to achieve hemostasis. Duration 3 weeks. Concurrent medication/care: All participants were asked to continue with the same daily routines and not to take any analgesic, anti-inflammatory or muscle relaxant medications 24 hours prior to testing. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=15) Intervention 2: Sham acupuncture. Participants received three sessions of a sham needle procedure (one per week). At most, three active MTrPs were treated during each session with the sham needle procedure using a blunted needle with insertion tube. The needle was placed on the MTrP area and was pressed up and down against the skin without penetrating. Duration 3 weeks. Concurrent medication/care: All participants were asked to continue with the same daily routines and not to take any analgesic, anti-inflammatory or muscle relaxant medications 24 hours prior to testing. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=15) Intervention 3: No intervention - Acupuncture compared to no treatment. Control group participants did not receive any treatment, education or advice during the study. Duration 3 weeks. Concurrent medication/care: All participants were asked to continue with the same daily routines and not to take any analgesic, anti-inflammatory or muscle relaxant medications 24 hours prior to testing. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DRY NEEDLING versus SHAM ACUPUNCTURE Protocol outcome 1: Pain at ≤ 3 months	

- Actual outcome: Pain at 3 weeks; Group 1: mean 3.4 (SD 1.95); n=15, Group 2: mean 5.87 (SD 2.94); n=15; WOMAC-Pain 0-20 Top=High is poor outcome; Comments: Baseline: Dry needling group: 8.13 (3.09); sham group: 6.53 (3.29) Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Dry needling group had higher WOMAC pain and physical function scores at baseline but this was not significant; Blinding details: The participants in the dry needling and sham groups were blinded, but those in the no treatment groups were not; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: Physical function at 3 weeks; Group 1: mean 10.47 (SD 6.18); n=15, Group 2: mean 21.73 (SD 4.71); n=15; WOMAC-Physical function 0-68 Top=High is poor outcome; Comments: Baseline scores: Dry needling group: 25.4 (8.6); Sham group: 19.27 (7.08) Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Dry needling group had higher WOMAC pain and physical function scores at baseline but this was not significant; Blinding details: The participants in the dry needling and sham groups were blinded, but those in the no treatment groups were not; Group 1 Number missing: 0; Group 2 Number missing: 0 RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DRY NEEDLING versus ACUPUNCTURE COMPARED TO NO TREATMENT Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: Pain at 3 weeks; Group 1: mean 3.4 (SD 1.95); n=15, Group 2: mean 6.27 (SD 2.65); n=15; WOMAC-pain 0-20 Top=High is poor outcome; Comments: Baseline scores: Dry needling group: 8.13 (3.09); No treatment group: 6.8 (2.48) 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 ; Baseline details: Dry needling group had higher WOMAC pain and physical function scores at baseline but this was not significant; Blinding details: The participants in the dry needling and sham groups were blinded, but those in the no treatment groups were not; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: Physical function at 3 weeks; Group 1: mean 10.47 (SD 6.18); n=15, Group 2: mean 23.53 (SD 9.64); n=15; WOMAC - physical function 0-68 Top=High is poor outcome; Comments: Baseline scores: Dry needling group: 25.4 (8.6); No treatment group: 22.73 (9.72) 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 ; Baseline details: Dry needling group had higher WOMAC pain and physical function scores at baseline but this was not significant; Blinding details: The participants in the dry needling and sham groups were blinded, but those in the no treatment groups were not; 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; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Chen 2013[19]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=213)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks with an additional 14 weeks of follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with pain in 1 or both knee joints for more than 6 months with radiological Kellgren-Lawrence grade 2-3 osteoarthritic changes
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 40 years or older to focus on classic knee osteoarthritis; pain in 1 or both knee joints for more than 6 months and moderate pain >4/10 for more than 5 out of 7 consecutive days in the week before enrollment. Between the enrollment and first treatment visit, the degree of knee osteoarthritis was radiologically confirmed as KellgrenLawrence score 2 or 3, in 1 or both knees on a radiograph obtained within the last year or on an x-ray performed as part of the study
Exclusion criteria	If they had other diseases known to affect the knee including gout, rheumatoid arthritis and significant trauma; neurologic, cardiac or psychiatric disease that would interfere with a standard EPT program; pregnancy; significant coagulopathy or taking anti-coagulants that would interfere with the safe administration of acupuncture; previous acupuncture treatment within the last 12 months
Recruitment/selection of patients	People were recruited from 3 physical therapy sites in Philadelphia, Pennsylvania based on any person being referred for physical therapy from any discipline of medicine
Age, gender and ethnicity	Age - Mean (SD): 60.5 (11.4). Gender (M:F): 103:110. Ethnicity: White = 62, African America = 141, Other = 10
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3, median grade 3 Duration of symptoms (mean [SD]): 9.5 (9.6) years
Indirectness of population	No indirectness

Interventions	(n=105) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture sessions with the puncturing needles were adminsitered following every exercise session one or twice a week by fully trained acupuncturists, without electrical stimulation or co-intervention. The penetrating (gauge 8 x 1.2") and identical appearing non-penetrating Streitberger needles were used. Nine acupuncture points for each knee were chosen to be consistent with the traditional Chinese Bi syndrome therapy for knee pain. The primary knee points were GB 34, SP 9, ST 36, ST 35 and Xiyan, and the distal points (located near the ankles) selected were UB 60, GB 39, SP 6, and KI 3 for a total of 9 points. The same points were used for each affected leg. If both knees had pain >3/10, both were treated. The insertion depth for standard needles was between 0.2 to 3cm depending on the location of the point and patient's body size. The needles were left in place for 20 minutes, with a brief manipulation at the beginning and end of the treatment. The de qi sensation was not required and not specifically recorded.. Duration 12 weeks. Concurrent medication/care: All people received exercise-based physical therapy once or twice a week for a maximum of 12 total treatments. The standardized program was as vigorous as the person could tolerate with routine encouragement and included range of motion exercises, muscle strengthening and aerobic conditioning. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=109) Intervention 2: Sham acupuncture. The Streitberger non-penetrating needle was used. The needle appears identical except that it is blunt and retracts into the handle when it is pressed against the skin, giving the appearance and sensation of needle insertion. Both puncturing and non0puncturing needles were played in the same points and held in place by being inserted through a single-layer gauze retaining mechanism held on by a small doughnut-shaped bandage. Acupuncturists were instructed to not to attempt to stimulate with the Streitberger needle and did not ask about the achievement of de qi to minimize the interaction between the acupuncturist and the patient.. Duration 12 weeks. Concurrent medication/care: All people received exercise-based physical therapy once or twice a week for a maximum of 12 total treatments. The standardized program was as vigorous as the person could tolerate with routine encouragement and included range of motion exercises, muscle strengthening and aerobic conditioning. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling:
Funding	Academic or government funding (This study is supported by a grant from the National Institutes of Health/National Center for Complementary and Alternative Medicine (NCCAM) R01-AT000304. Dr Mao is also supported by NCCAM K23 AT004112. The funding agency had no role in the design and conduct of this study.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical subscale at 12 weeks; Group 1: mean 3.01 (SD 10); n=104, Group 2: mean 4.3 (SD 9.03); n=109; SF-36 physical subscale 0-100 Top=High is good outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: 3.01 (1.09, 4.93). Reported sham: 4.30 (2.61, 6.00). Baseline acupuncture: 31.7 (30.2, 33.3). Baseline sham: 32.1 (30.5, 33.7).
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, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 18, Reason: 8 lost to follow up, 4 withdrew, 2 violations, 3 adverse events, 1 other; Group 2 Number missing: 15, Reason: 7 lost to follow up, 6 withdrew, 2 violations
- Actual outcome: SF-36 mental subscale at 12 weeks; Group 1: mean 5.01 (SD 11.1); n=104, Group 2: mean 3.33 (SD 10.5); n=109; SF-36 mental subscale 0-100 Top=High is good outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: 5.01 (2.88, 7.15). Reported sham: 3.33 (1.30, 5.24). Baseline acupuncture: 46.2 (43.8, 48.5). Baseline sham: 49.7 (47.2, 52.2).
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, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 18, Reason: 8 lost to follow up, 4 withdrew, 2 violations, 3 adverse events, 1 other; Group 2 Number missing: 15, Reason: 7 lost to follow up, 6 withdrew, 2 violations

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean -2.83 (SD 4.6); n=104, Group 2: mean -2.35 (SD 3.25); n=109; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: -2.83 (-3.71, -1.94). Reported sham: -2.35 (-2.96, -1.74). Baseline acupuncture: 10.31 (9.65, 11.0). Baseline sham: 9.47 (8.83, 10.1).
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, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 18, Reason: 8 lost to follow up, 4 withdrew, 2 violations, 3 adverse events, 1 other; Group 2 Number missing: 15, Reason: 7 lost to follow up, 6 withdrew, 2 violations

Protocol outcome 3: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean -1.91 (SD 3.77); n=104, Group 2: mean -1.35 (SD 3.76); n=109; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: -1.91 (-2.63, -1.18). Reported sham: -1.35 (-2.05, -0.64). Baseline acupuncture: 10.31 (9.65, 11.0). Baseline sham: 9.47 (8.83, 10.1).
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: Reported age, gender, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 34, Reason: 18 lost to follow up, 4 withdrew, 2 violations, 2 surgery/injection/injury, 3 missed, 3 adverse events, 2 other; Group 2 Number missing: 27, Reason: 14 lost to follow up, 6 withdrew, 3 surgery/injection/injury, 2 violations, 1 missed, 1 other

Protocol outcome 4: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 12 weeks; Group 1: mean -7.02 (SD 13.16); n=104, Group 2: mean -6.93 (SD 10.97); n=109; WOMAC function 0-68 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: -7.02 (-9.55, -4.49). Reported sham: -6.93 (-8.99, -4.87). Baseline acupuncture: 32.9 (30.8, 35.0). Baseline sham: 30.3 (28.1, 32.6). 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, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 18, Reason: 8 lost to follow up, 4 withdrew, 2 violations, 3 adverse events, 1 other; Group 2 Number missing: 15, Reason: 7 lost to follow up, 6 withdrew, 2 violations Protocol outcome 5: Physical function at > 3 months - Actual outcome: WOMAC function at 26 weeks; Group 1: mean -4.02 (SD 13.35); n=104, Group 2: mean -3.74 (SD 10.76); n=109; WOMAC function 0-68 Top=High is poor outcome; Comments: Reported change scores and 95% confidence intervals. Reported acupuncture: -4.02 (-6.59, -1.46). Reported sham: -3.74 (-5.76, -1.72). Baseline acupuncture: 32.9 (30.8, 35.0). Baseline sham: 30.3 (28.1, 32.6). 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: Reported age, gender, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 34, Reason: 18 lost to follow up, 4 withdrew, 2 violations, 2 surgery/injection/injury, 3 missed, 3 adverse events, 2 other; Group 2 Number missing: 27, Reason: 14 lost to follow up, 6 withdrew, 3 surgery/injection/injury, 2 violations, 1 missed, 1 other Protocol outcome 6: Serious adverse events at > 3 months - Actual outcome: Adverse events at 26 weeks; Group 1: 47/105, Group 2: 31/109; Comments: Denominator unclear, but assumed to be number of participants. Acupuncture: agitation = 2, bruising = 1, fatigue = 1, increased pain = 22, redness/infection = 1, muscle soreness = 6, swelling = 6, weakness = 1, other = 7. Sham: bruising = 1, fatigue = 1, increased pain = 16, muscle soreness = 2, swelling = 5, tearfulness = 1, other = 5. 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: Reported age, gender, race, marital status, occupation, BMI, Kellgren Lawrence grade, pain duration, and baseline values of outcomes; Group 1 Number missing: 34, Reason: 18 lost to follow up, 4 withdrew, 2 violations, 2 surgery/injection/injury, 3 missed, 3 adverse events, 2 other; Group 2 Number missing: 27, Reason: 14 lost to follow up, 6 withdrew, 3 surgery/injection/injury, 2 violations, 1 missed, 1 other	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	Dunning 2018[26]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=242)
Countries and setting	Conducted in USA; 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 meeting the American College of Rheumatology criteria for the diagnosis of knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People meeting the American College of Rheumatology criteria for the diagnosis of knee osteoarthritis and have had chronic pain in the knee joint for >3 months. People had to have at least 3 of the following criteria: 1) above 50 years of age; 2) <30 minutes of morning stiffness; 3) crepitus on active motion; 4) bony tenderness; 5) bony enlargement; 6) no palpable warmth of synovium. In addition, participants had to have a minimum knee pain intensity score of 2 points and be older than 18 years of age.
Exclusion criteria	History of surgery to the painful knee; a history of surgery to either of the lower extremities in the last 6 months; any red flags to manual therapy, dry needling or exercise; had received physical therapy, acupuncture, massage therapy, chiropractic, or intra-articular injections for the painful knee in the last 3 months; presented with at least 2 positive neurological signs; had involvement in litigation or worker's compensation regarding their knee pain; pregnancy.
Recruitment/selection of patients	People were recruited from 18 outpatient physical therapy clinics in 10 different states in the United States of America
Age, gender and ethnicity	Age - Mean (SD): 57.6 (13.2). Gender (M:F): 111:111. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 4.6 (4.9) years
Indirectness of population	No indirectness
Interventions	(n=121) Intervention 1: Electroacupuncture. This group received electrical dry needling using a standardized 9-point protocol for 20 to 30 minutes on each treatment session. 8 to 10 sessions of periosteal electrical dry needling at a frequency of 1 to 2

	times per week over 6 weeks. Electrical dry needling included a 9-point standardized protocol. In addition, people were also permitted to insert needles at up to 4 additional locations based on the presence of the symptoms. Sterilized disposable stainless steel acupuncture needles were used with 3 sizes: 0.25mmx30mm, 0.30mmx40mm, and 0.30mmx50mm. The depth of needle insertion ranged from 15 to 45mm and depended on the point selected (intramuscular, periosteal, joint line, intra/periarticular) and the patient's physical constitution. Following topical skin cleansing, all needles were inserted and then manipulated bidirectionally to illicit a sensation of aching, tingling, deep pressure, heaviness or warmth. In addition, at least 3 of the 9 obligatory needles (ie, over the posteromedial aspect of the medial tibial condyle, within the depression posterior to the femoral epicondyle, and over the anterolateral crest of the tibia 1 fingerbreadth lateral to the tibial tuberosity) were repeatedly thrusted and tapped on to the respective bone using a "periosteal stimulation" technique. Notably, with the exception of 2 obligatory needles inserted at the level of the tibiofemoral joint margin within the medial or lateral infrapatellar sulcus, and depending on the patient's physical constitution, the needle length selected by the practitioner and the patient's tolerance to such, the remaining needles were also advanced toward the underlying bone to facilitate direct mechanical and electrical "periosteal stimulation". The needles were then left in situ for 20 to 30 minutes with electrical stimulation in pains (crossing through the knee joint in a superior-inferior and diagonal orientation) using 4 channels to 8 of the needles using a low frequency (2Hz), moderate pulse duration (250 microseconds), biphasic continuous waveform at a maximum tolerable intensity. In cases of bilateral knee osteoarthritis, both knees were treated, but only the most painful side at baseline was recorded. Duration 6 weeks. Concurrent medication/care: Both groups received manual therapy (passive joint mobilizations and muscle stretching) and exercise (riding a stationary bicycle, range of motion, and strengthening exercises to the lower extremity) on each session. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=121) Intervention 2: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. No acupuncture. Duration 6 weeks. Concurrent medication/care: Both groups received manual therapy (passive joint mobilizations and muscle stretching) and exercise (riding a stationary bicycle, range of motion, and strengthening exercises to the lower extremity) on each session. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: WOMAC pain at 3 months; Group 1: mean -5.9 (SD 3.3); n=111, Group 2: mean -2.8 (SD 3.2); n=111; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 8.7 (3.2). Baseline no treatment: 8.0 (3.3). 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: Reported sex, age, weight, height, years with knee pain, medication intake, number of treatment sessions, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 did not return the follow up questionnaire; Group 2 Number missing: 4, Reason: 1 did not return, 3 did not return the follow up questionnaire Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: WOMAC physical function at 3 months; Group 1: mean -18.8 (SD 10.6); n=111, Group 2: mean -9.4 (SD 9.8); n=111; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 28.9 (10.6). Baseline no treatment: 28.1 (11.1). 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: Reported sex, age, weight, height, years with knee pain, medication intake, number of treatment sessions, and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 did not return the follow up questionnaire; Group 2 Number missing: 4, Reason: 1 did not return, 3 did not return the follow up questionnaire	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	**Farazdaghi 2021[30]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Iran; Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention + follow up: 1 week + 2 weeks

Method of assessment of guideline condition	Unclear method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Moderate osteoarthritis symptoms (grade 2-3 of Kellgren-Lawrence Classification System) in muscles around the hip and knee joint
Exclusion criteria	Severe knee inflammation/swelling, any history of joint injection/aspiration within one year, any previous spinal or lower extremity fracture, related neurological pathology, vestibular pathology, systemic disease (such as rheumatoid arthritis or diabetes), immune deficiency diseases, bleeding disorders (such as hemophilia or thalassemia), cancer, trypanphobia, or if they were taking antiplatelet medications
Recruitment/selection of patients	Patients were recruited from orthopedic and rehabilitation clinics associated with Shiraz University of Medical Sciences
Age, gender and ethnicity	Age - Mean (SD): Dry needling group: 61.00 (8.19); sham group: 56.20 (6.03). Gender (M:F): All female. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Hip and knee).
Extra comments	Severity of symptoms: Grade 2-3 of Kellgren-Lawrence Classification criteria Duration of symptoms: Not reported
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Acupuncture/dry needling - Dry needling. The therapist searched for trigger points around hip and knee joint by the algometer to quantify pain sensitivity on MTrPs. Dry needling involved three repetitive measures at each site of MTrP. At least two hyperalgesic points showing radicular pain, jumping sign or abrupt response were marked for treatment. The dry needling technique consisted of insertion of a disposable 0.25 x 40 mm stainless steel acupuncture needle. The "sparrow pecking technique" (in-and-out motion) was performed on selected MTrPs in multiple directions with "coning technique". Dry needling was performed in supine position for hip adductors/abductors, flexors, and knee extensors. The technique continued in prone positions for hip adductors/abductors, extensors, popliteous and knee flexors. The therapist held the guide of the needle between the thumb and index finger of non-dominant hand perpendicular to the MTrP and then inserted the needles with the swift push by the second finger tip of the dominant hand. The therapist probed the needle in different angles until he perceived a switch response, or pain response, or referral pain of the MTrP, otherwise the technique lasted about 5-10 seconds depending on the patients

	tolerance. Patients received three sessions over one week. Duration 1 week. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=20) Intervention 2: Sham acupuncture. A plastic cover of a needle was used. The plastic cover was pushed against the skin with a quick force to mimic sensation of a needle insertion. Patients received three sessions over one week . Duration 1 week. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Academic or government funding (Funded by Shiraz University of Medical Sciences)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DRY NEEDLING versus SHAM ACUPUNCTURE Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: Pain at 3 weeks; Group 1: mean 4.73 (SD 2.18); n=20, Group 2: mean 7.53 (SD 1.68); n=20; VAS 0-10 Top=High is poor outcome; Comments: Baseline scores: Dry needling group: 7.4 (1.72); sham group: 6.47 (1.30) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: Knee impairement at 3 weeks; Group 1: mean 49.75 (SD 9.62); n=20, Group 2: mean 37.91 (SD 10.94); n=20; KOOS 0-100 Top=High is good outcome; Comments: Baseline values: Dry needling group: 36.55 (11.04); sham group 43.35 (12.46) Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; 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; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Fink 2001[34]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=65)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People with pain, discomfort and movement restriction in the hip with radiographic changes of at least 2 on a Kellgren-Lawrence score
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Discomfort or painful areas in the gluteal and trochanter region, pain irradiation tot he lateral side of the upper limb, complaints for at least 6 months with pain on most days of the previous month; movement restriction with internal-rotation less than 15 degrees or painful internal rotation more than 15 degrees and flexed hip less than 115 degrees; radiographic changes of the hip on a pelvic x-ray, not older than 1 year of at least 2 on a Kellgren-Lawrence score
Exclusion criteria	Scars or sensibility problems around the acupuncture area; skin implantation around the acupuncture area; acute dermatosis or wounds around the acupuncture area; serious circulatory problems (e.g. chronic venous insufficiency, gangrene); dermatitis, contact allergies, psoriasis, herpes; immune deficiency syndrome (i.e. AIDS, iatrogenic following transplants); damaged or implanted heart valves; systemic illnesses which during their duration could relate to the hip joint (e.g. chronic polyarthritis, metabolic illness such as gout or chondrocalcinosis); treatment within the last 4 weeks which could lead to misinterpretation of the outcome, physical therapy, regular intake of analgesics or NSAID
Recruitment/selection of patients	People were recruited through advertisements in a local newspaper
Age, gender and ethnicity	Age - Mean (SD): 62.6 (9.1). Gender (M:F): 22:43. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Hip
Extra comments	Severity: Kellgren Lawrence grade 2-4, median grade 3 Duration of symptoms (mean [SD]): 5.2 (3.8) years
Indirectness of population	No indirectness

Interventions	(n=33) Intervention 1: Acupuncture/dry needling - Acupuncture. Ten individual treatments which was performed within 3 weeks. Site selection included: six pressure sensitive locations ('Ah-Shi' points). In addition, the regional meridian points 'GB-30', 'GB-31', 'BL-37' and the distal meridian points 'ST-40' and 'BL-54' were chosen as well as the master point for tendons and muscles 'GB-34'. Needle treatment was continued for 20 minutes with twisting of the needles to cause a mechanical stimulation. It was the aim of this manipulation to elicit the 'deqi'-sensation. Needle manipulation was carried out 2 to 3 times during a treatment session. Needle acupuncture was performed using identical sets of sterile, steel, disposable needles (0.3 x 60mm).. Duration 3 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=34) Intervention 2: Sham acupuncture. Acupuncture performed in the same way as the acupuncture group, apart from selected puncture sites were at least 5cm away from the classical acupuncture points and their interconnecting lines (meridian) and also clear of the painful pressure points (Ah-Shi or trigger points).. Duration 3 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling:
Funding	Study funded by industry (This study was supported by a grant from the PharmaMED foundation Germany)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE Protocol outcome 1: Serious adverse events at ≤ 3 months - Actual outcome: Adverse events at 3 weeks; Group 1: 0/33, Group 2: 0/32 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: Reported age, height, weight, BMI, duration of symptoms, radiographic stage, psychological scores and baseline values of outcomes; Group 1 Number missing: 16, Reason: 1 person did not turn up, 7 total hip replacement, 8 excluded for other reason; Group 2 Number missing: 10, Reason: 2 treatment discontinued, 2 did not turn up, 4 total hip replacement, 2 excluded for other reasons	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at ≤ 3 months; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Foster 2007[37] (Whitehurst 2011[153])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=352)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks of intervention, 12 months of follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical diagnosis of knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Male and female subjects aged 50 years and above with pain (with or without stiffness) in one or both knees presented to primary care. They must be naive to acupuncture treatment (i.e. have never experienced acupuncture before for their present or any past complaints) and considered suitable for referral to a physiotherapy outpatients department by their general practitioner. People must be able to read and write English, be willing to consent to participation and able to give full informed consent. They must also be available for telephone contact.
Exclusion criteria	People with potentially serious pathology (e.g. inflammatory arthritis, malignancy etc) on the basis of general practice or physiotherapy diagnosis or from past medical history; those who have had a knee or hip replacement on the affected side(s); are already on a surgical waiting list for total knee replacement; or for whom the trial interventions are contraindicated; those who have received an exercise program, from a physiotherapist, for their knee problem within the last 3 months; an intra-articular injection to the knee in the last 6 months
Recruitment/selection of patients	People were recruited from 37 NHS physiotherapy centers providing services for general practices within the Midlands and Cheshire regions of the United Kingdom. People referred for physiotherapy by their GP.
Age, gender and ethnicity	Age - Mean (SD): 63.2 (8.8). Gender (M:F): 136:216. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: <1 - at least 10 years, median time 1 to <5 years.
Indirectness of population	No indirectness

Interventions	(n=117) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture using 6-10 points of 16 local and distal points per session: Sp 9, Sp 10, St 34, St 35, St 36, Xiyan, Gb 34 and trigger points,. Distal points available include: LI 4, tH 5, Sp 6, Liv 3, St 44, Ki 3, Bl 60 and Gb 41. Treatment was performed with sterilised steel needles, 30 x 0.3mm. The depth of the needle insertion was between 0.5-2.5cm dependent on the points selected for treatment and the needles were manipulation until de-qi sensation was achieved. The needles were left for 25-35 minutes and could be manipulated to elicit the de-qi sensation. 6 sessions over 3 weeks. Duration 6 weeks (3 weeks for the acupuncture). Concurrent medication/care: Advice and exercise. Advice is given by a leaflet that contains standard advice on the use of analgesia. If people are taking NSAIDs, they were permitted to continue their stable dose. Exercise was conducted as a program with a maximum of 6x 30 minute sessions over 6 weeks including concentric, eccentric, isometric and balance exercises. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=119) Intervention 2: Sham acupuncture. Sham acupuncture using needles that collapse in the handle creating an illusion of insertion. A minimum of 6 needles were inserted for 6x 30 minute sessions over 3 weeks. Duration 6 weeks (3 weeks for the acupuncture). Concurrent medication/care: Advice and exercise. Advice is given by a leaflet that contains standard advice on the use of analgesia. If people are taking NSAIDs, they were permitted to continue their stable dose. Exercise was conducted as a program with a maximum of 6x 30 minute sessions over 6 weeks including concentric, eccentric, isometric and balance exercises. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=116) Intervention 3: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. Advice and exercise only. Duration 6 weeks. Concurrent medication/care: Advice and exercise. Advice is given by a leaflet that contains standard advice on the use of analgesia. If people are taking NSAIDs, they were permitted to continue their stable dose. Exercise was conducted as a program with a maximum of 6x 30 minute sessions over 6 weeks including concentric, eccentric, isometric and balance exercises. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (This study was supported by a project grant from the Arthritis Research Campaign, UK (grant H0640) and Support for Science funding secured by the North Staffordshire Primary Care Research Consortium for NHS service support costs. NEF is funded by a primary care career scientist award from the

	Department of Health and NHS research and development, UK. JCH is funded by a lectureship in physiotherapy from the Arthritis Research Campaign, UK.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean -2.83 (SD 4); n=113, Group 2: mean -3.02 (SD 3.6); n=115; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 9.3 (4.0). Baseline sham: 8.9 (3.3).
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 4, Reason: 1 withdrawal, 3 no questtionnaire

Protocol outcome 2: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean -2.37 (SD 4.2); n=99, Group 2: mean -2.82 (SD 4.1); n=105; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 9.3 (4.0). Baseline sham: 8.9 (3.3).
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 18, Reason: 3 withdrawal, 13 no questionnaire (some questionnaires were not valid); Group 2 Number missing: 14, Reason: 2 withdrawal, 11 no questionnaire (some questionnaires were not valid)

Protocol outcome 3: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 6 weeks; Group 1: mean -8.18 (SD 11.5); n=113, Group 2: mean -9.32 (SD 11.4); n=110; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 30.8 (13.9). Baseline sham: 31.1 (12.8).
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 4, Reason: 1 withdrawal, 3 no questtionnaire

Protocol outcome 4: Physical function at > 3 months
- Actual outcome: WOMAC function at 12 months; Group 1: mean -6.61 (SD 13.8); n=100, Group 2: mean -8.24 (SD 13.5); n=104; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 30.8 (13.9). Baseline sham: 31.1 (12.8).
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 17, Reason: 3 withdrawal, 13 no questionnaire (some questionnaires were not valid); Group 2 Number missing: 15, Reason: 2 withdrawal, 11 no questionnaire (some questionnaires were not valid)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT

COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: EQ-5D at 6 weeks; Group 1: mean 0.663 (SD 0.24); n=117, Group 2: mean 0.639 (SD 0.27); n=116; EQ-5D 0-1 Top=High is good outcome; Comments: Baseline acupuncture: 0.562 (0.28). Baseline no treatment: 0.603 (0.25)
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes. EQ-5D value different at baseline.; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 11, Reason: 3 withdrawals, 8 no questionnaire

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: EQ-5D at 12 months; Group 1: mean 0.621 (SD 0.3); n=117, Group 2: mean 0.616 (SD 0.3); n=116; EQ-5D 0-1 Top=High is good outcome; Comments: Baseline acupuncture: 0.562 (0.28). Baseline no treatment: 0.603 (0.25)
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: Reported age, gender, BMI, socioeconomic class, employment, drug use and baseline values of outcomes. EQ-5D value different at baseline.; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 11, Reason: 3 withdrawals, 8 no questionnaire

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 6 weeks; Group 1: mean -2.83 (SD 4); n=113, Group 2: mean -2.1 (SD 3.5); n=105; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 9.3 (4.0). Baseline no treatment: 9.1 (3.7).
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, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 11, Reason: 3 withdrawals, 8 no questionnaire

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean -2.37 (SD 4.2); n=99, Group 2: mean -2.57 (SD 4.3); n=98; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 9.3 (4.0). Baseline no treatment: 9.1 (3.7).
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, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 18, Reason: 3 withdrawals, 13 no questionnaire (some questionnaires were not valid); Group 2 Number missing: 18, Reason: 6 withdrawals, 10 no questionnaire (some questionnaires were not valid)

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 6 weeks; Group 1: mean -8.18 (SD 11.5); n=113, Group 2: mean -6.21 (SD 11.4); n=105; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 30.8 (13.9). Baseline no treatment: 29.0 (12.9).
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, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 4, Reason: 4 no questionnaire; Group 2 Number missing: 11, Reason: 3 withdrawals, 8 no questionnaire Protocol outcome 6: Physical function at > 3 months - Actual outcome: WOMAC function at 12 months; Group 1: mean -6.61 (SD 13.8); n=100, Group 2: mean -5.36 (SD 11.9); n=97; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 30.8 (13.9). Baseline no treatment: 29.0 (12.9). 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, BMI, socioeconomic class, employment, drug use and baseline values of outcomes; Group 1 Number missing: 18, Reason: 3 withdrawals, 13 no questionnaire (some questionnaires were not valid); Group 2 Number missing: 18, Reason: 6 withdrawals, 10 no questionnaire (some questionnaires were not valid)	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Hinman 2014[47]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=282)
Countries and setting	Conducted in Australia; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks of treatment, 1 year total follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee pain on most days with an average severity of 4 or more out of 10 on a NRS and had morning stiffness lasting less than 30 minutes (consistent with a clinical diagnosis of osteoarthritis)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 50 years or older with knee pain for longer than 3 months duration, knee pain on most days with an average severity of 4 or more out of 10 on a NRS and had morning stiffness lasting less than 30 minutes (consistent with a clinical diagnosis of osteoarthritis)
Exclusion criteria	History of any systemic arthritic condition; history of knee arthroplasty on the most painful knee; wait-listed for any knee surgery for either knee; history of any knee surgery in previous 6 months; any other condition affecting lower limb function (eg trauma, malignancy, neurological condition); history of any knee injection in the past 6 months (eg cortisone, hyaluronic acid); current use of oral or injectable anticoagulant medication; use of acupuncture in past 12 months; any bleeding disorder; allergy to light; referral to pain clinic and use of morphine or pethidine within past 6 months; any other medical condition precluding participation in the trial (eg kidney or liver disease, deep vein thrombosis); knee pain subject to compensation claim; unable to give written informed consent
Recruitment/selection of patients	People recruited from metropolitan Melbourne and regional Victoria via advertisements in the community, media and medical/physical therapy clinics
Age, gender and ethnicity	Age - Mean (SD): 63.6 (8.4). Gender (M:F): 143:139. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: <1 to at least 10 years, median 5 to 10 years
Indirectness of population	No indirectness

Interventions	(n=70) Intervention 1: Acupuncture/dry needling - Acupuncture. Twenty minute treatments delivered once or twice weekly for 12 weeks, with 8 to 12 sessions in total permitted. Acupuncturists treated participants according to usual practice using a standardized set of acupuncture points from around the knee as well as distal points (SP 9 or 10, ST 34, 35 and 35, LR 7, 8 and 9, KI10, BL39, 40 and 57, GB34, 35 and 36 with local extra points in the hamstring muscle, ST40, LR3, SP6, GB41, BL60, BL21, 22 and 23, GB30 and 31, Ear Knee point, DU20, LI11, GV14 and BL11). Single used Seirin needles (0.25 x 40mm) were used for needle acupuncture (administered with the person lying down and needles left in situ while the person rested). . Duration 12 weeks. Concurrent medication/care: Not stated. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Mixed (Combined Western and traditional Chinese medicine style of acupuncture). (n=71) Intervention 2: No intervention - Acupuncture compared to no treatment. No treatment. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable (n=70) Intervention 3: Sham acupuncture. Sham laser acupuncture. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: As this group involved sham laser acupuncture rather than sham needle acupuncture, this group was not included in the final analysis as it did not fulfill the inclusion criteria in the protocol (n=71) Intervention 4: Other. Laser acupuncture. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (This trial was funded by the National Health and medical Research Council (project 566783). Drs Hinman and Bennell are both funded in part by Australian Research Council Future Fellowships (FTI30100175 and FT0991413, respectively). Dr McCrory is funded in party by a National Health and medical Research Council Practitioner Fellowship (1026383). Dr Pirotta is funded in part by a National Health and medical Research Council Career Development Fellowship (1050830). Dr Williamson was funded in part by a National Health and medical Research Council grant (1004233).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE COMPARED TO 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 40.7 (SD 9.6); n=64, Group 2: mean 39.5 (SD 10.7); n=69; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Baseline needle: 36.6 (9.0). Baseline control: 39.2 (9.0).
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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 6, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem); Group 2 Number missing: 2, Reason: 2 lost to follow up, 2 not interested
- Actual outcome: SF-12 mental component summary at 12 weeks; Group 1: mean 51.5 (SD 11); n=64, Group 2: mean 55.8 (SD 9.1); n=69; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Baseline needle: 51.3 (11.4). Baseline control: 55.6 (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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 6, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem); Group 2 Number missing: 2, Reason: 2 lost to follow up, 2 not interested

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-12 physical component summary at 12 months; Group 1: mean 41.7 (SD 10.8); n=59, Group 2: mean 38.9 (SD 11.2); n=62; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Baseline needle: 36.6 (9.0). Baseline control: 39.2 (9.0).
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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 11, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up at 12 weeks (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem). 5 lost to follow-up at 1 year, 4 not interested, 1 family illness.; Group 2 Number missing: 9, Reason: 2 lost to follow up at 12 weeks, 2 not interested. 8 lost to follow up at 1 year, 5 not interested, 2 family illness, 1 other medical problem (1 rejoined at 1 year after having been lost to follow up at 12 weeks)
- Actual outcome: SF-12 mental component summary at 12 months; Group 1: mean 51.1 (SD 11); n=59, Group 2: mean 54.4 (SD 10.2); n=62; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Baseline needle: 51.3 (11.4). Baseline control: 55.6 (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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 11, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up at 12 weeks (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem). 5 lost to follow-up at 1 year, 4 not interested, 1 family illness.; Group 2 Number missing: 9, Reason: 2 lost to follow up at 12 weeks, 2 not interested. 8 lost to follow up at 1 year, 5 not interested, 2 family illness, 1 other medical problem (1 rejoined at 1 year after having been lost to follow up at 12 weeks)

Protocol outcome 3: Pain at ≤ 3 months

- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 6.7 (SD 3.8); n=64, Group 2: mean 7.3 (SD 3.9); n=69; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline needle: 9.0 (3.3). Baseline control: 7.8 (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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 6, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem); Group 2 Number missing: 2, Reason: 2 lost to follow up, 2 not interested

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean 6.7 (SD 4); n=59, Group 2: mean 7.4 (SD 4.1); n=62; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline needle: 9.0 (3.3). Baseline control: 7.8 (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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 11, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up at 12 weeks (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem). 5 lost to follow-up at 1 year, 4 not interested, 1 family illness.; Group 2 Number missing: 9, Reason: 2 lost to follow up at 12 weeks, 2 not interested. 8 lost to follow up at 1 year, 5 not interested, 2 family illness, 1 other medical problem (1 rejoined at 1 year after having been lost to follow up at 12 weeks)

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 22.5 (SD 13.1); n=64, Group 2: mean 23 (SD 13.2); n=69; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline needle: 31.3 (11.8). Baseline control: 26.1 (12.4).
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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 6, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem); Group 2 Number missing: 2, Reason: 2 lost to follow up, 2 not interested

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 12 months; Group 1: mean 22.4 (SD 14.1); n=59, Group 2: mean 23.6 (SD 13.4); n=62; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline needle: 31.3 (11.8). Baseline control: 26.1 (12.4).
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: Differences in SF-12 and WOMAC outcomes at baseline; Group 1 Number missing: 11, Reason: 13 people decline invitation for treatment but were followed up (10 people at week 12 reassessment). 6 were lost to follow up at 12 weeks (3 not interested, 1 time commitment, 1 increased pain, 1 other medical problem). 5 lost to follow-up at 1 year, 4 not interested, 1 family illness.; Group 2 Number missing: 9, Reason: 2 lost to follow up at 12 weeks, 2 not interested. 8 lost to follow up at 1 year, 5 not interested, 2 family illness, 1 other medical problem (1 rejoined at 1 year after having been lost to follow up at 12 weeks)

Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Ju 2015[54]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=80)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoathritis according to the criteria in the American College of Rheumatology and the presence of a severity grade of 2 or 3 according to the radiological Kellgren classification
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed with knee osteoarthritis according to the criteria in the American College of Rheumatology and the presence of a severity grade of 2 or 3 according to the radiological Kellgren classification; male or female between 38 and 80 years of age; participants will be informed of the research and signed the informed consent form is required for each participant
Exclusion criteria	People who received medical treatment with steroids, physical therapy or acupuncture within the past 4 weeks; participants who had experienced a malignancy of any kind, psychiatric disease or suffered from serious life-threatening disease, such as the heart disease or disease of brain and blood vessels, liver, kidney, and hematopoietic system; participants who complicated with serious genu varus/valgus and flexion contraction or had vascular or nerve injury history in ipsilateral limb; systemic inflammatory disease such as rheumatoid arthritis; patients during pregnancy and lactation period
Recruitment/selection of patients	Conducted at Shu Guang Hospital affiliated with the Shanghai Traditional Medicine university
Age, gender and ethnicity	Age - Mean (SD): 61.5 (8.2). Gender (M:F): 24:53. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grade 2-3 Duration of symptoms: Not stated
Indirectness of population	No indirectness

Interventions	(n=40) Intervention 1: Electroacupuncture. Electroacupuncture using single use, sterile, 30-mm long and 30-gauge acupuncture needles into the local points GB34, ST34, EX-le4, EX-LE5, ST36 and SP9 at the affected lower limb of the knee osteoarthritis patients. De qi sensation was achieved at each point through lifting and thrusting movements combined with twisting and rotating needles. Then, the HANS-200E stimulators were used to stimulate the needles in pairs GB34-ST34, EX-LE4-EX-LE5 and ST36-SP9. The three pairs of six acupoints were simultaneously stimulated. The stimulation was 30 minutes per session. In the high intensity group, the intensity of the stimulation was strong enough to reach the patients' tolerance threshold value (5-6mA). People will receive 16 electroacupuncture treatments: five times per week (once a day for 5 days continuously, followed by a 2 day interval) during the first 2 weeks, and three times a week (once every 2-3 days) during the following 2 weeks. Duration 4 weeks. Concurrent medication/care: All people were encouraged to self-exercise and pay attention to maintaining good posture. Meanwhile, all people received 30mg etoricoxib tablets once a day during the study. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=40) Intervention 2: Sham acupuncture. Electroacupuncture using single use, sterile, 30-mm long and 30-gauge acupuncture needles into the local points GB34, ST34, EX-le4, EX-LE5, ST36 and SP9 at the affected lower limb of the knee osteoarthritis patients. De qi sensation was achieved at each point through lifting and thrusting movements combined with twisting and rotating needles. Then, the HANS-200E stimulators were used to stimulate the needles in pairs GB34-ST34, EX-LE4-EX-LE5 and ST36-SP9. The three pairs of six acupoints were simultaneously stimulated. The stimulation was 30 minutes per session. In the low intensity group, the intensity was relatively weak so the participants could begin to feel the electroacupuncture stimulus plus 1mA (2-2.5mA). People will receive 16 electroacupuncture treatments: five times per week (once a day for 5 days continuously, followed by a 2 day interval) during the first 2 weeks, and three times a week (once every 2-3 days) during the following 2 weeks. Duration 4 weeks. Concurrent medication/care: All people were encouraged to self-exercise and pay attention to maintaining good posture. Meanwhile, all people received 30mg etoricoxib tablets once a day during the study. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (This study was supported in part by National Natural Science Foundation of China (Grant no. 81202767 and 81202748), and Shang-hai municipal Health Bureau (Grant no. 2012QL016A).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 4 weeks; Group 1: mean 13.7 (SD 6.42); n=40, Group 2: mean 16.24 (SD 7.18); n=37; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported electroacupuncture: 13.70 (12.71-16.69). Reported sham: 16.24 (13.93-18.56). Baseline electroacupuncture: 19.15 (17.59-20.71). Baseline sham: 18.84 (16.63-21.05).
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 gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 taking other medical therapies

Protocol outcome 2: Physical function at ≤ 3 months
- Actual outcome: WOMAC physical function at 4 weeks; Group 1: mean 46.18 (SD 11.44); n=40, Group 2: mean 55.76 (SD 19.51); n=37; WOMAC physical function 0-68 Top=High is poor outcome; Comments: Reports final values and 95% confidence intervals. Reported electroacupuncture: 46.18 (42.63-49.72). Reported sham: 55.76 (49.49-62.06). Baseline electroacupuncture: 63.22 (59.05-67.40). Baseline sham: 63.59 (57.16-70.03).
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 gender, age, BMI and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 taking other medical therapies

Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Lam 2021[63]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=86)
Countries and setting	Conducted in Hong Kong (China); Setting: Not reported
Line of therapy	Unclear
Duration of study	Intervention time: 4 week + 6 weeks
Method of assessment of guideline condition	Unclear method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 50 years or above; fell within the American College of Rheumatology clinical classification criteria for osteoarthritis of the knee, had present knee pain, and had less than 30 minutes of morning stiffness or crepitus on active motion and osteophytes, as determined by history and physical examination; had either unilateral knee pain or bilateral knee pain; and knee pain intensity over 40mm on a visual analogue scale; and were able to read and write Chinese and sign the informed consent form
Exclusion criteria	Unable to walk; had a serious infection of the knee; had a history of knee trauma, ligament damage, fracture or surgery in the past 6 months; had a history of prolotherapy, hyaluronic acid injections, or corticosteroid injections within the past 3 months; had received acupuncture, electroacupuncture, tui-na therapy, massage or physiotherapy within the past 8 weeks
Recruitment/selection of patients	Not reported
Age, gender and ethnicity	Age - Mean (SD): Acupuncture group: 62.7 (7.0); Sham group: 63.4 (6.7). Gender (M:F): 24/59. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): Not stated / Unclear 2. Diagnosis: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity of symptoms (VAS, mean, SD): Acupuncture group: 71.2 (16.0); Sham group: 70.1 (19.7) Duration of symptoms (n, %): Acupuncture group: ≤1 year - 1 (2.4), <1 to 5 years - 21 (50), <5 to 10 - 12 (28.6), >10 years - 8 (19.0); Sham acupuncture group: ≤1 year - 2 (4.9), <1 to 5 years - 13 (31.7), <5 to 10 - 18 (43.9), >10 years - 8 (19.5)
Indirectness of population	No indirectness

Interventions	(n=43) Intervention 1: Acupuncture/dry needling - Acupuncture. One style of superficial needling acupuncture was employed in the study and performed by acupuncturists with over 3 years of experience in acupuncture practice. Participants received acupuncture treatment in the sitting position with the knee joint flexed at the most comfortable angle closest to 90 degrees. A hospital trolley table was set up over the knees to prevent the participant from seeing the acupuncture treatment. Acupuncturists examined the painful points and points of tenderness and spasm in muscles along the foot meridians in affected knees and treated 5-8 affected points for each painful knee. The brief treatment procedure was as follows: after the disinfection of the acupoints the acupuncturists punctured the points using sterile disposable needles that were 0.3 mm x 40 mm. The needles were inserted into muscle in a length of 10-20mm at an angle of 0-10 degrees to the skin. The needles were adjusted by extension and flexion of the knee joint to ensure that all the needles would not cause pain during movement. The needles were then covered with hypoallergenic bandages. Participants were then advised to walk for 10 minutes, followed by stepping up and down from an 18cm step for 12 rounds per knee and sitting for 5 minutes. The needles were removed and bandages were applied for continuous blinding. The intervention lasted for 30 minutes. Duration 4 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=43) Intervention 2: Sham acupuncture. Participants in the sham group underwent the same procedures as those in the acupuncture group except that non-insertion sham acupuncture was employed. Briefly, the participants were in the sitting position and blinded by the trolley table. After disinfection of the acupoints, the acupuncturists applied needles (0.30 mm x 40 mm) on acupoints without penetrating the skin, followed by the cover of bandages. The other procedures were identical to those in the acupuncture group. Duration 4 weeks. Concurrent medication/care: Not reported. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (The study was supported by the Chinese Medicine Department, Hospital Authority, Hong Kong, and the Hong Kong Tuberculosis Association Chinese Medicine Clinic come Training Centre of the University of Hong Kong)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE	

Protocol outcome 1: Health-related quality of life at ≤ 3 months - Actual outcome: Quality of life - physical component at 10 weeks; Group 1: mean 4.1 (SD 8.76); n=42, Group 2: mean 3.4 (SD 8.82); n=41; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Study reported 95% confidence intervals so standard deviations have been calculated. Baseline scores: Acupuncture group: 42.1 (14.3); Sham group: 40.2 (15.8) Risk of bias: All domain - --, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment (54.8 versus 34.1 employed), target knees (11.9% versus 31.7% lateral); Group 1 Number missing: 1, Reason: Did not receive allocated intervention (declined to participate); Group 2 Number missing: 2, Reason: Did not receive allocated intervention (declined to participate) - Actual outcome: Quality of life - mental component at 10 weeks; Group 1: mean 1.2 (SD 10.42); n=42, Group 2: mean 3.8 (SD 10.62); n=41; SF-36 - mental component summary 0-100 Top=High is good outcome; Comments: Study reported 95% confidence intervals so standard deviations have been calculated. Baseline scores: Acupuncture group: 42.1 (14.3); Sham group: 40.2 (15.8) Risk of bias: All domain - --, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment (54.8 versus 34.1 employed), target knees (11.9% versus 31.7% lateral); Group 1 Number missing: 1, Reason: Did not receive allocated intervention (declined to participate); Group 2 Number missing: 2, Reason: Did not receive allocated intervention (declined to participate) Protocol outcome 2: Pain at ≤ 3 months - Actual outcome: Pain at 10 weeks; Group 1: mean -55.4 (SD 94.07); n=42, Group 2: mean -52.5 (SD 95.39); n=41; WOMAC - pain Unclear Top=High is poor outcome; Comments: Study reported 95% confidence intervals so standard deviations have been calculated. Baseline values: Acupuncture group 210.5 (88.9); Sham group 221.5 (104.2) Risk of bias: All domain - --, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment (54.8 versus 34.1 employed), target knees (11.9% versus 31.7% lateral); Group 1 Number missing: 1, Reason: Did not receive allocated intervention (declined to participate); Group 2 Number missing: 2, Reason: Did not receive allocated intervention (declined to participate) Protocol outcome 3: Physical function at ≤ 3 months - Actual outcome: Physical function at 10 weeks; Group 1: mean -227.4 (SD 319.74); n=42, Group 2: mean -182.1 (SD 324.08); n=41; WOMAC - physical function Unclear Top=High is poor outcome; Comments: Study reported 95% confidence intervals so standard deviations have been calculated. Baseline scores: Acupuncture group: 774.6 (378.8); Sham group: 785.5 (357.4) Risk of bias: All domain - --, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: Difference in employment (54.8 versus 34.1 employed), target knees (11.9% versus 31.7% lateral); Group 1 Number missing: 1, Reason: Did not receive allocated intervention (declined to participate); Group 2 Number missing: 2, Reason: Did not receive allocated intervention (declined to participate)	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3

	months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Lansdown 2009[64]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=30)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 10 weeks of treatment, 12 months of follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People over 50 years old who had consulted their GP in the last 3 years with knee pain) capturing the clinical symptoms of osteoarthritis of the knee, but no radiographically confirmed diagnosis)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People over 50 years old who had consulted their GP in the last 3 years with knee pain with ongoing pain and stiffness in their knee
Exclusion criteria	Under cancer care review; currently receiving acupuncture; having had a knee or hip replacement; involved in any insurance claim or litigation related to their knee pain; suffering from rheumatoid arthritis or haemophilia
Recruitment/selection of patients	Recruited from a York-based GP practice with a list size of 15,927 patients
Age, gender and ethnicity	Age - Mean (SD): 63.5 (8.2). Gender (M:F): 12:18. Ethnicity: All participants were white
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: Not stated.
Indirectness of population	No indirectness
Interventions	(n=15) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture by a flexible approach. This meant that the number of needles inserted, depth of needle insertion, needle responses elicited, needle stimulation used, needle retention time and needle type varied. Treatments were usually weekly for 10 sessions. Duration 10 weeks. Concurrent medication/care: Both groups received usual care, which included any appointments, medications (prescribed or over the counter) and interventions sought by participants from any health practitioner. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture

	(n=15) Intervention 2: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. No acupuncture. Duration 10 weeks. Concurrent medication/care: Both groups received usual care, which included any appointments, medications (prescribed or over the counter) and interventions sought by participants from any health practitioner. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Equipment / drugs provided by industry (This study was funded in part by the Medical Research Council's Health Services Research Collaboration. The medical Research Council provided an MSC scholarship for Harriet Landsdown. Hugh MacPherson is funded by a National Institute for Health Research Career Scientist Award.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: EQ-5D at 12 weeks; Group 1: mean 0.71 (SD 0.26); n=15, Group 2: mean 0.66 (SD 0.25); n=15; EQ-5D -0.11-1.0 Top=High is good outcome; Comments: Baseline acupuncture: 0.61 (0.25). Baseline sham: 0.67 (0.15).
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, sex, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 failed to return questionnaire

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: EQ-5D at 12 months; Group 1: mean 0.66 (SD 0.24); n=15, Group 2: mean 0.63 (SD 0.19); n=15; EQ-5D -0.11-1.0 Top=High is good outcome; Comments: Baseline acupuncture: 0.61 (0.25). Baseline sham: 0.67 (0.15).
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: Reported age, sex, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 2, Reason: 1 returned questionnaire with no quantitative data, 1 failed to return questionnaire; Group 2 Number missing: 7, Reason: 7 failed to return questionnaire

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 3.6 (SD 2.92); n=15, Group 2: mean 6.57 (SD 4.54); n=15; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 7.33 (2.82). Baseline sham: 7.40 (3.66).
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, sex, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 failed to return questionnaire

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean 4.7 (SD 2.3); n=15, Group 2: mean 5.3 (SD 3.9); n=15; WOMAC pain 0-20 Top=High is poor

outcome; Comments: Baseline acupuncture: 7.33 (2.82). Baseline sham: 7.40 (3.66). 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, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 2, Reason: 1 returned questionnaire with no quantitative data, 1 failed to return questionnaire; Group 2 Number missing: 7, Reason: 7 failed to return questionnaire Protocol outcome 5: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 12 weeks; Group 1: mean 13.4 (SD 12.12); n=15, Group 2: mean 21.86 (SD 11.99); n=15; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 20.53 (12.71). Baseline sham: 26.27 (13.98). 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, sex, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: 1 failed to return questionnaire Protocol outcome 6: Physical function at > 3 months - Actual outcome: WOMAC function at 12 months; Group 1: mean 17.4 (SD 13.9); n=15, Group 2: mean 17.6 (SD 12.6); n=15; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 20.53 (12.71). Baseline sham: 26.27 (13.98). 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, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 2, Reason: 1 returned questionnaire with no quantitative data, 1 failed to return questionnaire; Group 2 Number missing: 7, Reason: 7 failed to return questionnaire Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Major adverse events at 12 months; Group 1: 0/15, Group 2: 0/15 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 age, sex, ethnicity, employment history, education and baseline values of outcomes; Group 1 Number missing: 2, Reason: 1 returned questionnaire with no quantitative data, 1 failed to return questionnaire; Group 2 Number missing: 7, Reason: 7 failed to return questionnaire	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	Lev-ari 2011[67]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=41)
Countries and setting	Conducted in Israel; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks of treatment, 12 weeks in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People diagnosed as having osteoarthritis of the knee of at least 6 months duration with pain
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 45 years or older; diagnosed as having osteoarthritis of the knee of at least 6 months duration; have had moderate to severe pain during most days throughout the past month for which they had used analgesics for at least 1 month; were willing and able to complete the study protocol
Exclusion criteria	Intra-articular corticosteroid injection into the knees within 4 weeks preceding the study and severe unstable chronic illness (e.g. congestive heart failure, chronic renal failure, cancer).
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 71.2 (8.9). Gender (M:F): 17:38. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis without imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=28) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture using selected acupuncture points on the basis of traditional Chinese medicine treatment methods found to be effective for osteoarthritis of the knee (all people were needled on the following: GB34 on the opposite side, SP5, heading, ST35, Xi Yan on the painful side and LI11 or close Ah-shi point on the opposite side). Shu stream point was needled on the meridian involved with the knee pain and a local point around the needle was added according to the treated meridian). The standard intervention entailed the insertion of exposable sterile 0.16mm thick needles. Needles were left in place for a period of 20 minutes and manually manipulated every 5 minutes. Carried

	out twice weekly for 8 weeks. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=27) Intervention 2: Sham acupuncture. Sham acupuncture at the same frequency and according to the same protocol as that used on the intervention group but without insertion of needles into the skin. An empty needle tube was taped to the skin at acupoints to produce sensations similar to those of needle insertion, after which the needles were inserted into a piece of adhesive foam taped to the skin. Carried out twice weekly for 8 weeks. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: KSS pain score at 12 weeks; Group 1: mean 24 (SD 13.2); n=28, Group 2: mean 21.1 (SD 12.7); n=27; KSS pain score 0-50 Top=High is good outcome; Comments: Baseline acupuncture: 16.3 (12.1). Baseline sham: 17.3 (10.0). 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 KSS function score at baseline. Similar for pain score, gender and age.; Group 1 Number missing: 7, Reason: 4 lost during treatment, 3 lost during follow up; Group 2 Number missing: 7, Reason: 6 lost during treatment, 1 lost during follow up Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: KSS function score at 12 weeks; Group 1: mean 67.4 (SD 24.2); n=28, Group 2: mean 54.7 (SD 15); n=27; KSS function score 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 61.1 (20.2). Baseline sham: 48.7 (19.9). 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 KSS function score at baseline. Similar for pain score, gender and age.; Group 1 Number missing: 7, Reason: 4 lost during treatment, 3 lost during follow up; Group 2 Number missing: 7, Reason: 6 lost during treatment, 1 lost during 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; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Lin 2018[73]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks of treatment, 26 weeks of follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical diagnosis of knee osteoarthritis according to the NICE 2014 guideline edit ion with chronic knee pain for the past 6 months and radiologic confirmation of unilateral or bilateral knee osteoarthritis (Kellgren Lawrence score of 2-3)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 45-75 years (either sex); be diagnosed with knee osteoarthritis; chronic knee pain during the past 6 months; radiologic confirmation of unilateral or bilateral knee osteoarthritis; an average pain intensity of 40 or more on a 100mm visual analogue scale in the previous seven days; agree to refrain from the use of any analgesics during the trial
Exclusion criteria	History of knee surgery or arthroscopy; pain in the knee caused by floating cartilage, joint effusion, inflammatory, malignant, or autoimmune disease; serious acute or chronic organic disease or mental disorder; pregnancy or breastfeeding; and history of bleeding disorder; if they had acupuncture treatment or participated in other clinical trials in the past 3 months
Recruitment/selection of patients	Recruited from 3 sites: Beijing Hospital of Traditional Chinese medicine Affiliated to Capital Medical University, Beijing Friendship Hospital and Beijing Jishuitan Hospital
Age, gender and ethnicity	Age - Mean (SD): 59.8 (7.4). Gender (M:F): 5:37. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence score of 2-3 Duration of symptoms: At least 6 months
Indirectness of population	No indirectness
Interventions	(n=21) Intervention 1: Acupuncture/dry needling - Acupuncture. Traditional Chinese acupuncture that was semi-standardized. Used the 10 commonly used local points (ST34, ST35, ST36, EX-LE2, EX-LE5, GB33, GB34, SP9, SP10, LR8) and between

	three and four acupuncture points from 11 distal points (GB31, GB36, GB39, GB41, ST40, ST41, LR3, BL60, SP6, KI3, LI4). Therefore, the minimum number of needles is 8 and the maximum is 10 (unilateral). The acupuncturists inserted 0.30mmx40mm (for local points) and 0.30mmx25mm (for distal points) acupuncture needles to a conventional depth of approximately 10-30mm, depending on point location. People were treated with manipulations of twirling, lifting and thrusting on the basis of traditional Chinese acupuncture theory. People had to achieve "de qi" sensation with needles stimulation manually at least 10 seconds. Duration 8 weeks. Concurrent medication/care: Celebrex was given to people with a pain score greater than and equal to 8/10. People were advised to not have any other treatments. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=21) Intervention 2: Sham acupuncture. Sham acupuncture through minimal insertion into non-acupoints with no manipulation of the needle. Duration 8 weeks. Concurrent medication/care: Celebrex was given to people with a pain score greater than and equal to 8/10. People were advised to not have any other treatments. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (The study was supported by Beijing Municipal Administration of Hospitals Clinical medicine Development of Special Funding Support (XMLX201607) and Beijing Municipal Science & Technology Commission (D171100003217003).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: Physical health (SF-12) at 8 weeks; Group 1: mean 40.7 (SD 9.6); n=21, Group 2: mean 40.2 (SD 10.1); n=21; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 36.6 (9.1). Baseline sham: 37.9 (9.5).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 scheduling conflict; Group 2 Number missing: 1, Reason: 1 lack of treatment effect
- Actual outcome: Mental health (SF-12) at 8 weeks; Group 1: mean 51.5 (SD 11.1); n=21, Group 2: mean 53.2 (SD 10.4); n=21; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 51.3 (11.4). Baseline sham: 52.4 (9.5).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 scheduling conflict; Group 2 Number missing: 1, Reason: 1 lack of treatment effect

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: Physical health (SF-12) at 26 weeks; Group 1: mean 38.2 (SD 9.2); n=21, Group 2: mean 37.6 (SD 9.3); n=21; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 36.6 (9.1). Baseline sham: 37.9 (9.5).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 scheduling conflict, 1 relocation, 1 lack of treatment effect ; Group 2 Number missing: 3, Reason: 1 lack of treatment effect, 1 lost to contact, 1 other
- Actual outcome: Mental health (SF-12) at 26 weeks; Group 1: mean 51.2 (SD 11.1); n=21, Group 2: mean 52.2 (SD 9.8); n=21; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 51.3 (11.4). Baseline sham: 52.4 (9.5).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 scheduling conflict, 1 relocation, 1 lack of treatment effect ; Group 2 Number missing: 3, Reason: 1 lack of treatment effect, 1 lost to contact, 1 other

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean -3.3 (SD 2.5); n=21, Group 2: mean -3.1 (SD 2.6); n=21; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 5.8 (2.8). Baseline sham: 7.2 (3.7).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 scheduling conflict; Group 2 Number missing: 1, Reason: 1 lack of treatment effect

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean 4.1 (SD 3.4); n=21, Group 2: mean 4.6 (SD 3.2); n=21; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 5.8 (2.8). Baseline sham: 7.2 (3.7).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 scheduling conflict, 1 relocation, 1 lack of treatment effect ; Group 2 Number missing: 3, Reason: 1 lack of treatment effect, 1 lost to contact, 1 other

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean -11 (SD 7.3); n=21, Group 2: mean -9.5 (SD 6.6); n=21; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 20.1 (9.8). Baseline sham: 21.6 (10.4).
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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 scheduling conflict; Group 2 Number missing: 1, Reason: 1 lack of treatment effect

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 26 weeks; Group 1: mean 12.1 (SD 8.8); n=21, Group 2: mean 14 (SD 8.1); n=21; WOMAC function 0-68 Top=High is

poor outcome; Comments: Baseline acupuncture: 20.1 (9.8). Baseline sham: 21.6 (10.4). 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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 scheduling conflict, 1 relocation, 1 lack of treatment effect ; Group 2 Number missing: 3, Reason: 1 lack of treatment effect, 1 lost to contact, 1 other Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Serious adverse events at 26 weeks; Group 1: 0/21, Group 2: 0/21; Comments: Acupuncture: 2 adverse events (needling pain after treatment, hematoma). Sham acupuncture: 1 adverse event (needling pain after treatment). But 0 serious adverse events. 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: Reported age, sex, duration of disease, BMI, and baseline values of outcomes; Group 1 Number missing: 3, Reason: 1 scheduling conflict, 1 relocation, 1 lack of treatment effect ; Group 2 Number missing: 3, Reason: 1 lack of treatment effect, 1 lost to contact, 1 other	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	Lv 2019[82]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=301)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical criteria for knee osteoarthritis formulated by the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were 50 years or older meeting the clinical criteria for knee osteoarthritis formulated by the American College of Rheumatology
Exclusion criteria	People who experienced adverse reactions to acupuncture prior to our study; who had comorbidities including severe cardiovascular, cerebral, hepatic, renal, or hematopoietic diseases; who had other disorders that might affect the knee (e.g. rheumatoid arthritis, gouty arthritis); who were pregnant or attempting to become pregnant or were lactating; who had a history of mental illness. All people were required not to take analgesic medications and electroacupuncture 48 hours before each treatment session.
Recruitment/selection of patients	People were recruited from 5 hospitals in Wuhan, China
Age, gender and ethnicity	Age - Mean (SD): 63.7 (9.9). Gender (M:F): 69:223. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: Between <0.5 years and at least 5 years, median 0.5 to 3 years
Indirectness of population	No indirectness
Interventions	(n=225) Intervention 1: Electroacupuncture. Weak and strong electroacupuncture. Electroacupuncture delivered in 10 30-minute sessions over 2 weeks. This was completed with participants in a supine position, with a pillow under each knee for support. Sterile disposable needles (30 gauge with an outer diameter of 0.32mm and a length of 40mm). The same four acupoints were used: Neixiyan (EX-LE5), Dubi (ST35), Liangqiu (ST34), and Xuehai (SP10) unilaterally based on traditional Chinese

	medicine meridian theory. Needles were inserted to a depth of 25 to 40mm vertically. De qi sensation was elicited by lifting and thrusting combined with twirling and rotating the needles. Electrical stimulation was applied using an electroacupuncture apparatus with a pair of electrodes connecting points EX-LE5 with ST35 and another pair connecting ST34 and SP10. Stimulation parameters were direct current, continuous wave, 2Hz frequency and 0.5ms pulse width for 30 minutes. The strong group received a current between 2 and 5mA, while the low-intensity group received a current of 0-0.5mA.. Duration 2 weeks. Concurrent medication/care: All people were required not to take analgesic medications and electroacupuncture 48 hours before each treatment session. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture Comments: The strong and weak electroacupuncture groups were combined due to class effect (n=76) Intervention 2: Sham acupuncture. Sham electroacupuncture where the number of acupoints, electroacupuncture apparatus and stimulation parameters were the same as for the true electroacupuncture groups. However, the needles used in the sham group were fine and short (35-gauge needle with an outer diameter of 0.20mm and a length of 25mm). The needles were inserted only superficially into non-acupoint sites, each 2cm lateral to each of the four acupoints to an approximate depth of 5 to 10mm. In addition, the needles were not manipulated to avoid obtaining de qi sensation. Electrical stimulation was delivered with the same low intensity as the weak electroacupuncture group.. Duration 2 weeks. Concurrent medication/care: All people were required not to take analgesic medications and electroacupuncture 48 hours before each treatment session. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (The trial was sponsored by a grant from the National Natural Science Foundation of China (No. 81473768) and the Fundamental Research Funds for the Central Public Welfare Research Institutes (No. ZZKF08007))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: VAS at 2 weeks; Group 1: mean -2.9 (SD 1.23); n=217, Group 2: mean 1.19 (SD 1.21); n=75; VAS 0-10 Top=High is poor outcome; Comments: Stated that it reports change scores and standard deviation. However, the size of the standard deviations is too small and correlated more with standard errors. Therefore, they have been treated as standard errors. Reported strong electroacupuncture: -2.97 (0.10). Reported weak electroacupuncture: -2.75 (0.15). Reported sham: 1.19 (0.14).
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: Reported gender, age, symptom duration, height, weight, BMI and previous treatment; Group 1 Number missing: 17, Reason: Strong electroacupuncture: 9 dropped out (5 reason unclear, 2 coronary heart disease, 1 stroke, 1 fracture). Weak electroacupuncture: 8 dropped out (3 reasons unclear, 2 coronary heart disease, 1 pulmonary embolism, 1 pulmonary infection, 1 nephritis).; Group 2 Number missing: 4, Reason: 4 dropped out (2 reason unclear, 2 coronary heart disease). Protocol outcome 2: Serious adverse events at ≤ 3 months - Actual outcome: Adverse events at 2 weeks; Group 1: 32/217, Group 2: 11/75; Comments: Strong electroacupuncture: 22 (15 subcutaneous haemorrhage or bleeding, 7 needling pain and nausea). Weak electroacupuncture: 10 (7 subcutaneous haemorrhage and 3 needling pain or nausea). Sham electroacupuncture: 11 (7 subcutaneous haemorrhage or bleeding, 4 needling pain). 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: Reported gender, age, symptom duration, height, weight, BMI and previous treatment; Group 1 Number missing: 17, Reason: Strong electroacupuncture: 9 dropped out (5 reason unclear, 2 coronary heart disease, 1 stroke, 1 fracture). Weak electroacupuncture: 8 dropped out (3 reasons unclear, 2 coronary heart disease, 1 pulmonary embolism, 1 pulmonary infection, 1 nephritis).; Group 2 Number missing: 4, Reason: 4 dropped out (2 reason unclear, 2 coronary heart disease).	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study	Mavrommatis 2012[92]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=120)
Countries and setting	Conducted in Greece; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks of treatment, 12 weeks follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People had to have met the American College of Rheumatology criteria for diagnosis of knee osteoarthritis with Kellgren Lawrence scores of at least 2 and chronic pain in the knee joint for more than 3 months
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People who met the American College of Rheumatology criteria for diagnosis of knee osteoarthritis; people with a Kellgren-Lawrence score of at least 2; chronic pain in the knee joint for more than 3 months
Exclusion criteria	Intra-articular corticosteroids or hyaluronate injection during the previous 3 months; corticosteroids per os; antiplatelet drugs (apart from acetylsalicylic acid 100mg); immunosuppressive drugs; pregnancy; people who had experienced a malignancy of any kind; psychiatric disease; stroke; heart attack; kidney failure; active gastric or duodenal ulcer; gastrorrhagia; previously received acupuncture; other forms of arthritis; arthroplasty during the previous year
Recruitment/selection of patients	People were recruited from the Orthopaedic Clinic of the General Hospital of Florina
Age, gender and ethnicity	Age - Mean (SD): 61.8 (10.6). Gender (M:F): 29:91. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grade of at least 2 Duration of symptoms: At least 3 months
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Electroacupuncture. Acupuncture using single use, sterile, 30mm long and 30-gauge acupuncture needles into the local points ST36, SP9, SP10, GB34, Ex-LE 2, and Ex-LE5 as well as the distal points Li4, Ki3, ST40, and SP6. At each point, the person confirmed the de qi sensation. The treatment was given biweekly for 8 weeks. Starting from the third session, the ES-160 electrostimulator ITO

	cb. (6Hz, 150ms for 20 minutes) was used to stimulate the needles in pairs ST36-SP9 and GB34-SP10. Duration 8 weeks. Concurrent medication/care: Everyone was treated with etoricoxib 60mg only on a daily basis and were examined biweekly. People with risk factors for upper gastrointestinal tract complications received proton pump inhibitors. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=40) Intervention 2: Sham acupuncture. Sham acupuncture for the same duration and frequency, but with retractable needles in small adhesive cylingers, placed in the same points. The same pairs of electrodes were used to simulate the electrical connection. Duration 8 weeks. Concurrent medication/care: Everyone was treated with etoricoxib 60mg only on a daily basis and were examined biweekly. People with risk factors for upper gastrointestinal tract complications received proton pump inhibitors. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=40) Intervention 3: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. No acupuncture. Duration 8 weeks. Concurrent medication/care: Everyone was treated with etoricoxib 60mg only on a daily basis and were examined biweekly. People with risk factors for upper gastrointestinal tract complications received proton pump inhibitors. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical component summary at 8 weeks; Group 1: mean 45.8 (SD 6.9); n=40, Group 2: mean 35.2 (SD 5.4); n=40; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 29.3 (5.2). Baseline sham: 28.2 (6.8).
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 gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 0
- Actual outcome: SF-36 mental component summary at 8 weeks; Group 1: mean 52.2 (SD 8); n=40, Group 2: mean 51.5 (SD 6.1); n=40; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 45.9 (8.5). Baseline sham: 42.8 (9.8).
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: Reported gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 0

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 77.9 (SD 32.9); n=40, Group 2: mean 145.9 (SD 35.5); n=40; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline acupuncture: 221.5 (44.2). Baseline sham: 212.2 (44.4).
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 gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 1, Reason: 1 lost to follow up

Protocol outcome 3: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 321.6 (SD 141); n=40, Group 2: mean 603.2 (SD 126); n=40; WOMAC function 0-1800 Top=High is poor outcome; Comments: Baseline acupuncture: 948.0 (203.0). Baseline sham: 890.3 (158.0).
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 gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 1, Reason: 1 lost to follow up

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical component summary at 8 weeks; Group 1: mean 45.8 (SD 6.9); n=40, Group 2: mean 35.3 (SD 4.5); n=40; SF-36 physical component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 29.3 (5.2). Baseline no treatment: 28.0 (5.7).
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: Reported gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 2, Reason: 2 ceased due to increase in arterial blood pressure
- Actual outcome: SF-36 mental component summary at 8 weeks; Group 1: mean 52.2 (SD 8); n=40, Group 2: mean 50.7 (SD 7.4); n=40; SF-36 mental component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 45.9 (8.5). Baseline no treatment: 45.2 (8.0).
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 gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 2, Reason: 2 ceased due to increase in arterial blood pressure

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 77.9 (SD 32.9); n=40, Group 2: mean 153.8 (SD 41.2); n=40; WOMAC pain 0-500 Top=High is poor outcome; Comments: Baseline acupuncture: 221.5 (44.2). Baseline no treatment: 198.9 (45.6).
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, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 2, Reason: 2 ceased

due to increase in arterial blood pressure Protocol outcome 3: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 12 weeks; Group 1: mean 321.6 (SD 141); n=40, Group 2: mean 653.4 (SD 157); n=40; WOMAC function 0-1800 Top=High is poor outcome; Comments: Baseline acupuncture: 948.0 (203.0). Baseline no treatment: 860.0 (165.0). 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 gender, age, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 ceased due to start of treatment with clopidogrel; Group 2 Number missing: 2, Reason: 2 ceased due to increase in arterial blood pressure	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Miller 2011[95]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=55)
Countries and setting	Conducted in Israel; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks of intervention, additional 1 month of follow up (12 weeks in total)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosed as having osteoarthritis of the knee at least of 6 months duration with moderate to severe pain during most days throughout the past month
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	45 years or older; diagnosed as having osteoarthritis of the knee of at least 6 months duration; had been suffering from moderate to severe pain during most days throughout the past month for which they had used analgesics for at least 1 month; were willing and able to complete the study protocol.
Exclusion criteria	Intra-articular corticosteroid injection into the knee within 4 weeks preceding the study; severe unstable chronic illness (e.g., congestive heart failure, chronic renal failure, cancer).
Recruitment/selection of patients	People were recruited from July 2002 to October 2003 at the Department of Orthopaedics "B" of the Tel Aviv Sourasky Medical Center, a large university-affiliated institution.
Age, gender and ethnicity	Age - Mean (SD): 71.2 (8.9). Gender (M:F): 17:38. Ethnicity: Not stated/unclear
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated/unclear Duration of symptoms: At least 6 months
Indirectness of population	No indirectness
Interventions	(n=28) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture following Traditional Chinese Medicine treatment methods using the following acupuncture points: GB34 on the opposite side, SP5, Heading, ST35, Xi Yan on the painful side and LI11 or close Ah-shi point on the opposite side, the Shu Stream point and a local point around the knee was added according to the treated meridian (ST34 to treat pain

	on the stomach meridian, KI10 to treat pain on the kidney meridian). The standard acupuncture intervention entailed the insertion of exposable sterile 0.16mm thick needles. Acupuncture was performed after alcohol wipe of the skin at the specific points. Needles were left in place for a period of 20 minutes and manually manipulated every 5 minutes. Treatment was twice weekly for 8 weeks.. Duration 8 weeks. Concurrent medication/care: All people received standard therapy, which included treatment with NSAIDs.. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=27) Intervention 2: Sham acupuncture. Sham acupuncture performed at the same frequency and according to the same protocol as that used for the intervention group but without insertion of needles into the skin. An empty needle tube was taped to the skin at acupoints to produce sensations similar to those of needle insertion, after which the needles were inserted into a piece of adhesive foam taped to the skin. Carried out twice weekly for 8 weeks.. Duration 8 weeks. Concurrent medication/care: All people received standard therapy, which included treatment with NSAIDs.. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: KSS pain score at 12 weeks; Group 1: mean 24 (SD 13.2); n=21, Group 2: mean 21.1 (SD 12.7); n=20; KSS pain score Unclear Top=High is good outcome; Comments: Baseline acupuncture: 16.3 (12.1). Baseline sham: 17.3 (10.0).
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 the KSS function score at baseline which is likely to have an important effect; Group 1 Number missing: 7, Reason: 4 lost during treatment, 3 lost during follow up; Group 2 Number missing: 7, Reason: 6 lost during treatment, 1 lost during follow up

Protocol outcome 2: Physical function at ≤ 3 months
- Actual outcome: KSS function score at 12 weeks; Group 1: mean 67.4 (SD 24.2); n=21, Group 2: mean 54.7 (SD 15); n=20; Comments: Baseline acupuncture: 61.1 (20.2). Baseline control: 48.7 (19.9).
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 the KSS function score at baseline which is likely to have an important effect; Group 1 Number missing: 7, Reason: 4 lost during treatment, 3 lost during follow up; Group 2 Number missing: 7, Reason: 6 lost during treatment, 1 lost during 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; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Min 2009[96]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=78)
Countries and setting	Conducted in South Korea; 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: Osteoarthritis of the knee according to the American College of Rheumatology with documented radiographic changes of osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Female or male over the age of 50; diagnosis of osteoarthritis of the knee; duration of osteoarthritis more than 6 months; documented radiographic changes of osteoarthritis (Kellgren-Lawrence grade of 1 or more); signed informed consent
Exclusion criteria	Female or male over the age of 70; intra-articular corticosteroid injection in the knee within four weeks immediately preceding entry into the study; severe chronic or uncontrolled concomitant illness; diagnosis of rheumatoid arthritis of the knee; history or clinical indications of bleeding diathesis and cardiovascular disease, including current use of anticoagulants; allergy to metal; previous treatment with acupuncture within four weeks prior to entry into the study; taking hormone medications
Recruitment/selection of patients	People were recruited by advertisements
Age, gender and ethnicity	Age - Mean (SD): 59.4 (5.3). Gender (M:F): 14:64. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grades 1-4, median grade 2 Duration of symptoms: At least 6 months.
Indirectness of population	No indirectness
Interventions	(n=40) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture using disposable stainless steel needles (0.25 x 40mm). The depth of Sa-am acupuncture was about 1mm-5mm, and the duration was about 20 minutes with acupuncture two times per week over 4 weeks. Twirling reinforcement reduction and nine six reinforcement reduction methods were used as a reinforcement reduction method, and the people reported feeling deqi. The acupoints are determined by a method

	passed down through generations by Sa-am practitioners and were appointed by "Five Phases". Sa-am uses a set of four acupoints (two for tonifying, two for purging).. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=38) Intervention 2: Sham acupuncture. Sham acupuncture using park sham needles but otherwise the same position, frequency and duration. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (This work was supported by a Korea Research Foundation Grant funded by the Korean Government (mOETHD) (KRS-2005-005-J00702))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical health at 4 weeks; Group 1: mean 8.5 (SD 13.1); n=40, Group 2: mean 5 (SD 13.3); n=38; SF-36 physical health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 33.4 (18.1). Baseline sham: 36.7 (17.1).
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, smoking, drinking, medication use, Kellgren Lawrence grade, sasang constitution, target knee, acupoints, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 3 withdrew consent, 1 personal reason, 2 other disease; Group 2 Number missing: 7, Reason: 3 withdrew consent, 4 personal reason
- Actual outcome: SF-36 mental health at 4 weeks; Group 1: mean 7 (SD 19.7); n=40, Group 2: mean 6.4 (SD 12.8); n=38; SF-36 mental health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 44.7 (20.8). Baseline sham: 44.8 (18.4).
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, gender, weight, height, BMI, smoking, drinking, medication use, Kellgren Lawrence grade, sasang constitution, target knee, acupoints, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 3 withdrew consent, 1 personal reason, 2 other disease; Group 2 Number missing: 7, Reason: 3 withdrew consent, 4 personal reason

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: KWOMAC pain at 4 weeks; Group 1: mean -2 (SD 4.3); n=40, Group 2: mean -0.9 (SD 3); n=38; KWOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 10.4 (4.6). Baseline sham: 8.6 (4.4).
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, gender, weight, height, BMI, smoking, drinking, medication use, Kellgren Lawrence grade, sasang constitution, target knee, acupoints, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 3 withdrew consent, 1 personal reason, 2 other disease; Group 2 Number missing: 7, Reason: 3 withdrew consent, 4 personal reason

Protocol outcome 3: Physical function at ≤ 3 months - Actual outcome: KWOMAC function at 4 weeks; Group 1: mean -6.8 (SD 12.7); n=40, Group 2: mean -1.4 (SD 10.9); n=38; KWOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 38.1 (14.7). Baseline sham: 31.7 (15.6). 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, smoking, drinking, medication use, Kellgren Lawrence grade, sasang constitution, target knee, acupoints, and baseline values of outcomes; Group 1 Number missing: 6, Reason: 3 withdrew consent, 1 personal reason, 2 other disease; Group 2 Number missing: 7, Reason: 3 withdrew consent, 4 personal reason	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Sanchez romero 2020[111]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=62)
Countries and setting	Conducted in Spain; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks of treatment, 12 months follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnostics
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 62 years or older with knee pain and uni- or bilateral dysfunction, primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnostics, and at least one active or one latent myofascial trigger points elicited by palpation ipsilateral to the painful knee that was situated in a taut band of a skeletal muscle of the lower limbs that usually leads to pain
Exclusion criteria	Myofascial or neuropathic pain in the lower limb, such as lumbar radioculopathy, saphenous nerve entrapment, or parestheticameralgia; previous total replacement of the same knee; previous simultaneous total replacements of both knees; any other surgical procedure of the lower limbs in the previous 6 months; prior diagnoses or prescriptions in the medical record for myopathy or lumbo-sacral neuropathy; rheumatoid arthritis, with initiation of opioid analgesia or corticosteroid or analgesic injection intervention for hip or knee pain within the previous 30 days; alcohol or drug consumption; uncontrolled hypertension or moderate to high risk for cardiac complications during exercise; conservative or invasive physical therapy (previous 6 months or during follow-up); fibromyalgia syndrome or other altered affective/cognitive modulation processes of pain perception; physical impairments unrelated to the hip or knee that prevented safe participation in exercise and walking, such as vision problems that affect mobility, body weight >155kg, neurogenic disorder, primary or significantly limiting back pain, advanced osteoporosis, inability to walk 10 meters without an assistive device, 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; a mean pain intensity score higher than 7 on the NRS
Recruitment/selection of patients	Enrolled from older adult centers

Age, gender and ethnicity	Age - Mean (SD): 72.3 (5.7). Gender (M:F): 44:18. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): Mixed 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 65.6 (36.0) months
Indirectness of population	Serious indirectness: People had to have osteoarthritis and myofascial trigger points
Interventions	(n=31) Intervention 1: Acupuncture/dry needling - Acupuncture. Dry needling with 6 sessions for 6 weeks. This used a fast-in fast-out technique, consisting of 15 times of manipulating the needle upwards and downwards inside the muscle. A 0.3x40mm, 0.4x60mm, or 0.3x0.75mm needle inserted perpendicularly into the myofascial trigger point located under the index and middle fingers of the nondominant hand. After removing the guide tube, the area was traversed in different directions using the metacarpophalangeal flexion/extension of the first and second fingers of the dominant hand, trying to obtain one or several local twitch responses, a local pain response, and generally the referred pain pattern of myofascial trigger points. Ischaemic compression was applied manually for one minute after removing the needle from the dominant hand. Duration 6 weeks of acupuncture. Concurrent medication/care: Therapeutic exercise for 1 hour, twice a week for 12 weeks. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=31) Intervention 2: Sham acupuncture. Sham acupuncture in the same areas but with a simulated device that does not penetrate the skin. Duration 6 weeks of acupuncture. Concurrent medication/care: Therapeutic exercise for 1 hour, twice a week for 12 weeks. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Academic or government funding (This study was funded by the winnings of the research prize for the Colegio Profesional de Fisioterapeutas de la Comunidad de Madrid)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: EQ-5D at 12 weeks; Group 1: mean 6 (SD 1.6); n=31, Group 2: mean 5.87 (SD 1.12); n=31; EQ-5D 5-15 Top=High is good outcome; Comments: Baseline acupuncture: 7.84 (1.64). Baseline sham: 7.16 (1.59).
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, duration of pain, weight, height, BMI,

and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: EQ-5D at 12 months; Group 1: mean 6.35 (SD 1.56); n=31, Group 2: mean 6.2 (SD 1.36); n=31; EQ-5D 5-15 Top=High is good outcome; Comments: Baseline acupuncture: 7.84 (1.64). Baseline sham: 7.16 (1.59).
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, duration of pain, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 2.81 (SD 2.48); n=31, Group 2: mean 3.68 (SD 3.12); n=31; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 7.58 (2.23). Baseline sham: 8.03 (2.93).
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, duration of pain, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 12 months; Group 1: mean 4.23 (SD 2.56); n=31, Group 2: mean 4.03 (SD 4.25); n=31; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 7.58 (2.23). Baseline sham: 8.03 (2.93).
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, duration of pain, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 9.74 (SD 6.96); n=31, Group 2: mean 10.03 (SD 8.4); n=31; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 26.6 (8.27). Baseline sham: 23.81 (12.21).
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, duration of pain, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 12 months; Group 1: mean 11.71 (SD 7.71); n=31, Group 2: mean 12.1 (SD 10.25); n=31; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 26.6 (8.27). Baseline sham: 23.81 (12.21).

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, duration of pain, weight, height, BMI, and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 underwent total knee replacement before the 12 months follow-up; Group 2 Number missing: 2, Reason: 2 lost to follow up (defunction or had personal or health issues)	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Sanchez-romero 2018[110]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=20)
Countries and setting	Conducted in Spain; 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 and uni- or bilateral dysfunction with primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People aged 65 years or older with knee pain and uni- or bilateral dysfunction, primary knee osteoarthritis fulfilling the American College of Rheumatology criteria for clinical and radiographic diagnosis, and at least 1 active or 1 latent myofascial trigger point elicited by palpation ipsilateral to the painful knee(s), situated in a taut band of a skeletal muscle of the lower limb(s), which usually has referred pain
Exclusion criteria	any other condition that could cause myofascial or neuropathic pain in the lower limbl; previous total replacement of the same knee; any other surgical procedure of the lower limbs in the previous 6 months; prior diagnoses or prescriptions in the medical record for myopathy or lumbosacral neuropathy; rheumatoid arthritis; initiation of opioid analgesia or corticosteroid or analgesic injection intervention for hip or knee pain within the previous 30 days; alcohol or drug consumption; uncontrolled hypertension or moderate to high risk for cardiac complications during exercise; conservative or invasive physical therapy (previous 6 months for during follow-up); or physical impairments unrelated to the hip or knee preventing safe participation in exercise and walking, such as vision problems that affect mobility; body weight grater than 155kg; neurogenic disorder; primary or significantly limiting back pain; advanced osteoporosis; inability to walk 10m without an assistive device; inability to comprehend and complete study assessments or comply with study instructions; stated inability to attend or complete the proposed course of intervention and follow-up schedule; fibromyalgia syndrome; other altered affected/cognitive modulation processes of pain perception
Recruitment/selection of patients	Recruited from older adult care centers
Age, gender and ethnicity	Age - Mean (SD): 71.4 (4.2). Gender (M:F): 8:12. Ethnicity: Not stated

Further population details	1. Age (≤/> 75 years): ≤ 75 years (Confidence intervals fall just under this). 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of pain (mean [SD]): 28.5 (25.2) months
Indirectness of population	Serious indirectness: Required to have myofascial trigger points
Interventions	(n=11) Intervention 1: Acupuncture/dry needling - Dry needling. 6 dry needling sessions (once a week for the first 6 weeks) at all myofascial trigger points of the involved symptomatic lower limb using the fast-in and fast-out technique with multiple rapid needle insertions. Needle insertion was repeated 15 times. A headless 0.30 x 40mm needle, 0.30 x 60mm needle, and 0.30x0.75mm needle (AGU-PUNT) was inserted perpendicularly directly to the selected muscle in the lower limb toward the myofascial trigger point located between the fingers of the subdominant hand, and the guide tube was removed. The area was probed in different directions until a minimum of 1 local twitch response, a local pain response, and usually the referred pain pattern was obtained. Duration 6 weeks of acupuncture, 12 weeks of exercise. Concurrent medication/care: All people received a therapeutic exercise program in 1 hour, group based, supervised sessions twice weekly over 12 weeks. On average, about 10 people attended each training session. A total of 24 sessions were conducted consisting of aerobic exercise (20-25 minutes warm up), lower limb muscle strengthening (20-25 minutes), and lower-limb muscle stretching (10-15 minutes). Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=9) Intervention 2: Sham acupuncture. Sham dry needling once a week for the first 6 weeks using a park sham device. The sham looked like the real needle, except it penetrated only a few millimeters of the skin without inducing any LTR. Duration 6 weeks of acupuncture, 12 weeks of exercise. Concurrent medication/care: All people received a therapeutic exercise program in 1 hour, group based, supervised sessions twice weekly over 12 weeks. On average, about 10 people attended each training session. A total of 24 sessions were conducted consisting of aerobic exercise (20-25 minutes warm up), lower limb muscle strengthening (20-25 minutes), and lower-limb muscle stretching (10-15 minutes). Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Academic or government funding (Xi Award for Best Research Project awarded by the Ilustre Colegio Profesional de Fisioterapeutas de la Comunidad de Madrid (Spain))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: DRY NEEDLING versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 2.82 (SD 1.99); n=11, Group 2: mean 3.33 (SD 2.12); n=9; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 6.82 (2.31). Baseline sham: 7.78 (2.10).
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; Indirectness of outcome: No indirectness ; Baseline details: Reported age, sex, duration of pain, weight, height, BMI and baseline values of outcomes. Different for WOMAC pain and function at baseline.; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 10.27 (SD 7.07); n=11, Group 2: mean 7.78 (SD 6.59); n=9; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 26.73 (10.33). Baseline sham: 22.33 (8.01).
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, duration of pain, weight, height, BMI and baseline values of outcomes. Different for WOMAC pain and function 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; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Sangdee 2002[112]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=193)
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: Unilateral or bilateral osteoarthritis of the knee according to the criteria of the American College of Rheumatology for more than 3 months duration
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People of any sex, aged over 40 years, and who had been suffering from unilateral or bilateral osteoarthritis of the knee according to the criteria of the American College of Rheumatology for more than 3 months duration with a Lequesne's functional index of at least 6 points who were able to walk and give verbal and written consent
Exclusion criteria	An underlying inflammatory arthropathy; expectation of surgery in the future; recent injury in the area affected by osteoarthritis of the knee; intraarticular corticosteroid injections or electroacupuncture within the last 3 months; hypersensitivity to NSAIDs or paracetamol; abnormal liver or kidney function tests; evidence of leukopenia and coagulopathies screened by clinical laboratory; concomitantly receiving anticoagulants; history of peptic ulceration; anaemia; uncontrolled hypertension; congestive heart failure; hyperkalaemia; pregnancy; lactation; malignant tumours
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 62.9 (7.2). Gender (M:F): 43:150. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 4.9 (3.9) years
Indirectness of population	No indirectness
Interventions	(n=97) Intervention 1: Electroacupuncture. Electroacupuncture using four fine stainless steel needles inserted into acupuncture points around the affected knee. All needles were used in order to conduct an electrical current through the points and were inserted superficially (not more than 0.5 inch approximately in depth). Thus, an

	elicitation of needle sensation (de qi) during the insertion was not intended. The first pain of electrodes was connected to the Dubi and nearest adjacent point (medial Xiyan) and the second pair to the trigger point and Qu-quan. The electrical stimulation was applied slowly and simultaneously to each pair of needles until it reached the maximum toleration level of the person. Biphasic pulses were used for the electrical stimulation at a frequency of 2Hz and it was administered for 20 minutes in each treatment. The people were treated 3 times a week for 4 weeks (12 times). People were also given either a) diclofenac 25mg three times a day for 4 weeks or b) placebo three times a day for 4 weeks. Duration 4 weeks. Concurrent medication/care: All additional therapies (e.g. oral or topical NSAIDs, intraarticular corticosteroid injection, other analgesics, chondro-protective agents, surgical procedures on the knee joint etc.) were not allowed. However, all other treatments for concomitant disorders that did not interfere with the study could be continued, but it had to be documented. Paracetamol was prescribed as rescue analgesic. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: The two groups were combined as both electroacupuncture and diclofenac compared to sham acupuncture and diclofenac, and electroacupuncture and placebo compared to sham acupuncture and procedure are the same comparison (n=95) Intervention 2: Sham acupuncture. Sham acupuncture by attached electrodes, but the electrodes were connected to a sound producing dummy mode that did not give a current. The timing and frequency of the treatment was the same. People were also given either a) diclofenac 25mg three times a day for 4 weeks or b) placebo three times a day for 4 weeks. Duration 4 weeks. Concurrent medication/care: All additional therapies (e.g. oral or topical NSAIDs, intraarticular corticosteroid injection, other analgesics, chondro-protective agents, surgical procedures on the knee joint etc.) were not allowed. However, all other treatments for concomitant disorders that did not interfere with the study could be continued, but it had to be documented. Paracetamol was prescribed as rescue analgesic. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: The two groups were combined as both electroacupuncture and diclofenac compared to sham acupuncture and diclofenac, and electroacupuncture and placebo compared to sham acupuncture and procedure are the same comparison
Funding	Academic or government funding (This work was supported by the Faculty of Medicine, Chiang Mai university, Thailand)
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE	

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 4 weeks; Group 1: mean -5.97 (SD 4.66); n=97, Group 2: mean -4.07 (SD 4.3); n=94; WOMAC pain 0-20 Top=High is poor outcome; Comments: Reports change scores and standard error. Reported combined electroacupuncture and diclofenac: -6.28 (0.77). Reported electroacupuncture and placebo: -5.65 (0.59). Reported sham acupuncture and diclofenac: -4.90 (0.53). Reported sham acupuncture and placebo: -3.31 (0.68). Baseline combined electroacupuncture and diclofenac: 10.50 (4.18). Baseline electroacupuncture and placebo: 10.25 (3.86). Baseline sham acupuncture and diclofenac: 11.02 (4.15). Baseline sham acupuncture and placebo: 10.19 (4.20).
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, weight, height, duration of symptoms, localisation, knee affected, number of paracetamol tablets taken per week and baseline values of outcomes; Group 1 Number missing: 5, Reason: 1 flare of pain with joint swelling, 3 severe GI side effects, 1 flare of pain from an accidental fall not related to treatment; Group 2 Number missing: 2, Reason: 2 flare of pain with joint swelling (but only 1 missing from data)

Protocol outcome 2: Physical function at ≤ 3 months
- Actual outcome: WOMAC disability at 4 weeks; Group 1: mean -19.08 (SD 13.47); n=92, Group 2: mean -13.4 (SD 12.54); n=94; WOMAC disability 0-68 Top=High is poor outcome; Comments: Reports change scores and standard error. Reported combined electroacupuncture and diclofenac: -18.98 (1.92). Reported electroacupuncture and placebo: -19.17 (2.05). Reported sham acupuncture and diclofenac: -14.39 (1.77). Reported sham acupuncture and placebo: -12.33 (1.88). Baseline combined electroacupuncture and diclofenac: 37.94 (13.02). Baseline electroacupuncture and placebo: 38.00 (13.18). Baseline sham acupuncture and diclofenac: 35.65 (12.89). Baseline sham acupuncture and placebo: 37.04 (12.00).
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, duration of symptoms, localisation, knee affected, number of paracetamol tablets taken per week and baseline values of outcomes; Group 1 Number missing: 5, Reason: 1 flare of pain with joint swelling, 3 severe GI side effects, 1 flare of pain from an accidental fall not related to treatment; Group 2 Number missing: 2, Reason: 2 flare of pain with joint swelling (but only 1 missing from data)

Protocol outcome 3: Osteoarthritis flares at ≤ 3 months
- Actual outcome: Flare of pain at 4 weeks; Group 1: 2/97, Group 2: 2/95; Comments: Acupuncture: 1 flare of pain with joint swelling, 1 flare of pain from an accidental fall not related to the treatment. Placebo: 2 flare of pain with joint swelling
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: Serious indirectness, Comments: Withdrawal due to adverse events; Baseline details: Reported age, weight, height, duration of symptoms, localisation, knee affected, number of paracetamol tablets taken per week and baseline values of outcomes; Group 1 Number missing: 5, Reason: 1 flare of pain with joint swelling, 3 severe GI side effects, 1 flare of pain from an accidental fall not related to treatment; Group 2 Number missing: 2, Reason: 2 flare of pain with joint swelling (but only 1 missing from data)

Protocol outcome 4: Serious adverse events at ≤ 3 months
- Actual outcome: Severe GI side effects leaving to withdrawal from the trial at 4 weeks; Group 1: 3/97, Group 2: 0/95; Comments: Acupuncture: 3 severe GI adverse events
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: Serious indirectness, Comments: Withdrawal due to adverse events; Baseline details:

Reported age, weight, height, duration of symptoms, localisation, knee affected, number of paracetamol tablets taken per week and baseline values of outcomes; Group 1 Number missing: 5, Reason: 1 flare of pain with joint swelling, 3 severe GI side effects, 1 flare of pain from an accidental fall not related to treatment; Group 2 Number missing: 2, Reason: 2 flare of pain with joint swelling (but only 1 missing from data)	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study	Scharf 2006[113]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=1039)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks of intervention, 26 week follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Chronic pain in the knee joint for the last 6 months, according to the American College of Rheumatology criteria with radiologic confirmation of osteoarthritis in 1 or both knees (Kellgren Lawrence score 2-3)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 40 years and older; chronic pain in the knee joint for the last 6 months; radiologic confirmation of osteoarthritis in 1 or both knees (Kellgren-Lawrence score 2 or 3); WOMAC score of at least 3 points; a chronic pain score of at least 1, according to the criteria of von Korff and colleagues
Exclusion criteria	People with other diseases affecting the knee; neurologic and psychiatric diseases; severe coagulopathy; pregnancy; previous acupuncture treatment for osteoarthritis of the knee
Recruitment/selection of patients	Primary care practices selected from a group of experienced practitioners participating in a large cohort study on acupuncture
Age, gender and ethnicity	Age - Mean (SD): 62.81 (10.04). Gender (M:F): 346:693. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren-Lawrence score 0-4, median score 2 Duration of symptoms (mean [SD]): 65.08 (71.07) months
Indirectness of population	No indirectness
Interventions	(n=330) Intervention 1: Acupuncture/dry needling - Acupuncture. 10 acupuncture sessions administered over a 6 week period beginning 2 weeks after screening. Following traditional Chinese theory of the Bi syndrome to treat knee pain (ST34, ST36, Xiyan, SP9, SP10, GB34). In addition, 2 of 16 defined distal acupuncture points could be chosen, with a maximum of 4 Ahshi points allowed. Duration 6 weeks. Concurrent medication/care: All people received conservative therapy of 150mg of diclofenac per day during the first 2 treatment weeks up to a total of 1g until week 23.

	Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=367) Intervention 2: Sham acupuncture. Sham acupuncture using minimal-depth needling without stimulation at 10 points at defined distances from traditional Chinese acupuncture points. One point was between the gallbladder and stomach meridian on the distal part of the fibula, 2 cun above the malleolus lateralis toward the knee. Two points were 2 cun and 6 cun, respectively, above the malleolus medialis in the center of the tibia surface area, intracutaneous, without periosteum contact and in the direction of the knee. one point was in the center of the thigh on the connecting line from the center of the patella to the anterior superior iliac spine, in the direction of the hip. One point was on the highest spot of the tightened musculus biceps brachii. Duration 6 weeks. Concurrent medication/care: All people received conservative therapy of 150mg of diclofenac per day during the first 2 treatment weeks up to a total of 1g until week 23. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=342) Intervention 3: Other. Standard care including diclofenac, rofecoxib and physician visits (10 visits). Duration 26 weeks. Concurrent medication/care: All people received conservative therapy of 150mg of diclofenac per day during the first 2 treatment weeks up to a total of 1g until week 23. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: This group was not included as a no treatment comparison as the group included additional components not available to the acupuncture groups and so did not fulfill the inclusion criteria
Funding	Academic or government funding (Grants received: C. maier, H.-J. Trampisch, N. Victor (Consortium of Allgemeine Ortskrankenkassen, Betriebskrankenkassen, Innungskrankenkassen, Bundesknappschaft, Landwirtschaftliche, Krankenkassen, and See-Krankenkasser).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-12 physical subscale at 13 weeks; Group 1: mean 6.4 (SD 19.4); n=326, Group 2: mean 5.2 (SD 9.8); n=365; SF-12 physical subscale 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: 6.4 (4.3, 8.5). Reported sham: 5.2 (4.2, 6.2). Baseline acupuncture: 30.5 (0.87). Baseline sham: 30.7 (0.89).
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: Reported gender, age, bmI, affected knees, mean duration

of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks
- Actual outcome: SF-12 mental subscale at 13 weeks; Group 1: mean 2.1 (SD 11.1); n=326, Group 2: mean 3 (SD 11.7); n=365; SF-12 mental subscale 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: 2.1 (0.9, 3.3). Reported sham: 3.0 (2.1, 4.5). Baseline acupuncture: 48.9 (1.58). Baseline sham: 48.9 (1.52).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-12 physical subscale at 26 weeks; Group 1: mean 7 (SD 14.7); n=326, Group 2: mean 5.9 (SD 9.8); n=365; SF-12 physical subscale 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: 7.0 (5.9, 8.1). Reported sham: 5.9 (4.9, 6.9). Baseline acupuncture: 30.5 (0.87). Baseline sham: 30.7 (0.89).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks
- Actual outcome: SF-12 mental subscale at 26 weeks; Group 1: mean 1.6 (SD 11.1); n=326, Group 2: mean 3.1 (SD 11.7); n=365; SF-12 mental subscale 0-100 Top=High is good outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: 1.6 (0.4, 2.8). Reported sham: 3.1 (1.9, 4.3). Baseline acupuncture: 48.9 (1.58). Baseline sham: 48.9 (1.52).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 13 weeks; Group 1: mean -2.2 (SD 2.4); n=326, Group 2: mean -2 (SD 2.44); n=365; WOMAC pain 0-10 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: -2.2 (-2.47, -1.95). Reported sham: -2.0 (-2.22, -1.72). Baseline acupuncture: 5.3 (5.04, 5.46). Baseline sham: 5.3 (5.12, 5.53).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no

telephone interview after 26 weeks

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean -2.3 (SD 2.5); n=326, Group 2: mean -2.1 (SD 2.5); n=365; WOMAC pain 0-10 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: -2.3 (-2.60, -2.05). Reported sham: -2.1 (-2.37, -1.85). Baseline acupuncture: 5.3 (5.04, 5.46). Baseline sham: 5.3 (5.12, 5.53).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 13 weeks; Group 1: mean -2.1 (SD 2.4); n=326, Group 2: mean -1.9 (SD 2.5); n=365; WOMAC function 0-10 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: -2.1 (-2.36, -1.83). Reported sham: -1.9 (-2.14, -1.63). Baseline acupuncture: 5.4 (5.23, 5.64). Baseline sham: 5.6 (5.41, 5.80).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 26 weeks; Group 1: mean -2.2 (SD 2.5); n=326, Group 2: mean -2 (SD 2.5); n=365; WOMAC function 0-10 Top=High is poor outcome; Comments: Reports change scores and 95% confidence intervals. Reported acupuncture: -2.2 (-2.49, -1.94). Reported sham: -2.0 (-2.29, -1.77). Baseline acupuncture: 5.4 (5.23, 5.64). Baseline sham: 5.6 (5.41, 5.80).
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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks

Protocol outcome 7: Serious adverse events at > 3 months
- Actual outcome: Observed adverse events at 26 weeks; Group 1: 179/326, Group 2: 177/365; Comments: Including... arthralgia, bone pain, haematoma, back pain, joint lock, condition aggravated, localised osteoarthritis, influenza-like illness, sciatica, headache, meniscus lesion, contusion, joint effusion, nasopharyngitis, myalgia, cervical root pain, fall, bronchitis, acute, gastroenteritis, groin pain, bursitis, diarrhoea, joint sprain, neck pain, pain in extremity, periarthritis, rotator cuff syndrome, sinobronchitis, vertigo, vomiting, abdominal pain upper, epidcondylitis, gastritis, joint swelling, migraine, pneumonia, rhinitis, venous insufficiency, abdominal pain lower, angina pectoris, bone spur, carpal tunnel syndrome, cervicobrachial syndrome, cystitis, depressed mood, diabetes

mellitus, dyssomnia, gastroduodenitis, gout, hypertension, hypoaesthesia, knee arthropalsty, metatarsalgia, osteoarthritis, osteoporosis, pain, pharyngitis, phlebitis, post procedural haemorrhage, post procedural pain, tendonitis, tenosynovitis, acute sinusitis, allergy to chemicals, angioneurotic oedema, application site pain, arrhythmia, arterial bypass operation, arthropod bit,e, atrial fibrillation, carcinoembryonic antigen increased, cardiac death, cardiac aoperation, cervix carcinoma, cholecystitis, cholelithiasis, chondropathy, coccydynia, cough, deep vein thrombosis, dermatitis, dermatitis allergic, dermatitis contact, diverticulitis, dyspnoea, dystonia, ear discomfort, exanthem, excoriation, external ear disorders, eye haemorrhage, eyelid infection, facet joint syndrome, family stress, fatigue, fibrocystic breast disease, fibromyalgia, finger crushing, flatulence, foot fracture, foreign body trauma, furuncle, ganglion, gastric haemorrhage, gastroenteritis bacterial, goitre, hip dysplasia, hyperglycaemia, influenza, injury, iron deficiency anaemia, ischaemic stroke, joint range of motion decreased, localised oedema, localised skin reaction, lung nodule, melaena, metastases to lymph nodes, muscle rupture, muscle strain, myocardial infarction, neck shoulder and arm syndrome, nodal osteoarthritis, oedema peripheral, orchitis, pharyngolaryngeal pain, piriformis syndrome, plantar fasciitis, pleural effusion, polyarthritis, prostatits, pruritus, psychiatric symptom, psychosomatic disease, radicular pain, radius fracture, renal insufficiency, restless legs syndrome, rib fracture, skin fussures, small-cell lung cancer stage unspecified, spinal osteoarthritis, sudden hearing loss, synovial disorder, synovitis, temporal arteritis, tendon disorder, tenosynovitis stenosans, thoracic vertebral fracture, thrombophlebitis, thyroidectomy, toe operation, toothache, transaminases increased, transient ischaemic attack, traumatic haematoma, umbilical hernia, urinary tract infection, urticaria, varicose vein, viral upper respiratory tract infection 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: Reported gender, age, bmI, affected knees, mean duration of pain, Kellgren Lawrence score , medication use and baseline values of outcomes; Group 1 Number missing: 12, Reason: 4 withdrew immediately after randomisation, 8 had no telephone interview after 26 weeks; Group 2 Number missing: 7, Reason: 2 withdrew immediately after randomisation, 5 had no telephone interview after 26 weeks	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	Suarez-almazor 2010[123]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=560)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks of treatment, 12 weeks follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Knee osteoarthritis according to the American College of Rheumatology criteria with all people having a radiologic diagnosis of osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People who were at least 50 years of age and had knee osteoarthritis according to the American College of Rheumatology criteria. All people had a radiologic diagnosis of osteoarthritis. Additional inclusion criteria were: pain in the knee in the preceding 2 weeks at least 3/10 on a visual analog scale; no prior treatment with acupuncture; stable treatment with NSAIDs and analgesics in the previous months, if receiving glucosamine a stable dose for the past 2 months
Exclusion criteria	Intra-articular injections in the knee in the previous 2 months
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 64.5 (9.2). Gender (M:F): 186:338. Ethnicity: White = 358, African American = 98, Hispanic = 45, Other = 26
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms (mean [SD]): 9.2 (10.4) years
Indirectness of population	No indirectness
Interventions	(n=153) Intervention 1: Electroacupuncture. Electroacupuncture using TENS equipment. Traditional Chinese acupuncture using points on the basis of clinical practice. TENS was set to emit a dense disperse wave impulse at 50 HZ, dispersing at 15Hz, 20 cycles/minute. Voltage was increased slowly from 5V to 60V until maximal tolerance was achieved. People rested for 20 minutes with continuous TENS. Additionally the group was split into two sections, with some participants having a communication style meant to cause high expectations being used while others had a

	communication style meant to cause neutral expectations (these groups were combined for the analysis).. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=302) Intervention 2: Sham acupuncture. Sham acupuncture using a shallower needle placement and tENS device set to deliver a 40Hz adjustable wave instead with the voltage increased until the person could feel it, and then immediately turned off. Otherwise the treatments were comparable. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=72) Intervention 3: No intervention - Acupuncture compared to no treatment. Waiting list control. Duration 12 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (Supported by the National Institute of Arthritis and musculoskeletal and Skin Disorders (grant R01-AR49999). Dr. Suarez-Almazor holds a K24 career award from the National Institute of Arthritis and Musculoskeletal and skin disorders and is the Director of the Houston Center for Education and Research on Therapeutics, supported by the Agency for Health Research and Quality (grant U18-HS016093))

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-12 physical component summary at 12 weeks; Group 1: mean 39.5 (SD 9.7); n=153, Group 2: mean 38.7 (SD 10.1); n=302; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 35.0 (9.9). Baseline sham: 33.5 (8.7).
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, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 19, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Sham: 19 drop outs.
- Actual outcome: SF-12 mental component summary at 12 weeks; Group 1: mean 54.1 (SD 8.2); n=153, Group 2: mean 53.2 (SD 8.9); n=302; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 52.3 (9.4). Baseline sham: 53.4 (9.3).
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, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to

traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 19, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Sham: 19 drop outs.

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 30.8 (SD 17.9); n=153, Group 2: mean 31 (SD 19.1); n=302; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 44.5 (18.4). Baseline sham: 45.0 (18.2).
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, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 19, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Sham: 19 drop outs.

Protocol outcome 3: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 31.2 (SD 17.9); n=153, Group 2: mean 32.1 (SD 18.3); n=302; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 42.9 (19.0). Baseline sham: 42.9 (19.0).
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, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 19, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Sham: 19 drop outs.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE COMPARED TO 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 39.5 (SD 9.7); n=153, Group 2: mean 35.8 (SD 8.9); n=72; SF-12 physical component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 35.0 (9.9). Baseline no treatment: 35.3 (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: Reported sex, age, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 8, Reason: 8 drop outs
- Actual outcome: SF-12 mental component summary at 12 weeks; Group 1: mean 54.1 (SD 8.2); n=153, Group 2: mean 51.6 (SD 9.8); n=72; SF-12 mental component summary 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 52.3 (9.4). Baseline no treatment: 53.7 (10.7).
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 sex, age, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 8, Reason: 8 drop outs

Protocol outcome 2: Pain at ≤ 3 months - Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 30.8 (SD 17.9); n=153, Group 2: mean 42.4 (SD 16.8); n=72; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 44.5 (18.4). Baseline no treatment: 44.1 (15.2). 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: Reported sex, age, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 8, Reason: 8 drop outs Protocol outcome 3: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 12 weeks; Group 1: mean 31.2 (SD 17.9); n=153, Group 2: mean 41.7 (SD 18); n=72; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 42.9 (19.0). Baseline no treatment: 44.1 (15.2). 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: Reported sex, age, ethnicity, educational level, duration of knee pain, medication usage and baseline values of outcomes; Group 1 Number missing: 14, Reason: Difficult to say. 25 dropped out before allocation to traditional and sham groups. Traditional: 14 drop outs.; Group 2 Number missing: 8, Reason: 8 drop outs	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	Takeda 1994[127]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=40)
Countries and setting	Conducted in Canada; 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: Grade 1-4 radiographic osteoarthritis with pain in one or both knees
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People were volunteers with grade 1-4 osteoarthritis of the affected knee with: pain in one or both knees; radiological evidence of osteoarthritis; no change in medications for arthritis and other conditions in the last 3 months; no previous experience of acupuncture of the knee
Exclusion criteria	Serious systemic condition (such as diabetes); any neurologic or musculoskeletal condition (including fibromyalgia); had hemophilia; received intraarticular steroid injections in the previous 2 months; were receiving any treatment other than medication for their arthritis; had reconstructive surgery on the affected knee
Recruitment/selection of patients	No additional information
Age, gender and ethnicity	Age - Mean (SD): 61.6 (9.4). Gender (M:F): 20:20. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Grade 1-4, median grade 2 Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=20) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture three times a week for 3 weeks. 30mm needles with 0.23mm diameter into the five acupuncture points, specifically for knee and osteoarthritis pain. The needles were inserted, rotated, and inserted deeper until the subject experienced Te chi, or to the full depth of the needle if no Te chi was experienced. The needles were left in the subject for 30 minutes and each was rotated back and forth manually for 5 minutes. Duration 3 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness

	Further details: 1. Acupuncture or dry needling: Acupuncture (n=20) Intervention 2: Sham acupuncture. Sham acupuncture using the same type of needles but only inserted superficially approximately 1 inch from the acupuncture points in areas not considered active acupuncture points. The needles were only touched periodically to give the impression that movement of the needles was taking place. The location of the points was the same for all control subjects. Duration 3 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE Protocol outcome 1: Pain at ≤ 3 months - Actual outcome: WOMAC pain at 3 weeks; Group 1: mean 14.01 (SD 12.29); n=20, Group 2: mean 19.44 (SD 18.91); n=20; WOMAC pain 5-25 Top=High is poor outcome; Comments: Baseline acupuncture: 19.44. Baseline sham: 21.93 (8.71). 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, radiographic score and baseline values of outcomes. Different values for function at baseline.; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 3 weeks; Group 1: mean 48.03 (SD 43.58); n=20, Group 2: mean 60.02 (SD 45.85); n=20; WOMAC function 17-85 Top=High is poor outcome; Comments: Baseline acupuncture: 61.44 (43.15). Baseline sham: 77.80 (36.55). 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, height, weight, BMI, radiographic score and baseline values of outcomes. Different values for function 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; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Tu 2021[135] (Tu 2019[134], Wang 2021[145])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=480)
Countries and setting	Conducted in China; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks of treatment, 26 weeks of follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People reporting knee pain for longer than 6 months with a radiological confirmation of osteoarthritis (Kellgren Lawrence score 2-3)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People diagnosed with knee osteoarthritis according to the American College of Rheumatology clinical criteria; age 45 to 75 years; reported knee pain for longer than 6 months; had radiologic confirmation of osteoarthritis (Kellgren Lawrence score 2-3); pain score greater than 4 on the numeric rating scale.
Exclusion criteria	History of knee arthroplasty for the most painful knee or waiting for any knee surgery for either knee; knee pain caused by other diseases; arthroscopy in the last 12 months or intra-articular injection within the previous 6 months; acupuncture treatment in the last 3 months; serious acute or chronic organic diseases or psychiatric disorders; blood coagulation disorders; cardiac pacemaker; metal allergy or needle phobia; pregnancy or breastfeeding; participated in other clinical trials in the past 3 months
Recruitment/selection of patients	People were recruited from 9 study hospitals
Age, gender and ethnicity	Age - Mean (SD): 62.8 (7.1). Gender (M:F): 106:336. Ethnicity: Han = 430, minorities = 12
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Low comorbidity score (0 concomitant diseases: 226, 1 concomitant disease: 145, 2 concomitant diseases: 59, 3 or more concomitant diseases: 12). 4. Site of osteoarthritis: Knee
Extra comments	Severity: Radiological grade 2-3, median grade 2 Duration of symptoms: 6.6 (5.7) years

Indirectness of population	No indirectness
Interventions	(n=156) Intervention 1: Electroacupuncture. Electroacupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. Disposable sterile needles (0.25mm x 25-40mm) and HANS-200 electroacupuncture devices were used. The prescription was based on traditional Chinese medicine. Five obligatory acupoints and three adjunct acupoints were used, the obligatory points including: Dubi (ST35), Neixiyan (EX-LE5), Ququan (LR8), Xiyangguan (GB33) and an Ashi point (the point where the participant felt the most pain). Adjunct acupoints were selected from an acupoint pool. De qi was required. Electrodes should be attached to the handles of needles at LR8, GB33 and two adjunct acupoints in both groups. A dilatational wave of 2/100Hz was chosen and the electric current was gradually increased untilt he needles began to vibrate slightly. During treatment, the power light was switched on (as for all groups).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=155) Intervention 2: Acupuncture/dry needling - Acupuncture. Manual acupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. Disposable sterile needles (0.25mm x 25-40mm) and HANS-200 electroacupuncture devices were used. The prescription was based on traditional Chinese medicine. Five obligatory acupoints and three adjunct acupoints were used, the obligatory points including: Dubi (ST35), Neixiyan (EX-LE5), Ququan (LR8), Xiyangguan (GB33) and an Ashi point (the point where the participant felt the most pain). Adjunct acupoints were selected from an acupoint pool. De qi was required. Electrodes should be attached to the handles of needles at LR8, GB33 and two adjunct acupoints in both groups. No electrical stimulation was delivered. During treatment, the power light was switched on (as for all groups).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=157) Intervention 3: Sham acupuncture. Sham acupuncture in thirty minute sessions delivered three times weekly for 8 weeks, with 24 sessions in total. Disposable sterile needles (0.25mm x 25-40mm) and HANS-200 electroacupuncture devices were used. The prescription was based on traditional Chinese medicine. Eight non-acupoints were used. Electrodes should be attached to the handles of needles at 4 non-acupoints in both groups. During treatment, the power light was switched on (as for all groups).. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness

	Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (This trial was funded by Beijing Municipal Science & Technology Commission (D171100003217003) and Beijing Municipal Administration of Hospitals (XMLX201607).)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-12 physical health at 8 weeks; Group 1: mean 38.97 (SD 8.33); n=151, Group 2: mean 39.22 (SD 8.26); n=145; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 30.89 (8.01). Baseline acupuncture: 31.60 (8.08).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 8 weeks; Group 1: mean 53.73 (SD 9.23); n=151, Group 2: mean 54.67 (SD 8.51); n=145; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 51.47 (11.30). Baseline acupuncture: 51.11 (11.21).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-12 physical health at 26 weeks; Group 1: mean 39.18 (SD 8.79); n=151, Group 2: mean 39.2 (SD 8.61); n=145; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 30.89 (8.01). Baseline acupuncture: 31.60 (8.08).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 26 weeks; Group 1: mean 54.11 (SD 8.04); n=151, Group 2: mean 54.92 (SD 8.22); n=145; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 51.47 (11.30). Baseline acupuncture: 51.11 (11.21).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 2.5 (SD 1.95); n=151, Group 2: mean 2.79 (SD 1.91); n=145; WOMAC pain 0-20 Top=High is

poor outcome; Comments: Baseline electroacupuncture: 6.68 (2.93). Baseline acupuncture: 6.55 (2.75).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean 2.79 (SD 2.41); n=151, Group 2: mean 3.16 (SD 2.03); n=145; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline electroacupuncture: 6.68 (2.93). Baseline acupuncture: 6.55 (2.75).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 9.26 (SD 7.03); n=151, Group 2: mean 10.82 (SD 7.32); n=145; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 21.09 (9.00). Baseline acupuncture: 20.70 (8.92).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 26 weeks; Group 1: mean 9.87 (SD 7.04); n=151, Group 2: mean 11.29 (SD 7.7); n=145; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 21.09 (9.00). Baseline acupuncture: 20.70 (8.92).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 7: Serious adverse events at > 3 months
- Actual outcome: Adverse events at 26 weeks; Group 1: 18/156, Group 2: 22/155; Comments: Electroacupuncture: 18 (0 severe adverse events, 7 subcutaneous hematoma, 10 post-needling pain, 1 pantalgia). Manual acupuncture: 22 (0 severe adverse events, 10 subcutaneous hematoma, 13 post-needling pain, 0 pantalgia).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-12 physical health at 8 weeks; Group 1: mean 38.97 (SD 8.33); n=151, Group 2: mean 38.42 (SD 9.13); n=146; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 30.89 (8.01). Baseline sham: 30.94 (7.77).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 8 weeks; Group 1: mean 53.73 (SD 9.23); n=151, Group 2: mean 53.21 (SD 8.95); n=146; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 51.47 (11.30). Baseline sham: 51.35 (10.93).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-12 physical health at 26 weeks; Group 1: mean 39.18 (SD 8.79); n=151, Group 2: mean 38.05 (SD 8.11); n=146; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 30.89 (8.01). Baseline sham: 30.94 (7.77).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 26 weeks; Group 1: mean 54.11 (SD 8.04); n=151, Group 2: mean 51.34 (SD 9.37); n=146; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 51.47 (11.30). Baseline sham: 51.35 (10.93).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 2.5 (SD 1.95); n=151, Group 2: mean 3.57 (SD 2.59); n=146; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline electroacupuncture: 6.68 (2.93). Baseline sham: 6.40 (2.73).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean 2.79 (SD 2.41); n=151, Group 2: mean 3.94 (SD 2.7); n=146; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline electroacupuncture: 6.68 (2.93). Baseline sham: 6.40 (2.73).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 9.26 (SD 7.03); n=151, Group 2: mean 11.78 (SD 8.17); n=146; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 21.09 (9.00). Baseline sham: 20.77 (8.27).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 26 weeks; Group 1: mean 9.87 (SD 7.04); n=151, Group 2: mean 13.21 (SD 9.09); n=146; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 21.09 (9.00). Baseline sham: 20.77 (8.27).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 7: Serious adverse events at > 3 months
- Actual outcome: Adverse events at 26 weeks; Group 1: 18/156, Group 2: 17/157; Comments: Electroacupuncture: 18 (0 severe adverse events, 7 subcutaneous hematoma, 10 post-needling pain, 1 pantalgia). Sham acupuncture: 17 (0 severe adverse events, 9 subcutaneous hematoma, 10 post-needling pain, 0 pantalgia).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-12 physical health at 8 weeks; Group 1: mean 39.22 (SD 8.26); n=145, Group 2: mean 38.42 (SD 9.13); n=146; SF-12 physical health

0-100 Top=High is good outcome; Comments: Baseline acupuncture: 31.60 (8.08). Baseline sham: 30.94 (7.77).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 8 weeks; Group 1: mean 54.67 (SD 8.51); n=145, Group 2: mean 53.21 (SD 8.95); n=146; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline electroacupuncture: 51.47 (11.30). Baseline sham: 51.35 (10.93).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-12 physical health at 26 weeks; Group 1: mean 39.2 (SD 8.61); n=145, Group 2: mean 38.05 (SD 8.11); n=146; SF-12 physical health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 31.60 (8.08). Baseline sham: 30.94 (7.77).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).
- Actual outcome: SF-12 mental health at 26 weeks; Group 1: mean 54.92 (SD 8.22); n=145, Group 2: mean 51.34 (SD 9.37); n=146; SF-12 mental health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 51.11 (11.21). Baseline sham: 51.35 (10.93).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 2.79 (SD 1.91); n=145, Group 2: mean 3.57 (SD 2.59); n=146; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 6.55 (2.75). Baseline sham: 6.40 (2.73).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 26 weeks; Group 1: mean 3.16 (SD 2.03); n=145, Group 2: mean 3.94 (SD 2.7); n=146; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 6.55 (2.75). Baseline sham: 6.40 (2.73).
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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%). Protocol outcome 5: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 8 weeks; Group 1: mean 10.82 (SD 7.32); n=145, Group 2: mean 11.78 (SD 8.17); n=146; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline electroacupuncture: 21.09 (9.00). Baseline sham: 20.77 (8.27). 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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%). Protocol outcome 6: Physical function at > 3 months - Actual outcome: WOMAC function at 26 weeks; Group 1: mean 11.29 (SD 7.7); n=145, Group 2: mean 13.21 (SD 9.09); n=146; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 20.70 (8.92). Baseline sham: 20.77 (8.27). 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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%). Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Adverse events at 26 weeks; Group 1: 22/155, Group 2: 17/157; Comments: Manual acupuncture: 22 (0 severe adverse events, 10 subcutaneous hematoma, 13 post-needling pain, 0 pantalgia). Sham acupuncture: 17 (0 severe adverse events, 9 subcutaneous hematoma, 10 post-needling pain, 0 pantalgia). 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: Reported age, gender, race, BMI, duration, radiological grade, affected knee, treatment in the past, concomitant diseases, history of acupuncture, and baseline values of outcomes; Group 1 Number missing: -, Reason: Overall 407 people completed the study (84.8%).; Group 2 Number missing: -, Reason: Overall 407 people completed the study (84.8%).	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study (subsidiary papers)	Vas 2004[139] (Vas 2006[138])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=97)
Countries and setting	Conducted in Spain; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 11 weeks of treatment, 12 weeks of follow up total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Outpatient who had been clinically and radiologically diagnosed according to the criteria of the American College of Rheumatology
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People with an age of 45 years or older; with pain in one or both knees for the preceding 3 months or longer; radiological evidence of osteoarthritis of the knee (at least grade 1 according to the Ahlback classification)
Exclusion criteria	Previous treatment with acupuncture; contraindication to medication with diclofenac; inflammatory, metabolic, or neuropathic arthropathies; severe concomitant illnesses that might interfere with the clinical evaluation of the person; severe or generalised dermopathy; pregnancy or existing treatment with antineoplastic, corticoid, or immunosuppressive drugs
Recruitment/selection of patients	Recruited from a public primary care center in southern Spain, over a period of 2 years
Age, gender and ethnicity	Age - Mean (SD): 67.1 (10.2). Gender (M:F): 16:81. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Ahlback grade 1-4, median grade 2 Duration of symptoms (mean [SD]): 7.5 (8.6) years
Indirectness of population	No indirectness
Interventions	(n=48) Intervention 1: Electroacupuncture. Insertion of sterile, single use, 30 gauge and 45mm length acupuncture needles into the local points GB34, SP9, EX-LE5, and ST36 and the distal points KI3, SP6, LI4 and ST40. For each of the points the person determined the sensation of deqi. A WQ-10D1 electrostimulator was used to stimulate all the needles inserted into the local points electrically, in pairs. The treatment lasted 12 weeks, ending with visit 11, with the final evaluation on week 12. Duration 11

	weeks. Concurrent medication/care: All people were given a bag with 21 tablets of 50mg diclofenac for the week (50mg every 8 hours) with instruction to reduce the dose if symptoms improved. People with risk factors received gastroprotective drugs. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=49) Intervention 2: Sham acupuncture. Placebo acupuncture at the same frequency and for the same duration as for the group receiving the true intervention. Retractable needles that went into small adhesive cylinders, such that the needle was supported but did not perforate the skin. The acupuncturist then placed the needles over the same point as were used for the true acupuncture group. They connected the same pairs of electrodes and stimulated the electrical connection. Duration 11 weeks. Concurrent medication/care: All people were given a bag with 21 tablets of 50mg diclofenac for the week (50mg every 8 hours) with instruction to reduce the dose if symptoms improved. People with risk factors received gastroprotective drugs. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (This study was partly financed by Servicio Andaluz de Salud (Grant No 192/99). The acupuncture materials and the drugs used in the study were provided by the Sevilla-Sur health district authorities)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: Profile of quality of life in the chronically ill (PLQC) physical capability at 12 weeks; Group 1: mean 2.8 (SD 0.7); n=48, Group 2: mean 2.5 (SD 0.8); n=49; PLQC physical capability 0-4 Top=High is good outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 2.1 (0.6). Baseline sham: 1.9 (0.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons
- Actual outcome: Profile of quality of life in the chronically ill (PLQC) psychological functioning at 12 weeks; Group 1: mean 2.7 (SD 0.4); n=48, Group 2: mean 2.5 (SD 0.6); n=49; PLQC psychological functioning 0-4 Top=High is good outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 2.2 (0.5). Baseline sham: 2.2 (0.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons

- Actual outcome: Profile of quality of life in the chronically ill (PLQC) negative mood at 12 weeks; Group 1: mean 3.2 (SD 0.7); n=48, Group 2: mean 3.1 (SD 0.7); n=49; PLQC negative mood 0-4 Top=High is good outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 2.8 (0.8). Baseline sham: 2.8 (0.8).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons
- Actual outcome: Profile of quality of life in the chronically ill (PLQC) social functioning at 12 weeks; Group 1: mean 2.8 (SD 0.5); n=48, Group 2: mean 2.7 (SD 0.7); n=49; PLQC social functioning 0-4 Top=High is good outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 2.4 (0.5). Baseline sham: 2.2 (0.7).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons
- Actual outcome: Profile of quality of life in the chronically ill (PLQC) social wellbeing at 12 weeks; Group 1: mean 3.2 (SD 0.5); n=48, Group 2: mean 3.2 (SD 0.5); n=49; PLQC social wellbeing 0-4 Top=High is good outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 3.1 (0.4). Baseline sham: 3.0 (0.6).
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - High, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low; Indirectness of outcome: No indirectness ; Baseline details: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 1.7 (SD 2.6); n=48, Group 2: mean 6.4 (SD 5.8); n=49; WOMAC pain 0-20 Top=High is poor outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 12.4 (3.4). Baseline sham: 12.1 (4.0).
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: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons

Protocol outcome 3: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 7.4 (SD 10.3); n=48, Group 2: mean 24.9 (SD 20.4); n=49; WOMAC function 0-68 Top=High is poor outcome; Comments: Results are after a bivariate analysis. Baseline acupuncture: 40.5 (12.2). Baseline sham: 41.5 (13.9).
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: Reported age, gender, marital status, education, diagnosis, Ahlback score, mean duration of osteoarthritis, mean BmI, knee affected and baseline values of outcomes; Group 1 Number missing: 1, Reason: 1 fear of acupuncture; Group 2 Number missing: 8, Reason: 6 lack of improvement, 2 personal reasons

Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	**Wang 2020[146]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in China; Setting: Three hospitals in Bejing
Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks + 8 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Participants needed to be aged 45-75 years and have Kellgren-Lawrence grade II or III (mild or moderate) radiographically confirmed KOA affecting one or both knees with a duration of more than 6 months and pain intensity of 40 or more on a 100-point visual analogue scale
Exclusion criteria	A history of knee surgery or arthroscopy; pain in the knee caused by floating cartilage, joint effusion or inflammatory, malignant or autoimmune disease; serious acute or chronic organic disease or mental disorder; pregnancy or breast feeding; and history of bleeding disorder. Participants were also ineligible if they have received acupuncture treatment or participated in other clinical trials in the past 3 months
Recruitment/selection of patients	Participants were recruited via the community through media, outpatient, and poster paper advertisements at three hospital centres
Age, gender and ethnicity	Age - Mean (SD): Electroacupuncture group: 58.89 (6.75); manual acupuncture group: 59.70 (7.36). Gender (M:F): Define. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Duration of symptoms (mean, SD): Electroacupuncture group: 69.93 (56.69) months; manual acupuncture group: 73.20 (56.71) months

	Severity of symptoms: All participants had to have a Kellgren-Lawrence grade II or III and score 40 or more on a pain intensity VAS
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Electroacupuncture. Acupuncturists had Chinese medicine practitioner licenses and at lest 3 years of clinical experience. Huatuo brand disposable sterile steel needles (0.2 x 40 mm) were used. All participants underwent acupuncture needling at a selection of local and distant traditional acupuncture points or ah shi points chosen by the acupuncturists according to the principles of traditional Chinese medicine. Needles were inserted at 6-7 local points. Individual syndrome differentiation was used. If pain occurred on the outside of the affected knee joint, GB points were mainly selected. If pain occurred in front of the affected knee joint, ST points were selected. If pain occured in the interior of the affected knee joint, SP, LR and KI points were chosen. If pain occurred in the rear of the affected knee, BL joints were used. Needles were stimulated manually for 1- seconds to achieve di qi sensation. There was a total of 24 sessions lasting 30 minutes each, over a period of 8 weeks. In the EA group, an electrical apparatus producing a density wave with a frequency of 2/100Hz was connected to the needles with alligator clips to stimulate pairs of needles inserted at ST-36-GB34 and ST34-SP10. The fixed current intensity was uniformly 0.2mA.. Duration 8 weeks. Concurrent medication/care: Participants were advised not to take any NSAIDs or analgesics except for a 'rescue medication'. The use of NSAIDs was recorded. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=30) Intervention 2: Acupuncture/dry needling - Acupuncture. Acupuncture points were selected as described for the electroacupuncture group. Participants in the manual acupuncture group had the same schedule as the electroacupuncture except that the electrical apparatus featured a working power indicator and sound without actual current output. The middle wire was cut, although the appearance of the unit was identical. Thus, the EA instrument appeared to be "on", but the actual power was not energised. After elicitation of de qi sensation by MA, needles were retained for 30 minutes. Although no manual manipulation of the needles was performed after initially achieving de qi sensation, the stimulation associated with needle retention was still expected to induce therapeutic effects. Therefore the only difference between the two groups was the electrical current in the electroacupuncture group. Duration 8 weeks. Concurrent medication/care: Participants were advised not to take any NSAIDs or analgesics except for a 'rescue medication'. The use of NSAIDs was recorded. Indirectness: No indirectness

	Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (Supported by grants from Bejing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding Support, and Bejing Municipal Science and Technological Comission)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: Quality of life at 12 weeks; Group 1: mean 62.21 (SD 16.9); n=28, Group 2: mean 60.87 (SD 18.42); n=30; SF12 0-100 Top=High is good outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 57.65 (4.98); Manual acupuncture group: 56.98 (5.45)
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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: Quality of life at 16 weeks; Group 1: mean 63.64 (SD 16.63); n=28, Group 2: mean 61.87 (SD 18.67); n=30; SF12 0-100 Top=High is good outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 57.65 (4.98); Manual acupuncture group: 56.98 (5.45)
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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: Pain at 12 weeks; Group 1: mean 2.54 (SD 2.08); n=28, Group 2: mean 3.43 (SD 2.31); n=30; WOMAC - pain 0-20 Top=High is poor outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 7.0 (3.34); Manual acupuncture group: 6.77 (2.42)
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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0

Protocol outcome 4: Pain at > 3 months
- Actual outcome: Pain at 16 weeks; Group 1: mean 2.79 (SD 2.21); n=28, Group 2: mean 3.6 (SD 2.84); n=30; WOMAC 0-20 Top=High is poor outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 7.0 (3.34); Manual acupuncture group: 6.77 (2.42)
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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet

computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0 Protocol outcome 5: Physical function at ≤ 3 months - Actual outcome: Physical function at 12 weeks; Group 1: mean 10.04 (SD 7.07); n=28, Group 2: mean 11.07 (SD 6.93); n=30; WOMAC - physical function 0-68 Top=High is poor outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 23.64 (7.37); Manual acupuncture group: 23.80 (9.71) 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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0 Protocol outcome 6: Physical function at > 3 months - Actual outcome: Physical function at 16 weeks; Group 1: mean 9.39 (SD 7.83); n=28, Group 2: mean 12.33 (SD 8.62); n=30; WOMAC - physical function 0-68 Top=High is poor outcome; Comments: Standard deviations were calculated as the study reported 95% Confidence intervals. Baseline scores: Electroacupuncture group: 23.64 (7.37); Manual acupuncture group: 23.80 (9.71) 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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0 Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Hemarthrosis at 16 weeks; Group 1: 9/28, Group 2: 11/30; Comments: Hemarthrosis 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 ; Group 1 Number missing: 2, Reason: For two participants there was a tablet computer error which prevented a random number to be obtained in a timely fashion - these were excluded from the ITT analysis; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	**Wang 2021[147]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=60)
Countries and setting	Conducted in China; Setting: Five hospitals in Bejing

Line of therapy	Unclear
Duration of study	Intervention + follow up: 8 weeks + 18 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 45-75 years old, have Kellgren-Lawrence grade II or III (mild or moderate) radiographically confirmed KOA affecting one or both knees with a duration of more than six months and pain intensity ≥4 on a 10 point numerical rating scale
Exclusion criteria	Pain in the knee caused by joint effusion, floating cartilage, or inflammatory, malignant or autoimmune disease; a history of knee surgery or arthroscopy; pregnancy or breastfeeding; serious acute chronic organic disease or mental disorder and history of bleeding disorder. Participants were also ineligible if they had participated in other clinical trials or received acupuncture treatment in the past three months
Recruitment/selection of patients	Participants were recruited via the community through media, outpatient and poster paper advertisements at three hospitals
Age, gender and ethnicity	Age - Mean (SD): Electroacupuncture group: 64.73 (5.39); Sham acupuncture group 66.10 (7.42). Gender (M:F): 21/39. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Duration of symptoms (mean, SD): Electroacupuncture group: 85.73 (74.15); Sham acupuncture group: 104.37 (96.45) Severity of disease: All participants had Kellgren-Lawrence grade II or III. Pain score on numerical rating scale (mean, SD): Electroacupuncture group 6 (1.34); Sham group 6.13 (1.33)
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Electroacupuncture. All acupuncturists had Chinese medicine practitioner licenses and at least three years of clinical experience. Huato brand disposable, sterile steel needles (0.3 x 40 mm) were used. Acupuncture treatment was semi-standardised. Acupoints were chosen according to traditional Chinese medicine and were localised according to the WHO Standard Participants Acupuncture Locations. Needles were stimulated manually for 10 seconds to achieve de qi sensation. Treatment consisted of 24 sessions lasting 30 minutes each over 8 weeks (usually three times per week). In the electroacupuncture group, an electrical apparatus (HANS-200A acupoint nerve stimulator, Nanjing Jisheng Medical Co. Ltd.)

	producing a density wave with frequency of 2/100Hz was connected to the needles with alligator clips to stimulate pairs of needles inserted at LR8 and GB33 and another two adjunct acupoints by the research assistant. The fixed current intensity was uniformly 0.2mA. Duration 8 weeks. Concurrent medication/care: All participants were advised not to take any NSAIDs or analgesics except for a "rescue analgesic" (1 tablet of 200mg Acetaminophen orally as needed, once per day). The use of NSAIDs was recorded. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=30) Intervention 2: Sham acupuncture. Eight non-acupoints that were separate from conventional acupoints or meridians were used for the sham group. The schedule, electrode placements and other treatment settings were the same as for the electroacupuncture group but with superficial skin penetration (2-3 mm in depth) and no electricity output or needle manipulation for de qi. Duration 8 weeks. Concurrent medication/care: All participants were advised not to take any NSAIDs or analgesics except for a "rescue analgesic" (1 tablet of 200mg Acetaminophen orally as needed, once per day). The use of NSAIDs was recorded. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Academic or government funding (Supported by a grant from Bejing Municipal Science and Technology Commission)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: Quality of life - physical at 8 weeks; Group 1: mean 37.13 (SD 5.57); n=30, Group 2: mean 35.22 (SD 6.31); n=30; SF36 - physical component summary 0-100 Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 33.19 (6.84); Sham group: 33.58 (5.65)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)
- Actual outcome: Quality of life - mental at 8 weeks; Group 1: mean 43.86 (SD 7.6); n=30, Group 2: mean 43.5 (SD 6.74); n=30; SF36 - mental component summary 0-100 Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 41.62 (6.95); Sham group: 42.41 (8.24)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: Quality of life - physical at 26 weeks; Group 1: mean 36.05 (SD 4.31); n=30, Group 2: mean 35.33 (SD 7.09); n=30; SF36 - physical

component summary 0-100 Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 33.19 (6.84); Sham group: 33.58 (5.65)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)
- Actual outcome: Quality of life - mental at 26 weeks; Group 1: mean 44.09 (SD 6.9); n=30, Group 2: mean 42.84 (SD 6.41); n=30; SF36 - mental component summary 0-100 Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 41.62 (6.95); Sham group: 42.41 (8.24)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: Pain at 8 weeks; Group 1: mean 2.73 (SD 1.72); n=30, Group 2: mean 4.17 (SD 2.98); n=30; WOMAC - pain 0-20 Top=High is poor outcome; Comments: Baseline scores: Electroacupuncture group: 6.97 (2.68); Sham group: 7.00 (2.60)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 4: Pain at > 3 months
- Actual outcome: Pain at 26 weeks; Group 1: mean 3.63 (SD 1.67); n=30, Group 2: mean 4.03 (SD 2.24); n=30; WOMAC - pain 0-20 Top=High is poor outcome; Comments: Baseline scores: Electroacupuncture group: 6.97 (2.68); Sham group: 7.00 (2.60)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: Physical function at 8 weeks; Group 1: mean 10.63 (SD 5.88); n=30, Group 2: mean 14.2 (SD 9.65); n=30; Woman - function 0-68 Top=High is poor outcome; Comments: Baseline scores: Electroacupuncture group: 21.63 (9.08); Sham group: 22.40 (8.25)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: Physical function at 26 weeks; Group 1: mean 15 (SD 6.58); n=30, Group 2: mean 13.57 (SD 6.52); n=30; WOMAC - physical function 0-68 Top=High is poor outcome; Comments: Baseline scores: Electroacupuncture group: 21.63 (9.08); Sham group: 22.40 (8.25)
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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)

Protocol outcome 7: Serious adverse events at > 3 months - Actual outcome: Hemarthrosis at 26 weeks; Group 1: 4/30, Group 2: 6/30 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 ; Group 1 Number missing: 1, Reason: Lost to follow up (1 = housework); Group 2 Number missing: 3, Reason: Lost to follow up (1 = relocation, 2 = business trip)	
Protocol outcomes not reported by the study	Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months

Study	Weiner 2007[149]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=88)
Countries and setting	Conducted in USA; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks of intervention, 3 weeks follow up in total
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Chronic knee pain and radiographic knee osteoarthritis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	English-speaking community-dwelling older adults (age at least 65 years) with chronic knee pain (i.e. knee pain on most or all days for at least 3 months) and radiographic knee osteoarthritis (Kellgren Lawrence grade 2-4, read by a rheumatologist marked to group assignment). All people were cognitively intact and signed informed consent before their participation.
Exclusion criteria	Cognitive impairment (Golstein mini-mental state examination score <24 adjusted for age and education); severe visual or hearing impairment; acute illness or pain; prior knee surgery; non-osteoarthritis arthritides; nonambulatory or ambulatory only with a walker; pain in the lower body more severe than knee pain; a large knee effusion or severe mechanical knee instability; corticosteroid or hyaluronic acid injection during the prior 3 months; immunosuppressive or anticoagulant medications; pacemaker; prior electroacupuncture treatment; acute or terminal illness
Recruitment/selection of patients	They were recruited by way of newspaper advertisements and screened in two phases: over the telephone and on site by one investigator.
Age, gender and ethnicity	Age - Mean (SD): 71.5 (5.4). Gender (M:F): 40:48. Ethnicity: White = 82, Black = 6
Further population details	1. Age (≤/> 75 years): >75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: High comorbidity score (Comorbidity mean (SD): 1.95 (1.30)). 4. Site of osteoarthritis: Knee
Extra comments	Severity: Kellgren Lawrence grade 2-4, median grade 4 Duration of symptoms (mean [SD]): 8.0 (7.4) years
Indirectness of population	No indirectness
Interventions	(n=44) Intervention 1: Electroacupuncture. Periosteal stimulation therapy with four sterile, single-use 30-gauge acupuncture needles being inserted into the following

	locations of the symptomatic knee until they just touched bone: medial femoral condyle, lateral femoral condyle, flare of tibia, and head of fibula. The needles were stimulated with 100Hz for 30 minutes. The intensity was adjusted so that it was clearly felt but not uncomfortable and was adjusted so that it was perceptible throughout the session. Two additional needles were inserted into the soft tissue on either side of the upper third of the tibial shaft, which was stimulated with 100Hz for 1 minute. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=44) Intervention 2: Sham acupuncture. Same needle insertion, but the needles inserted into bone were not stimulated (the needles in soft tissue were still stimulated with 100Hz for 1 minute).. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling
Funding	Other author(s) funded by industry (This work was supported by Grant R21 AG024288 from the National Institute on Aging, National Institutes of Health. Dr. Morone was supported by the Roadmap multidisciplinary Clinical Research Career Development Award Grant (K12 RR023267) from the National Institutes of Health. C. Kent Kwoh has consulting relationships and participates in advisory panels for TAP Pharmaceutical Products Inc., Abbott, Centocor and Glaxo Smith Kline. He has grants from TAP Pharmaceutical Products Inc., Centocor, and the Beverage Institute for other projects unrelated to this manuscript)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 8.32 (SD 3.93); n=44, Group 2: mean 7.97 (SD 3.94); n=44; WOMAC pain 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 9.25 (3.14). Baseline sham: 9.06 (3.32)
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: Reported age, sex, race, comorbidity, BMI, radiographic score, pain duration, physical performance, disease burden, treatment credibility and baseline values of outcomes; Group 1 Number missing: 5, Reason: 2 lost to follow up refusing to return to clinic, 3 discontinued as they didn't like treatment; Group 2 Number missing: 4, Reason: 3 lost to follow up, 1 discontinued treatment

Protocol outcome 2: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 12 weeks; Group 1: mean 21.36 (SD 11.6); n=44, Group 2: mean 22.61 (SD 11.84); n=44; WOMAC function 0-68 Top=High is poor outcome; Comments: Baseline acupuncture: 26.82 (10.59). Baseline sham: 27.22 (10.62).

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, race, comorbidity, BMI, radiographic score, pain duration, physical performance, disease burden, treatment credibility and baseline values of outcomes; Group 1 Number missing: 5, Reason: 2 lost to follow up refusing to return to clinic, 3 discontinued as they didn't like treatment; Group 2 Number missing: 4, Reason: 3 lost to follow up, 1 discontinued treatment Protocol outcome 3: Psychological distress at ≤ 3 months - Actual outcome: Geriatric depression scale at 12 weeks; Group 1: mean 3.78 (SD 4.82); n=44, Group 2: mean 3.36 (SD 3.38); n=44; Geriatric depression scale 0-20 Top=High is poor outcome; Comments: Baseline acupuncture: 3.14 (3.84). Baseline sham: 3.38 (3.81). 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, race, comorbidity, BMI, radiographic score, pain duration, physical performance, disease burden, treatment credibility and baseline values of outcomes; Group 1 Number missing: 5, Reason: 2 lost to follow up refusing to return to clinic, 3 discontinued as they didn't like treatment; Group 2 Number missing: 4, Reason: 3 lost to follow up, 1 discontinued treatment Protocol outcome 4: Serious adverse events at ≤ 3 months - Actual outcome: Serious adverse events at 12 weeks; Group 1: 0/44, Group 2: 0/44 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, race, comorbidity, BMI, radiographic score, pain duration, physical performance, disease burden, treatment credibility and baseline values of outcomes; Group 1 Number missing: 5, Reason: 2 lost to follow up refusing to return to clinic, 3 discontinued as they didn't like treatment; Group 2 Number missing: 4, Reason: 3 lost to follow up, 1 discontinued treatment	
Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related 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; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study	White 2012[151]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=221)
Countries and setting	Conducted in USA; 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: Chronic osteoarthritis pain from a single joint (hip or knee)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age 18 to 80 years; suffering chronic osteoarthritis pain from a single joint (hip or knee); awaiting joint replacement surgery; having a mean score of at least 30mm during the baseline week (7 daily recordings)on a 100mm VAS; not on any current physical treatment (e.g. physiotherapy)
Exclusion criteria	Pregnancy; serious comorbidity (including severe back pain); history of prolonged or current steroid use; awaiting hip/knee revision (i.e. current prosthesis); needle phobia; allergy to sticking plaster
Recruitment/selection of patients	Recruitment was via joint replacement waiting lists at Southampton General and Salisbury District Hospitals
Age, gender and ethnicity	Age - Mean (SD): 66.75 (8.29). Gender (M:F): 94:127. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Mixed (Osteoarthritis of the hip or knee).
Extra comments	Severity: Not stated Duration of symptoms: Not stated
Indirectness of population	No indirectness
Interventions	(n=74) Intervention 1: Acupuncture/dry needling - Acupuncture. Real Western acupuncture using a flexible but prescribed range of points. A mean 6 points were used at each treatment, with deep needling, which lasted for 20 minutes during the 30 minute appointment twice a week for 4 weeks. Deqi was elicited for each needle through needle rotation. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling

	(n=73) Intervention 2: Sham acupuncture. Sham needling using steitberger needles that were nonpenetrating needles instead of real acupuncture needles. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Dry needling (n=74) Intervention 3: Other. Mock electrical stimulation. Duration 4 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: This group was not included as they did not fulfill the inclusion criteria
Funding	Academic or government funding (P.W. and this study was funded by a Department of Health Postdoctoral Research Award. C.S. was also funded by the same award. G.L.'s post is partially funded by the Rufford Maurice Laing Foundation. The Southampton Complementary medicine Research Trust contributed funding for this trial.)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: VAS at 4 weeks; Group 1: mean 43.5 (SD 25.5); n=74, Group 2: mean 44 (SD 21.7); n=73; VAS 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 60.5 (14.2). Baseline sham: 58.6 (14.6).
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 baseline values of outcomes, otherwise unclear; Group 1 Number missing: 2, Reason: 1 work commitment, 1 pain exacerbation; Group 2 Number missing: 6, Reason: 1 not improving, 2 transport problems, 1 pain exacerbation, 1 time commitment, 1 called for surgery

Protocol outcome 2: Serious adverse events at ≤ 3 months
- Actual outcome: Adverse events at 4 weeks; Group 1: 22/74, Group 2: 4/73; Comments: Acupuncture: Temporary increase in pain = 5, bleed or bruise at needle site = 15, tired posttreatment = 1, became tearful posttreatment = 1. Sham: Temporary increase in pain = 1, bleed or bruise at needle site = 2, tired posttreatment = 1.
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 baseline values of outcomes, otherwise unclear; Group 1 Number missing: 2, Reason: 1 work commitment, 1 pain exacerbation; Group 2 Number missing: 6, Reason: 1 not improving, 2 transport problems, 1 pain exacerbation, 1 time commitment, 1 called for surgery

Protocol outcomes not reported by the study	Health-related quality of life at ≤ 3 months; Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months;

	Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study	Williamson 2007[154]
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=181)
Countries and setting	Conducted in United Kingdom; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 6 weeks treatment, 12 weeks follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: People listed for knee arthroplasty due to osteoarthritis with unilateral or bilateral knee pain for more than 3 months
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	People listed for knee arthroplasty due to osteoarthritis; people with unilateral or bilateral knee pain; pain lasting more than 3 months
Exclusion criteria	Taking anticoagulants; within 2 months after receiving an intra-articular steroid injection; experiencing back pain associated with referred leg pain; suffering from ipsilateral osteoarthritis of the hip; suffering psoriasis or other skin disease in the region of the knee; suffering from rheumatoid arthritis if they had received acupuncture or physiotherapy treatment in the last year
Recruitment/selection of patients	People on the waiting list for knee replacement surgery
Age, gender and ethnicity	Age - Mean (SD): 70.7 (9.0). Gender (M:F): 84:97. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): Mixed 2. Diagnosis: Not stated / Unclear 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity: Not stated Duration of symptoms: At least 3 months
Indirectness of population	No indirectness
Interventions	(n=60) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture once a week for 6 weeks. The needles (1 inch, 0.25 gauge) were inserted and de chi achieved where possible and left in situ for 20 minutes. The acupuncture points used were those most commonly used in studies (SP10, ST35, Xiyan, ST36, SP9, GB34, and LIV3). Up to three additional needles were used in trigger or traditional points at the physiotherapist's discretion. This was conducted in a group setting of 6-10 people per session. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture

	(n=60) Intervention 2: No intervention - Acupuncture compared to no treatment. Exercise and advice leaflet only. Duration 6 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: (n=60) Intervention 3: Other. Physiotherapy (mixed modality supervised exercise) with 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: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: This group was not included as they did not fulfill the inclusion criteria
Funding	Academic or government funding (Funded by the Research and Development Grant, The Great Western Hospital, Swindon)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE COMPARED TO NO TREATMENT

Protocol outcome 1: Pain at ≤ 3 months
- Actual outcome: VAS at 12 weeks; Group 1: mean 6.58 (SD 2.29); n=60, Group 2: mean 7.24 (SD 2.07); n=60; VAS 0-10 Top=High is poor outcome; Comments: Baseline acupuncture: 7.25 (2.46). Baseline control: 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 ; Baseline details: Reported age, BMI, gender and baseline values of outcomes; Group 1 Number missing: 16, Reason: 16 lost to follow up; Group 2 Number missing: 26, Reason: 26 lost to follow up

Protocol outcome 2: Psychological distress at ≤ 3 months
- Actual outcome: HADS anxiety at 12 weeks; Group 1: mean 6.88 (SD 4.15); n=60, Group 2: mean 6.54 (SD 3.93); n=60; HADS anxiety 0-21 Top=High is poor outcome; Comments: Baseline acupuncture: 7.25 (4.27). Baseline control: 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 ; Baseline details: Reported age, BMI, gender and baseline values of outcomes; Group 1 Number missing: 16, Reason: 16 lost to follow up; Group 2 Number missing: 26, Reason: 26 lost to follow up
- Actual outcome: HADS depression at 12 weeks; Group 1: mean 6.72 (SD 3.18); n=60, Group 2: mean 7.13 (SD 3.54); n=60; HADS depression 0-21 Top=High is poor outcome; Comments: Baseline acupuncture: 7.1 (3.16). Baseline control: 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 ; Baseline details: Reported age, BMI, gender and baseline values

of outcomes; Group 1 Number missing: 16, Reason: 16 lost to follow up; Group 2 Number missing: 26, Reason: 26 lost to follow up Protocol outcome 3: Serious adverse events at ≤ 3 months - Actual outcome: Adverse responses to treatment at 12 weeks; Group 1: 0/60, Group 2: 0/60 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: Reported age, BMI, gender and baseline values of outcomes; Group 1 Number missing: 16, Reason: 16 lost to follow up; Group 2 Number missing: 26, Reason: 26 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; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at > 3 months

Study (subsidiary papers)	Witt 2005[155] (Brinkhaus 2007[12])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=300)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: Intervention for 8 weeks, follow up for 52 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Diagnosis with osteoarthritis to the American College of Rheumatology criteria with documented radiological alterations in the knee joint of grade 2 or more according to Kellgren Lawrence criteria
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Aged 50-75 years; had been diagnosed with osteoarthritis according to the American College of Rheumatology criteria; had documented radiological alterations in the knee joint of grade 2 or more according to Kellgren Lawrence criteria; had an average pain intensity of 40 or more on a 100mm visual analogue scale in the 7 days before baseline assessment, and if they gave written informed consent
Exclusion criteria	Pain in the knee caused by inflammatory, malignant or autoimmune disease; or other reasons for pain in the knee, such as serious valgus-defective or varus-defective position; if they had had knee surgery; arthroscopy of the affected knee in the past year; chondroprotective or intra-articualr injection in the past 4 months; systemic corticoid treatment or beginning of a new treatment for osteoarthritis in the past 4 weeks; local antiphlogistic treatment; acupuncture treatment during the past 12 months; physiotherapy or other treatments for osteoarthritis knee pain (with the exception of non-steroidal anti-inflammatory drugs) during the previous 4 weeks; application for pension or disability benefits; serious acute or chronic organic disease or mental disorder; pregnancy or breastfeeding; blood coagulation disorders or coagulation-inhibiting medication other than aspirin
Recruitment/selection of patients	Most people were recruited through reports in local newspapers, and few spontaneously contacted trial centres
Age, gender and ethnicity	Age - Mean (SD): 64.0 (6.5). Gender (M:F): 99:195. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee

Extra comments	Severity: Kellgren Lawrence grade 0-4, median grade 3 Duration of symptoms (mean [SD]): 9.2 (7.9) years
Indirectness of population	No indirectness
Interventions	(n=150) Intervention 1: Acupuncture/dry needling - Acupuncture. Acupuncture consisting of 12 sessions of 30 minute duration administered over 8 weeks (usually 2 sessions per week for the first 4 weeks, then 1 session per week for the remaining 4 weeks). For people with bilateral osteoarthritis in the acupuncture group, both knees were needled with at least eight out of ten proposed points (at least 16 needles altogether), whereas for people with unilateral osteoarthritis, the physician could choose a unilateral or bilateral approach. Acupuncture treatment was semi-standardised, with all people being treated by local and distant points including at least six of the following local points: stomach 34, 35, 36; spleen 9, 10; bladder 40; kidney 10; gall bladder 33, 34; liver 8; extraordinary points Heding and Xiyan. Additionally, at least two of the distant points including the following: spleen 4, 5, 6; stomach 6; bladder 20, 57, 58, 60, 62; kidney 3. Sterile one-time needles were used, but physicians were able to choose the needle length and diameter. People were instructed to achieve de qi if possible, with the needles being stimulated at least once during each session. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=76) Intervention 2: Sham acupuncture. Minimal acupuncture using the same methods as standard acupuncture, but only using superficial insertion of fine needles (20-40mm in length) at predefined, distant non-acupuncture points that were not in the area of the knee, with selection of at least 8 out of the 10 points. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=74) Intervention 3: No intervention - Acupuncture compared to no treatment. Waiting list control. Duration 8 weeks. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable Comments: The waiting list was only for 8 weeks with people having acupuncture at some point after this. Therefore, this comparison will only be examined up to 8 weeks.
Funding	Academic or government funding (Study activities at the Institute for Social Medicine, Epidemiology and Health Economics, Berlin were funded by the following social health insurance funds: Techniker Krankenkasse, BKK Aktir, Betriebskrankenkasse de

	Allianz Gesellschaften, Bertelsmann BKK, Bosch BKK, BKK BMW, DaimlerChrysler BKK, BKK Deutsche Bank, Ford Betriebskrankenkasse, BKK Hoechst, Hypo Vereinsbank Betriebskrankenkasse, Innungskrankenkasse, Handelskrankenkasse, Innungskrankenkasse Hamburg. Study activities at the Centre for Complementary Medicine Research, Munich were funded by the following social health insurance funds: Deutsche Angestellten-Krankenkasse; Barmer Ersatzkasse; Kaufmannische Krankenkasse, Hamburg-Munchener Krankenkasse; Hanseatische Krankenkasse; Grnunder Ersatzkasser; HZK Krankenkasse fur Bau-und Holzberufe; Bruhler Ersatzkasse; Krankenkasse Eintracts Heusenstamm; and Buchbrucker Krankenkasse)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus SHAM ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical health at 8 weeks; Group 1: mean 36.2 (SD 7.4); n=150, Group 2: mean 33.1 (SD 7); n=76; SF-36 physical health 0-100 Top=High is good outcome; Comments: Reported final values and standard error. Reported acupuncture: 36.2 (0.6). Reported sham: 33.1 (0.8). Baseline acupuncture: 30.0 (7.4). Baseline sham: 29.2 (8.2).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 3, Reason: 1 without baseline and acupuncture, 2 lost to follow up
- Actual outcome: SF-36 mental health at 8 weeks; Group 1: mean 53.6 (SD 8.6); n=150, Group 2: mean 51.9 (SD 8.7); n=76; SF-36 mental health 0-100 Top=High is good outcome; Comments: Reported final values and standard error. Reported acupuncture: 53.6 (0.7). Reported sham: 51.9 (1.0). Baseline acupuncture: 51.8 (12.1). Baseline sham: 51.1 (11.6).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 3, Reason: 1 without baseline and acupuncture, 2 lost to follow up

Protocol outcome 2: Health-related quality of life at > 3 months
- Actual outcome: SF-36 physical health at 52 weeks; Group 1: mean 35 (SD 10); n=150, Group 2: mean 32.8 (SD 9.5); n=76; SF-36 physical health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 30.0 (7.4). Baseline sham: 29.2 (8.2).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and

acupuncture, 4 lost to follow up
- Actual outcome: SF-36 mental health at 52 weeks; Group 1: mean 52.9 (SD 11); n=150, Group 2: mean 51.1 (SD 11.7); n=76; SF-36 mental health 0-100 Top=High is good outcome; Comments: Baseline acupuncture: 51.8 (12.1). Baseline sham: 51.1 (11.6).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and acupuncture, 4 lost to follow up

Protocol outcome 3: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 24.4 (SD 17.2); n=150, Group 2: mean 33.2 (SD 17.4); n=74; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 24.4 (1.4). Reported sham: 33.2 (2.0). Reported waiting list: 44.9 (2.1). Baseline values not reported.
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 3, Reason: 1 without baseline and acupuncture, 2 lost to follow up

Protocol outcome 4: Pain at > 3 months
- Actual outcome: WOMAC pain at 52 weeks; Group 1: mean 30 (SD 23.5); n=150, Group 2: mean 33.5 (SD 21.3); n=76; WOMAC pain 0-100 Top=High is poor outcome; Comments: Baseline values not reported
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and acupuncture, 4 lost to follow up

Protocol outcome 5: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 27 (SD 17.2); n=150, Group 2: mean 35.8 (SD 17.4); n=76; WOMAC function 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 27.0 (1.4). Reported sham: 35.8 (2.0). Baseline values not reported.
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 3, Reason: 1 without baseline and acupuncture, 2 lost to follow up

Protocol outcome 6: Physical function at > 3 months
- Actual outcome: WOMAC function at 52 weeks; Group 1: mean 33 (SD 23); n=150, Group 2: mean 38.9 (SD 23.8); n=76; WOMAC function 0-100 Top=High is poor outcome; Comments: Baseline values not reported
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and acupuncture, 4 lost to follow up

Protocol outcome 7: Psychological distress at ≤ 3 months
- Actual outcome: Depression (ADS) at 8 weeks; Group 1: mean 47.9 (SD 9.8); n=150, Group 2: mean 48.3 (SD 9.6); n=76; Depression (ADS) 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 47.9 (0.8). Reported sham: 48.3 (1.1). Baseline acupuncture: 51.2 (10.0). Baseline sham: 51.3 (7.9).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 3, Reason: 1 without baseline and acupuncture, 2 lost to follow up

Protocol outcome 8: Psychological distress at > 3 months
- Actual outcome: Depression (ADS) at 52 weeks; Group 1: mean 48.6 (SD 10.2); n=150, Group 2: mean 49.8 (SD 10.1); n=76; Depression (ADS) 0-100 Top=High is poor outcome; Comments: Baseline acupuncture: 51.2 (10.0). Baseline sham: 51.3 (7.9).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and acupuncture, 4 lost to follow up

Protocol outcome 9: Serious adverse events at > 3 months
- Actual outcome: Serious adverse events at 26 weeks; Group 1: 3/150, Group 2: 2/76; Comments: One person in the minimal acupuncture group died from myocardial infarction
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 5, Reason: 1 without baseline and acupuncture, 4 lost to follow up

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE COMPARED TO 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 36.2 (SD 7.4); n=150, Group 2: mean 31.8 (SD 7.7); n=74; SF-36 physical health 0-100 Top=High is good outcome; Comments: Reported final values and standard error. Reported acupuncture: 36.2 (0.6). Reported waiting list: 31.8 (0.9). Baseline acupuncture: 30.0 (7.4). Baseline waiting list: 29.8 (7.9).
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 gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 7, Reason: 4 without baseline, 3 lost to follow up
- Actual outcome: SF-36 mental health at 8 weeks; Group 1: mean 53.6 (SD 8.6); n=150, Group 2: mean 50.7 (SD 8.6); n=74; SF-36 mental health 0-100 Top=High is good outcome; Comments: Reported final values and standard error. Reported acupuncture: 53.6 (0.7). Reported waiting list: 50.7 (1.0). Baseline acupuncture: 51.8 (12.1). Baseline waiting list: 50.6 (12.1).
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: Reported gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 7, Reason: 4 without baseline, 3 lost to follow up

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 8 weeks; Group 1: mean 24.4 (SD 17.2); n=150, Group 2: mean 44.9 (SD 18.1); n=74; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 24.4 (1.4). Reported sham: 33.2 (2.0). Reported waiting list: 44.9 (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 gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 7, Reason: 4 without baseline, 3 lost to follow up

Protocol outcome 3: Physical function at ≤ 3 months
- Actual outcome: WOMAC function at 8 weeks; Group 1: mean 27 (SD 17.2); n=150, Group 2: mean 50.4 (SD 18.1); n=74; WOMAC function 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 27.0 (1.4). Reported waiting list: 50.4 (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 gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 7, Reason: 4 without baseline, 3 lost to follow up

Protocol outcome 4: Psychological distress at ≤ 3 months - Actual outcome: Depression (ADS) at 8 weeks; Group 1: mean 47.9 (SD 9.8); n=150, Group 2: mean 49.4 (SD 9.5); n=74; Depression ADS 0-100 Top=High is poor outcome; Comments: Reported final values and standard error. Reported acupuncture: 47.9 (0.8). Reported waiting list: 49.4 (1.1). Baseline acupuncture: 51.2 (10.0). Baseline waiting list: 51.2 (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 gender, age, BMI, education, Kellgren Lawrence score, duration of disease, days with pain, osteoarthritis bilateral, previous treatment and baseline values of outcomes (apart from WOMAC subscales); Group 1 Number missing: 4, Reason: 1 without baseline and acupuncture, 3 lost to follow up; Group 2 Number missing: 7, Reason: 4 without baseline, 3 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

Study (subsidiary papers)	Witt 2006[156] (Martins 2014[91], Reinhold 2008[107])
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=712)
Countries and setting	Conducted in Germany; Setting: Outpatient follow up
Line of therapy	Unclear
Duration of study	Intervention + follow up: 3 months of acupuncture, 6 months of total follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis: Clinical diagnosis of osteoarthritis-associated pain in the knee or hip with disease duration of >6 months with radiologic evidence of osteoarthritis (osteophyte formation)
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Age at least 40 years; clinical diagnosis of osteoarthritis associated pain in the knee or hip with disease duration of >6 months; radiologic evidence of osteoarthritis (osteophyte formation) and at least 15 days with pain in the preceding 30 days
Exclusion criteria	Knee or hip pain due to inflammation or malignancy
Recruitment/selection of patients	People with insurance with one of the participating statutory health insurance funds who had contacted participating physicians and requested acupuncture or were referred by the physician for acupuncture if they thought it was a suitable treatment option
Age, gender and ethnicity	Age - Mean (SD): 61.2 (10.4). Gender (M:F): 251:381. Ethnicity: Not stated
Further population details	1. Age (≤/> 75 years): ≤ 75 years 2. Diagnosis: 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]): 5.3 (6.2) years
Indirectness of population	No indirectness
Interventions	(n=357) Intervention 1: Acupuncture/dry needling - Acupuncture. 15 acupuncture sessions during the first 3 months of the study (no acupuncture over the last 3 months). The number of needles and acupuncture points used were chosen at the physician's discretion. Only needle acupuncture was allowed (laser, electroacupuncture and moxibustion were not allowed). In addition, only manual needle stimulation was allowed. Duration 3 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture

	(n=355) Intervention 2: No intervention - Acupuncture compared to no treatment. No acupuncture intervention during the first 3 months of treatment. After 3 months, people received acupuncture by the same protocol (therefore, any evidence from after 3 months will not be included in the results). Duration 3 months. Concurrent medication/care: No additional information. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Not applicable
Funding	Academic or government funding (Supported by the following statutory health insurance funds in Germany: Techniker Krankenkasse, Betriebskrankenkasse (BKK) Aktiv, Bosch BKK, DaimlerChrysler BKK, Bertelsmann BKK, BKK BMW, Siemens-Betriebskrankenkasse, BKK Deutsche Bank, BKK Hoechst, HypoVereinsbank BKK, Ford BKK, Betriebskrankenkasse der Allianz Gesellschaften, Vereins- und Westbank BKK, Handelskrankenkasse, and Innungskrankenkasse Hamburg)

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE COMPARED TO NO TREATMENT

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: SF-36 physical component score at 12 weeks; Group 1: mean 38.8 (SD 9); n=322, Group 2: mean 31.2 (SD 8.8); n=310; SF-36 physical component score 0-100 Top=High is good outcome; Comments: Reported mean (SE). Reported acupuncture: 38.8 (0.5). Reported no treatment: 31.2 (0.5). Baseline acupuncture: 30.6 (8.6). Baseline no treatment: 30.6 (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 gender, age, years of education, site of osteoarthritis, evaluated joint, reason for participating in the study, disease duration and baseline values of outcomes; Group 1 Number missing: 49, Reason: 308 completed questionnaires at 3 months; Group 2 Number missing: 66, Reason: 289 completed questionnaires at 3 months
- Actual outcome: SF-36 mental component score at 12 weeks; Group 1: mean 51.1 (SD 9); n=322, Group 2: mean 49.4 (SD 8.8); n=355; SF-36 mental component score 0-100 Top=High is good outcome; Comments: Reported mean (SE). Reported acupuncture: 51.1 (0.5). Reported no treatment: 49.4 (0.5). Baseline acupuncture: 49.9 (12.2). Baseline no treatment: 49.0 (12.0).
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: Reported gender, age, years of education, site of osteoarthritis, evaluated joint, reason for participating in the study, disease duration and baseline values of outcomes; Group 1 Number missing: 49, Reason: 308 completed questionnaires at 3 months; Group 2 Number missing: 66, Reason: 289 completed questionnaires at 3 months

Protocol outcome 2: Pain at ≤ 3 months
- Actual outcome: WOMAC pain at 12 weeks; Group 1: mean 27.3 (SD 17.9); n=322, Group 2: mean 45.7 (SD 19.4); n=310; WOMAC pain 0-100 Top=High is poor outcome; Comments: Reported mean (SE). Reported acupuncture: 27.3 (1.0). Reported no treatment: 45.7 (1.1). Baseline acupuncture: 48.5 (23.2). Baseline no treatment: 48.0 (22.4).
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 gender, age, years of education, site of

osteoarthritis, evaluated joint, reason for participating in the study, disease duration and baseline values of outcomes; Group 1 Number missing: 57, Reason: 300 completed WOMAC data at 3 months; Group 2 Number missing: 76, Reason: 279 completed WOMAC data at 3 months Protocol outcome 3: Physical function at ≤ 3 months - Actual outcome: WOMAC function at 12 weeks; Group 1: mean 30.8 (SD 17.9); n=322, Group 2: mean 47.1 (SD 17.6); n=310; WOMAC function 0-100 Top=High is poor outcome; Comments: Reported mean (SE). Reported acupuncture: 30.8 (1.0). Reported no treatment: 47.1 (1.0). Baseline acupuncture: 47.3 (24.5). Baseline no treatment: 47.7 (24.6). 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 gender, age, years of education, site of osteoarthritis, evaluated joint, reason for participating in the study, disease duration and baseline values of outcomes; Group 1 Number missing: 57, Reason: 300 completed WOMAC data at 3 months; Group 2 Number missing: 76, Reason: 279 completed WOMAC data at 3 months	
Protocol outcomes not reported by the study	Health-related quality of life at > 3 months; Pain at > 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Study	**Zhang 2019[162]**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	1 (n=90)
Countries and setting	Conducted in China; Setting: Conducted at a hospital (First affiliated Hospital of college medicine, Henan University of Science and technology in Luoyang, Henan Province
Line of therapy	Unclear
Duration of study	Intervention time: 2 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Overall
Subgroup analysis within study	Not applicable
Inclusion criteria	Male or female, age of at least 45 years with KOA diagnoses according to the American College of Rheumatology criteria including radiographic evidence of at least one osteophyte at the tibiofemoral joint in one or both knees (Kellgren-Lawrence score 2 or 3); pain score of at least 3 points on a 10 point visual analogue scale for most days during the previous month; willingness to sign the consent form and be randomly assigned to either a treatment of a placebo group

Exclusion criteria	Patient has had an adverse reaction to acupuncture or is unwilling to accept acupuncture treatment; patient conforms to the inclusion criteria, but does not follow prescribed treatment, which decreases the curative effects of electroacupuncture so that it cannot be judged, or patient has incomplete information that may interfere with his/her ability to accurately judge the effects of his/her treatment; patient has accompanying severe cardiovascular, cerebral, hepatic, renal or hemopoietic diseases; patient has inflammatory arthritis such as rheumatoid arthritis, gouty arthritis, etcetera or other diseases that may affect the condition of the knees; patient is pregnant, attempting to become pregnant or lactating; and patient has a mental disease
Recruitment/selection of patients	Not reported
Age, gender and ethnicity	Age - Mean (SD): Acupuncture group: 55.9 (5.8); usual care group: 55.2 (6.0); Electroacupuncture group: 54.9 (6.2). Gender (M:F): 47/43. Ethnicity: Not reported
Further population details	1. Age (≤/> 75 years): Not stated / Unclear 2. Diagnosis: Diagnosis with imaging 3. Multimorbidity: Not stated / Unclear 4. Site of osteoarthritis: Knee
Extra comments	Severity of symptoms (number with Kellgren Lawrence Grade II or III): Acupuncture group 15/15; Usual care group 13/17; Electroacupuncture group 11/19 Duration of pain (mean, SD): Acupuncture group 6.6 (2.0) years; usual care group 6.2 (1.8) years; electroacupuncture group 6.5 (1.8) years
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: No intervention - Acupuncture plus additional treatment compared to additional treatment alone. Participants were treated by usual medical care, the pharmacy species conclude analgesics, Ds-ABC, NSAIDs or COX2 inhibitors. Suitable drugs were prescribed according to the patients' specific situation, including Diclofenac Sodium Enteric-coated tablets, Etoricoxib tablets, Meloxicam capsules, Relaxing tendons and invigorate blood capsules and Pain Ning capsules etc.. Duration 2 weeks. Concurrent medication/care: Patients who had previously taken drugs for activating blood circulation (Ds-ABC) and non-steroidal anti-inflammatory drugs (NSAIDs) or COX2-inhibitors were allowed to continue to take these medications during the study period, however they were acted to avoid physical therapy as much as possible to ensure that the results reflected the role of acupuncture or electroacupuncture as much as possible rather than other forms of treatment. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=30) Intervention 2: Acupuncture/dry needling - Acupuncture. According to the

	Traditional Chinese Medicine meridian theory of treating pain of KOA, patients in this group were treated with acupuncture on eight ipsilateral acupoints once a day. The sterile disposable Hwato needles adopted in this study were made in Suzhou, China. Patients were positioned on the bed, supported by two pillows under the knees, and instructed to assume a comfortable position and not move during the 30 minute stimulation period. After the local area had been disinfected, the needles (30-gauge with an outer diameter of 0.3mm and a length of 40 mm) would be inserted at a depth of 25 to 40 mm vertically through lifting and thrusting combined with twirling and rotating the needles, De Qi (the feeling of fullness, numbness, heaviness sourness or dull aching) sensation were achieved. Patients were also allowed to accept pharmacological therapies according to the illness needs. Participants had 10 sessions over a period of two weeks, each lasting 30 minutes. Duration 2 weeks. Concurrent medication/care: Patients who had previously taken drugs for activating blood circulation (Ds-ABC) and non-steroidal anti-inflammatory drugs (NSAIDs) or COX2-inhibitors were allowed to continue to take these medications during the study period, however they were acted to avoid physical therapy as much as possible to ensure that the results reflected the role of acupuncture or electroacupuncture as much as possible rather than other forms of treatment. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture (n=30) Intervention 3: Electroacupuncture. On the basis of treatment for the acupuncture group, direct current and dilatational wave was delivered with a medical Hwato Electronic acupuncture Treatment Instrument (SDZ-II, Suzhou, China), at 2/100 Hz frequency and 0.2 ms pulse width for 30 minutes. Neixiyan (EX-LE 5) was connected to Dubi (ST 35) and Yin lingquan (SP 9) connected to Ashi point with a pair of electrodes. The intensity was not prescribed equally, but patients in the acupuncture and electroacupuncture groups were encouraged to increase them along with the physical fitness. Participants had 10 sessions over a period of two weeks, each lasting 30 minutes. Duration 2 weeks. Concurrent medication/care: Patients who had previously taken drugs for activating blood circulation (Ds-ABC) and non-steroidal anti-inflammatory drugs (NSAIDs) or COX2-inhibitors were allowed to continue to take these medications during the study period, however they were acted to avoid physical therapy as much as possible to ensure that the results reflected the role of acupuncture or electroacupuncture as much as possible rather than other forms of treatment. Indirectness: No indirectness Further details: 1. Acupuncture or dry needling: Acupuncture
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: AQoL-SF36 Physical functioning at 2 weeks (end of treatment); Group 1: mean 693 (SD 116); n=30, Group 2: mean 640 (SD 96); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture: 198 (94); usual care: 192 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to physical health at 2 weeks (end of treatment); Group 1: mean 243 (SD 122); n=30, Group 2: mean 233 (SD 137); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 67 (88); usual care group: 60 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to emotional problems at 2 weeks (end of treatment); Group 1: mean 243 (SD 82); n=30, Group 2: mean 207 (SD 101); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group 83 (75); Usual care group 77 (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; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Energy/fatigue at 2 weeks (end of treatment); Group 1: mean 331 (SD 53); n=30, Group 2: mean 307 (SD 51); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 184 (46); Usual care group: 183 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Emotional wellbeing at 2 weeks (end of treatment); Group 1: mean 355 (SD 69); n=30, Group 2: mean 331 (SD 69); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 220 (68); Usual care group 227 (59)
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: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Social functioning at 2 weeks (end of treatment); Group 1: mean 154 (SD 29); n=30, Group 2: mean 127 (SD 25); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 55 (28); Usual care group: 51 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Pain at 2 weeks (end of treatment); Group 1: mean 217 (SD 46); n=30, Group 2: mean 207 (SD 46); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 106 (58); Usual care group 105 (51)
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: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 General health at 2 weeks (end of treatment); Group 1: mean 387 (SD 94); n=30, Group 2: mean 312 (SD 64); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Acupuncture group: 173 (57); Usual care group 159 (59)
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: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE PLUS ADDITIONAL TREATMENT COMPARED TO ADDITIONAL TREATMENT ALONE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: AQoL-SF36 Physical functioning at 2 weeks (end of treatment); Group 1: mean 718 (SD 124); n=30, Group 2: mean 640 (SD 96); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 193 (91); Usual care group: 192 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to physical health at 2 weeks (end of treatment); Group 1: mean 247 (SD 117); n=30, Group 2: mean 233 (SD 137); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 73 (83); Usual care group: 60 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to emotional problems at 2 weeks (end of treatment); Group 1: mean 237 (SD 72); n=30, Group 2: mean 207 (SD 101); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture groups: 87 (78); Usual care group: 77 (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; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Energy/fatigue at 2 weeks (end of treatment); Group 1: mean 337 (SD 56); n=30, Group 2: mean 307 (SD 51); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 187 (55); Usual care group: 183 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Emotional wellbeing at 2 weeks (end of treatment); Group 1: mean 371 (SD 69); n=30, Group 2: mean 331 (SD 69); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 217 (77); Usual care group: 227 (59)
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: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Social functioning at 2 weeks (end of treatment); Group 1: mean 172 (SD 30); n=30, Group 2: mean 127 (SD 25); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 55 (30); Usual care group: 51 (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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Pain at 2 weeks (end of treatment); Group 1: mean 223 (SD 43); n=30, Group 2: mean 207 (SD 46); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 106 (58); Usual care group: 105 (51)
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: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 General health at 2 weeks (end of treatment); Group 1: mean 407 (SD 81); n=30, Group 2: mean 312 (SD 64); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 166 (73); Usual care group: 159 (59)
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: 0; Group 2 Number missing: 0

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: ELECTROACUPUNCTURE versus ACUPUNCTURE

Protocol outcome 1: Health-related quality of life at ≤ 3 months
- Actual outcome: AQoL-SF36 Physical functioning at 2 weeks (end of treatment); Group 1: mean 718 (SD 124); n=30, Group 2: mean 693 (SD 116); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline score: Electroacupuncture group: 193 (91); Acupuncture group: 198 (94)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to physical health at 2 weeks (end of treatment); Group 1: mean 247 (SD 117); n=30, Group 2: mean 243 (SD 122); n=30; AQoL-SF36 30 Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 73 (83); Acupuncture group: 67 (88)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Role limitations due to emotional problems at 2 weeks (end of treatment); Group 1: mean 237 (SD 72); n=30, Group 2: mean 243 (SD 82); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: electroacupuncture group 87 (78); acupuncture group 83 (75)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Energy/fatigue at 2 weeks (end of treatment); Group 1: mean 337 (SD 56); n=30, Group 2: mean 331 (SD 53); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 187 (55); Acupuncture group: 184 (46)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Emotional wellbeing at 2 weeks (end of treatment); Group 1: mean 371 (SD 69); n=30, Group 2: mean 355 (SD 69); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 217 (77); Acupuncture group: 220 (68)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Social functioning at 2 weeks (end of treatment); Group 1: mean 172 (SD 30); n=30, Group 2: mean 154 (SD 29); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 55 (30); Acupuncture group: 55 (28)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 Pain at 2 weeks (end of treatment); Group 1: mean 223 (SD 43); n=30, Group 2: mean 217 (SD 46); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 106 (58); Acupuncture group: 106 (58)
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 ; Group 1 Number missing: 0; Group 2 Number missing: 0
- Actual outcome: AQoL-SF36 General health at 2 weeks (end of treatment); Group 1: mean 407 (SD 81); n=30, Group 2: mean 387 (SD 94); n=30; AQoL-SF36 Unclear Top=High is good outcome; Comments: Baseline scores: Electroacupuncture group: 166 (73); Acupuncture group: 173 (57)
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 ; 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; Pain at ≤ 3 months; Pain at > 3 months; Physical function at ≤ 3 months; Physical function at > 3 months; Psychological distress at ≤ 3 months; Psychological distress at > 3 months; Osteoarthritis flares at ≤ 3 months; Osteoarthritis flares at > 3 months; Serious adverse events at ≤ 3 months; Serious adverse events at > 3 months

Appendix E – Forest plots

E.1 Acupuncture/dry needling compared to sham acupuncture

Figure 2: Quality of life (EQ-5D, 5-15, high is good, final value) at ≤3 months

Study or Subgroup	Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Mean Difference IV, Fixed, 95% CI
Sanchez romero 2019	0.13	0.3508	31	31	0.13 [-0.56, 0.82]

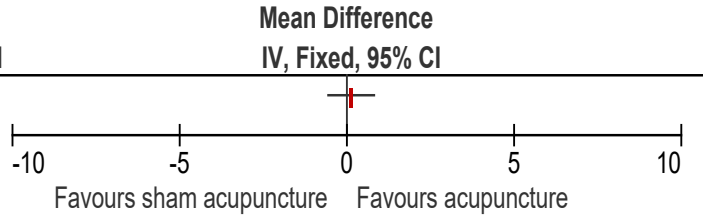


Figure 3: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change scores and final values) at ≤3 months

	Acupuncture			Sham acupuncture				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Chen 2013	3.01	10	104	4.3	9.03	109	16.9%	-1.29 [-3.85, 1.27]
Lin 2018	40.7	9.6	21	40.2	10.1	21	3.1%	0.50 [-5.46, 6.46]
Min 2009	8.5	13.1	40	5	13.3	38	3.2%	3.50 [-2.36, 9.36]
Scharf 2006	6.4	19.4	326	5.2	9.8	365	20.4%	1.20 [-1.13, 3.53]
Tu 2020	39.22	8.26	145	38.42	9.13	146	27.7%	0.80 [-1.20, 2.80]
Witt 2005	36.2	7.4	150	33.1	7	76	28.6%	3.10 [1.13, 5.07]
Total (95% CI)			**786**			**755**	**100.0%**	**1.26 [0.21, 2.32]**

Heterogeneity: Chi² = 7.99, df = 5 (P = 0.16); I² = 37%

Test for overall effect: Z = 2.35 (P = 0.02)

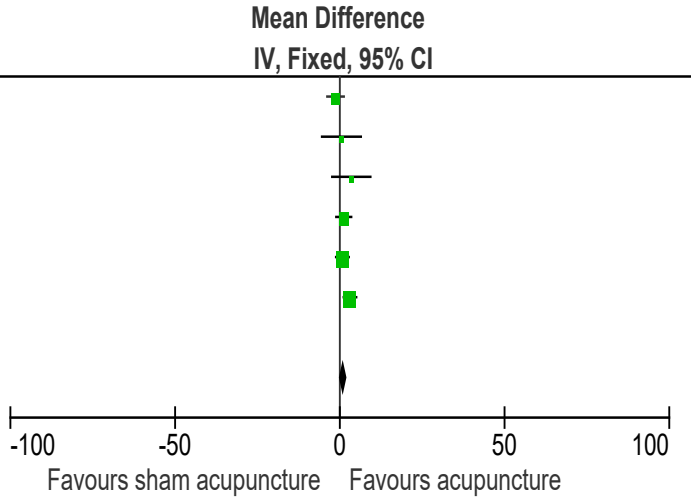


Figure 4: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change scores and final values) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Chen 2013	5.01	11.1	104	3.33	10.5	109	12.8%	1.68 [-1.22, 4.58]
Lin 2018	51.5	11.1	21	53.2	10.4	21	2.5%	-1.70 [-8.21, 4.81]
Min 2009	7	19.7	40	6.4	12.8	38	2.0%	0.60 [-6.74, 7.94]
Scharf 2006	2.1	11.1	326	3	11.7	365	37.2%	-0.90 [-2.60, 0.80]
Tu 2020	54.67	8.51	145	53.21	8.95	146	26.7%	1.46 [-0.55, 3.47]
Witt 2005	53.6	8.6	150	51.9	8.7	76	18.8%	1.70 [-0.69, 4.09]
Total (95% CI)			**786**			**755**	**100.0%**	**0.56 [-0.48, 1.60]**

Heterogeneity: Chi² = 5.51, df = 5 (P = 0.36); I² = 9%

Test for overall effect: Z = 1.05 (P = 0.29)

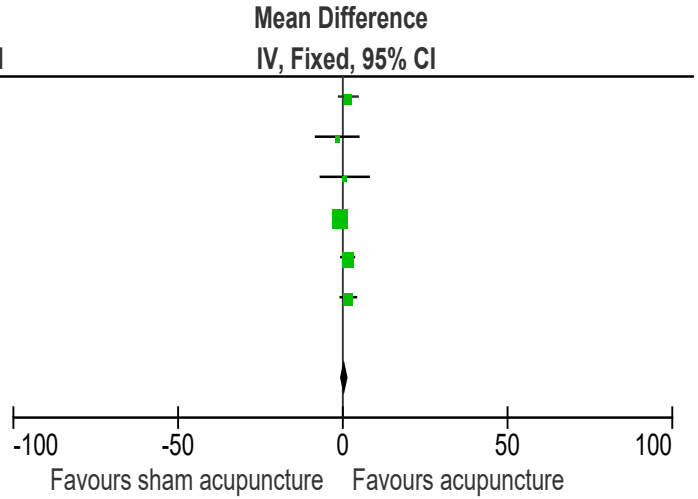


Figure 5: Quality of life (EQ-5D, 5-15, high is good, final value) at >3 months

Study or Subgroup	Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Mean Difference IV, Fixed, 95% CI
Sanchez romero 2019	0.15	0.3717	31	31	0.15 [-0.58, 0.88]

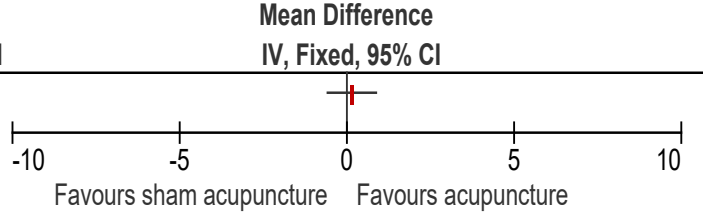


Figure 6: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score and final values) at >3 months

	Acupuncture			Sham acupuncture				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Lin 2018	38.2	9.2	21	37.6	9.3	21	4.4%	0.60 [-5.00, 6.20]
Scharf 2006	7	14.7	326	5.9	9.8	365	38.8%	1.10 [-0.79, 2.99]
Tu 2020	39.2	8.61	145	38.05	8.11	146	37.4%	1.15 [-0.77, 3.07]
Witt 2005	35	10	150	32.8	9.5	76	19.4%	2.20 [-0.47, 4.87]
Total (95% CI)			**642**			**608**	**100.0%**	**1.31 [0.13, 2.49]**

Heterogeneity: Chi² = 0.56, df = 3 (P = 0.90); I² = 0%

Test for overall effect: Z = 2.18 (P = 0.03)

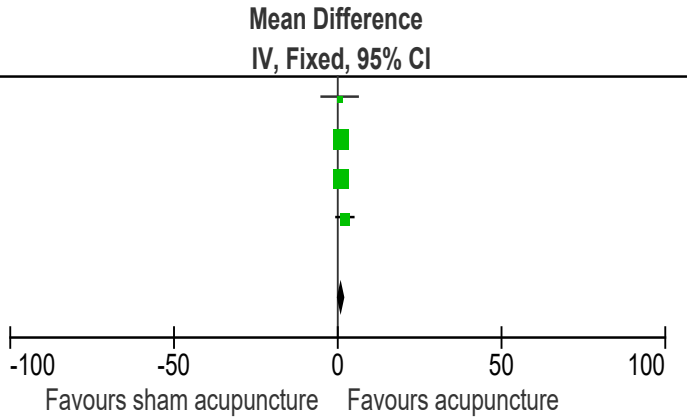


Figure 7: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change score and final values) at >3 months

	Acupuncture			Sham acupuncture				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Lin 2018	51.2	11.1	21	52.2	9.8	21	13.8%	-1.00 [-7.33, 5.33]
Scharf 2006	1.6	11.1	326	3.1	11.7	365	31.2%	-1.50 [-3.20, 0.20]
Tu 2020	54.92	8.22	145	51.34	9.37	146	29.9%	3.58 [1.56, 5.60]
Witt 2005	52.9	11	150	51.1	11.7	76	25.1%	1.80 [-1.37, 4.97]
Total (95% CI)			**642**			**608**	**100.0%**	**0.92 [-2.11, 3.95]**

Heterogeneity: Tau² = 6.92; Chi² = 14.94, df = 3 (P = 0.002); I² = 80%

Test for overall effect: Z = 0.59 (P = 0.55)

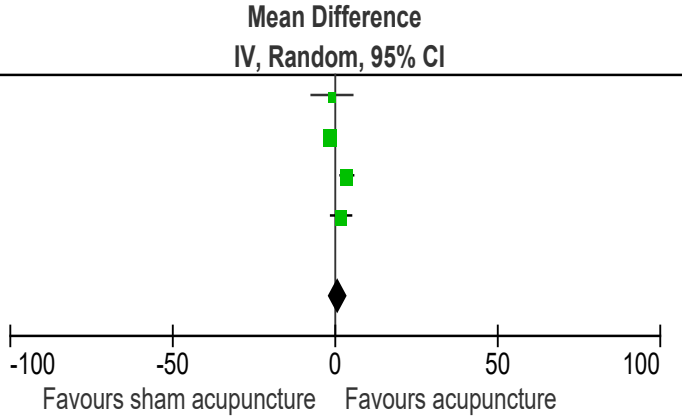


Figure 8: Pain (WOMAC [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Chen 2013	-2.83	4.6	104	-2.35	3.25	109	17.0%	-0.12 [-0.39, 0.15]
Foster 2007	-2.83	4	113	-3.02	3.6	115	18.3%	0.05 [-0.21, 0.31]
Lin 2018	-3.3	2.5	21	-3.1	2.6	21	3.4%	-0.08 [-0.68, 0.53]
Min 2009	-2	4.3	40	-0.9	3	38	6.2%	-0.29 [-0.74, 0.15]
Scharf 2006	-2.2	2.4	326	-2	2.44	365	55.2%	-0.08 [-0.23, 0.07]
Total (95% CI)			**604**			**648**	**100.0%**	**-0.08 [-0.19, 0.03]**

Heterogeneity: Chi² = 1.92, df = 4 (P = 0.75); I² = 0%

Test for overall effect: Z = 1.37 (P = 0.17)

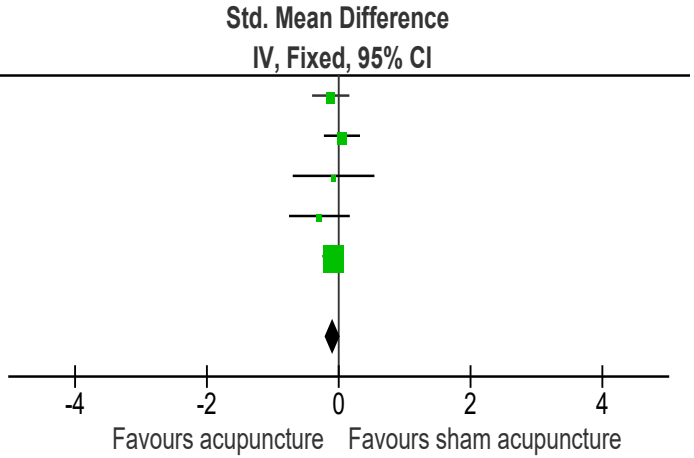


Figure 9: Pain (WOMAC, KSS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Lev-ari 2011	-0.2206	0.2706	28	27	6.5%	-0.22 [-0.75, 0.31]
Miller 2011	-0.2195	0.3135	21	20	4.8%	-0.22 [-0.83, 0.39]
Sanchez romero 2019	-0.3048	0.2556	31	31	7.3%	-0.30 [-0.81, 0.20]
Sanchez-romero 2018	-0.2384	0.4514	11	9	2.3%	-0.24 [-1.12, 0.65]
Takeda 1994	-0.3337	0.3187	20	20	4.7%	-0.33 [-0.96, 0.29]
Tu 2020	-0.3417	0.1181	145	146	34.1%	-0.34 [-0.57, -0.11]
White 2012	-0.021	0.165	74	73	17.5%	-0.02 [-0.34, 0.30]
Witt 2005	-0.5079	0.1441	150	74	22.9%	-0.51 [-0.79, -0.23]
Total (95% CI)			**480**	**400**	**100.0%**	**-0.30 [-0.44, -0.17]**

Heterogeneity: Chi² = 5.24, df = 7 (P = 0.63); I² = 0%
Test for overall effect: Z = 4.42 (P < 0.00001)

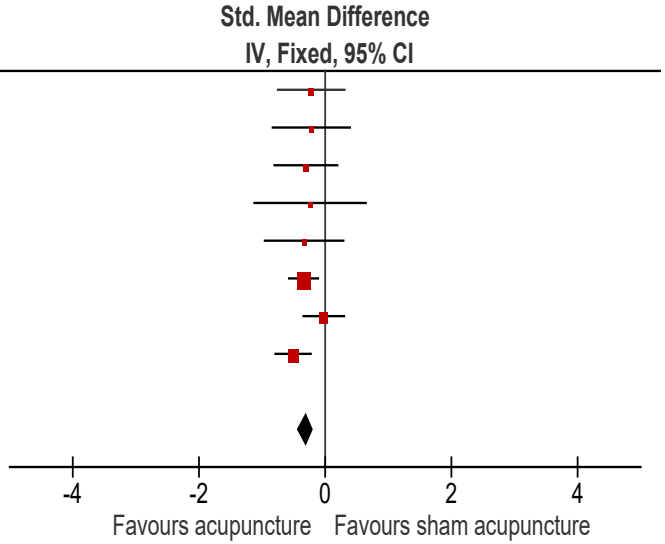


Figure 10: Pain (WOMAC [different scale ranges], high is poor, change scores) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Chen 2013	-1.91	3.77	104	-1.35	3.76	109	19.2%	-0.15 [-0.42, 0.12]
Foster 2007	-2.37	4.2	99	-2.82	4.1	105	18.4%	0.11 [-0.17, 0.38]
Scharf 2006	-2.3	2.5	326	-2.1	2.5	365	62.3%	-0.08 [-0.23, 0.07]
Total (95% CI)			**529**			**579**	**100.0%**	**-0.06 [-0.18, 0.06]**

Heterogeneity: Chi² = 1.92, df = 2 (P = 0.38); I² = 0%
Test for overall effect: Z = 0.97 (P = 0.33)

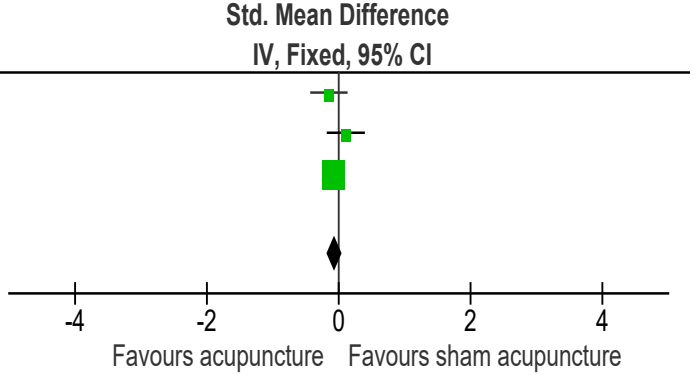


Figure 11: Pain (WOMAC [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Std. Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Lin 2018	-0.1486	0.3091	21	21	7.1%	-0.15 [-0.75, 0.46]
Sanchez romero 2019	0.0563	0.2541	31	31	10.5%	0.06 [-0.44, 0.55]
Tu 2020	-0.3255	0.118	145	146	48.5%	-0.33 [-0.56, -0.09]
Witt 2005	-0.1531	0.141	150	76	34.0%	-0.15 [-0.43, 0.12]
Total (95% CI)			**347**	**274**	**100.0%**	**-0.21 [-0.38, -0.05]**

Heterogeneity: Chi² = 2.26, df = 3 (P = 0.52); I² = 0%
Test for overall effect: Z = 2.61 (P = 0.009)

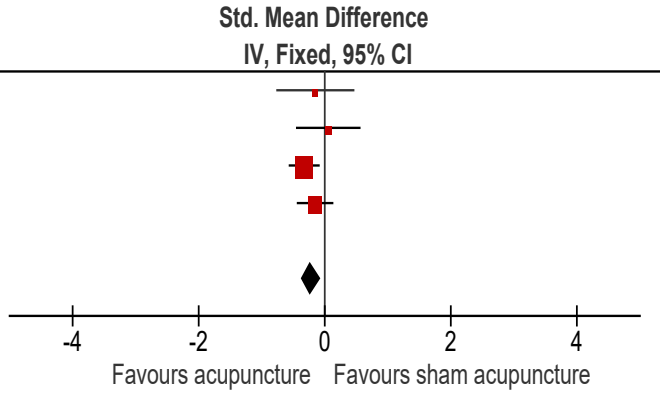


Figure 12: Physical function (WOMAC [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Chen 2013	-7.02	13.16	104	-6.93	10.97	109	17.1%	-0.01 [-0.28, 0.26]
Foster 2007	-8.18	11.5	113	-9.32	11.4	110	17.9%	0.10 [-0.16, 0.36]
Lin 2018	-11	7.3	21	-9.5	6.6	21	3.4%	-0.21 [-0.82, 0.40]
Min 2009	-6.8	12.7	40	-1.4	10.9	38	6.1%	-0.45 [-0.90, -0.00]
Scharf 2006	-2.1	2.4	326	-1.9	2.5	365	55.4%	-0.08 [-0.23, 0.07]
Total (95% CI)			**604**			**643**	**100.0%**	**-0.06 [-0.17, 0.05]**

Heterogeneity: Chi² = 4.77, df = 4 (P = 0.31); I² = 16%
Test for overall effect: Z = 1.12 (P = 0.26)

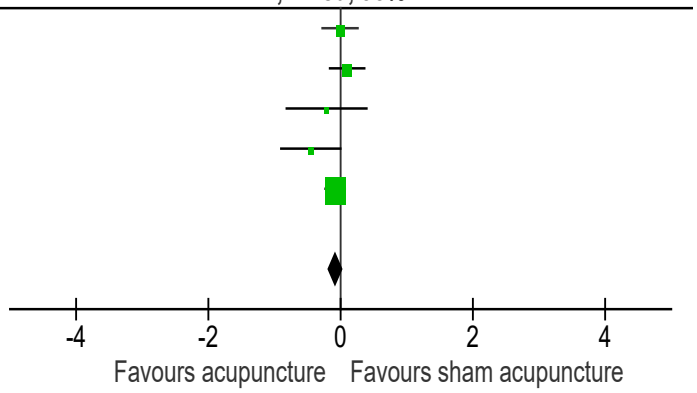


Figure 13: Physical function (WOMAC, KSS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Lev-ari 2011	-0.6193	0.2766	28	27	7.5%	-0.62 [-1.16, -0.08]
Miller 2011	-0.6151	0.3205	21	20	5.6%	-0.62 [-1.24, 0.01]
Sanchez romero 2019	-0.0371	0.254	31	31	8.9%	-0.04 [-0.53, 0.46]
Sanchez-romero 2018	0.3476	0.4536	11	9	2.8%	0.35 [-0.54, 1.24]
Takeda 1994	-0.2627	0.3177	20	20	5.7%	-0.26 [-0.89, 0.36]
Tu 2020	-0.1234	0.1174	145	146	41.5%	-0.12 [-0.35, 0.11]
Witt 2005	-0.5079	0.1428	150	76	28.1%	-0.51 [-0.79, -0.23]
Total (95% CI)			**406**	**329**	**100.0%**	**-0.28 [-0.43, -0.13]**

Heterogeneity: Chi² = 9.75, df = 6 (P = 0.14); I² = 38%
Test for overall effect: Z = 3.74 (P = 0.0002)

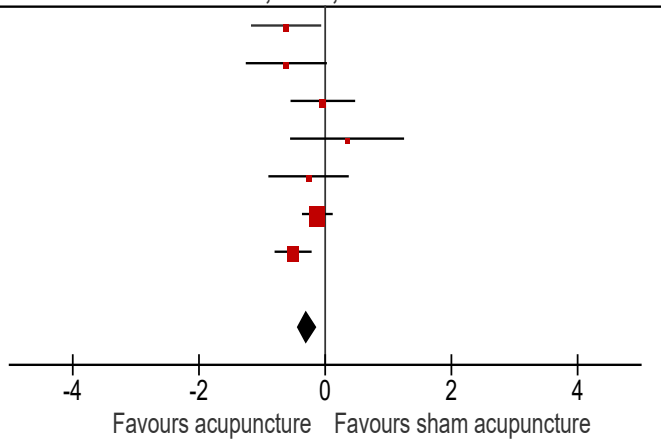


Figure 14: Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Chen 2013	-4.02	13.35	104	-3.74	10.76	109	19.3%	-0.02 [-0.29, 0.25]
Foster 2007	-6.61	13.8	100	-8.24	13.5	104	18.4%	0.12 [-0.16, 0.39]
Scharf 2006	-2.2	2.5	326	-2	2.5	365	62.3%	-0.08 [-0.23, 0.07]
Total (95% CI)			**530**			**578**	**100.0%**	**-0.03 [-0.15, 0.09]**

Heterogeneity: Chi² = 1.56, df = 2 (P = 0.46); I² = 0%
Test for overall effect: Z = 0.54 (P = 0.59)

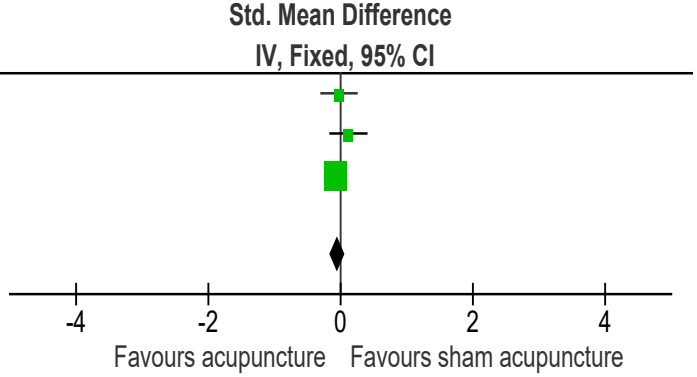


Figure 15: Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months

Study or Subgroup	Std. Mean Difference	SE	Acupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Lin 2018	-0.2204	0.3096	21	21	7.0%	-0.22 [-0.83, 0.39]
Sanchez romero 2019	-0.0425	0.254	31	31	10.4%	-0.04 [-0.54, 0.46]
Tu 2020	-0.2273	0.1176	145	146	48.7%	-0.23 [-0.46, 0.00]
Witt 2005	-0.2527	0.1413	150	76	33.8%	-0.25 [-0.53, 0.02]
Total (95% CI)			**347**	**274**	**100.0%**	**-0.22 [-0.38, -0.06]**

Heterogeneity: Chi² = 0.54, df = 3 (P = 0.91); I² = 0%
Test for overall effect: Z = 2.63 (P = 0.008)

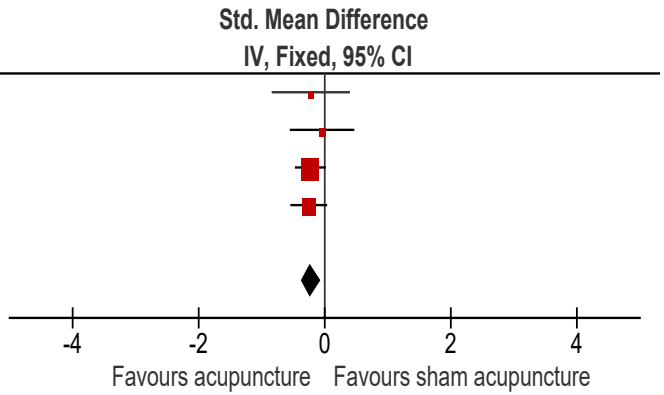


Figure 16: Psychological distress (Depression ADS, 0-100, high is poor, final value) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Witt 2005	47.9	9.8	150	48.3	9.6	76	-0.40 [-3.07, 2.27]

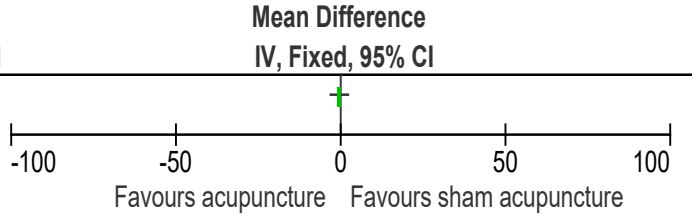


Figure 17: Psychological distress (Depression ADS, 0-100, high is poor, final value) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Witt 2005	48.6	10.2	150	49.8	10.1	76	-1.20 [-4.00, 1.60]

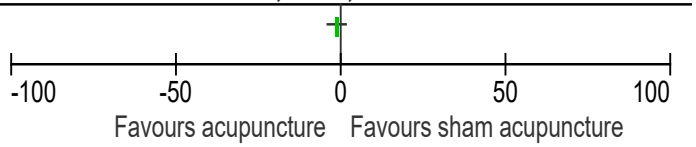


Figure 18: Serious adverse events at ≤3 months

Study or Subgroup	Acupuncture Events	Total	Sham acupuncture Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Fink 2001	0	33	0	32	50.9%	0.00 [-0.06, 0.06]
White 2012	22	74	4	73	49.1%	0.24 [0.13, 0.36]
Total (95% CI)		**107**		**105**	**100.0%**	**0.12 [-0.26, 0.50]**
Total events	22		4			

Heterogeneity: Tau² = 0.07; Chi² = 33.76, df = 1 (P < 0.00001); I² = 97%

Test for overall effect: Z = 0.62 (P = 0.54)

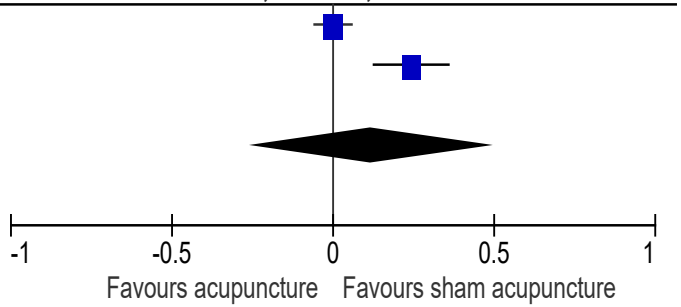
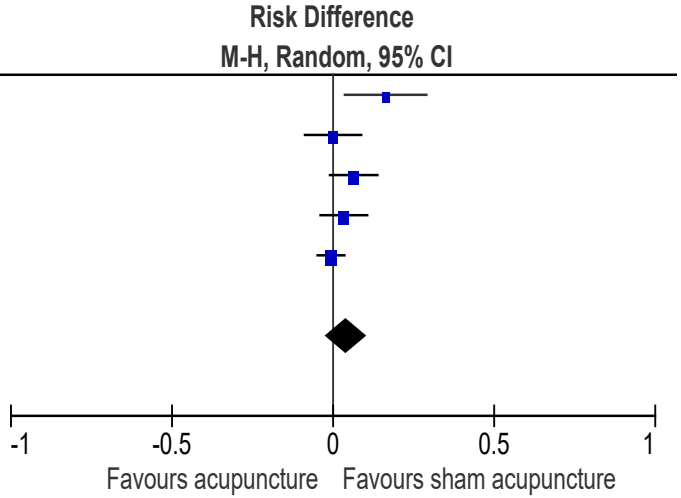


Figure 19: Serious adverse events at >3 months

Study or Subgroup	Acupuncture Events	Acupuncture Total	Sham acupuncture Events	Sham acupuncture Total	Weight	Risk Difference M-H, Random, 95% CI
Chen 2013	47	105	31	109	13.4%	0.16 [0.04, 0.29]
Lin 2018	0	21	0	21	18.7%	0.00 [-0.09, 0.09]
Scharf 2006	179	326	177	365	20.9%	0.06 [-0.01, 0.14]
Tu 2020	22	155	17	157	21.1%	0.03 [-0.04, 0.11]
Witt 2005	3	150	2	76	25.9%	-0.01 [-0.05, 0.04]
Total (95% CI)		**757**		**728**	**100.0%**	**0.04 [-0.02, 0.10]**
Total events	251		227			

Heterogeneity: Tau² = 0.00; Chi² = 14.40, df = 4 (P = 0.006); I² = 72%

Test for overall effect: Z = 1.26 (P = 0.21)



E.2 Acupuncture/dry needling compared to no treatment

Figure 20: Quality of life (EQ-5D, KOOS [different scale ranges], high is good, final values) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Foster 2007	0.663	0.24	117	0.639	0.27	116	78.1%	0.09 [-0.16, 0.35]
Kong 2018	50.7	18.8	34	47.4	16.7	12	11.9%	0.18 [-0.48, 0.84]
Lansdown 2009	0.71	0.26	15	0.66	0.25	15	10.0%	0.19 [-0.53, 0.91]
Total (95% CI)			**166**			**143**	**100.0%**	**0.11 [-0.11, 0.34]**

Heterogeneity: Chi² = 0.10, df = 2 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.98 (P = 0.33)

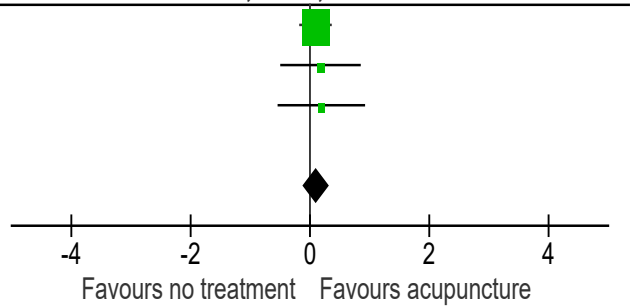


Figure 21: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Hinman 2014	40.7	9.6	64	39.5	10.7	69	28.3%	1.20 [-2.25, 4.65]
Witt 2005	36.2	7.4	150	31.8	7.7	74	34.5%	4.40 [2.28, 6.52]
Witt 2006	38.8	9	322	31.2	8.8	310	37.3%	7.60 [6.21, 8.99]
Total (95% CI)			**536**			**453**	**100.0%**	**4.69 [1.27, 8.11]**

Heterogeneity: Tau² = 7.67; Chi² = 14.71, df = 2 (P = 0.0006); I² = 86%
Test for overall effect: Z = 2.69 (P = 0.007)

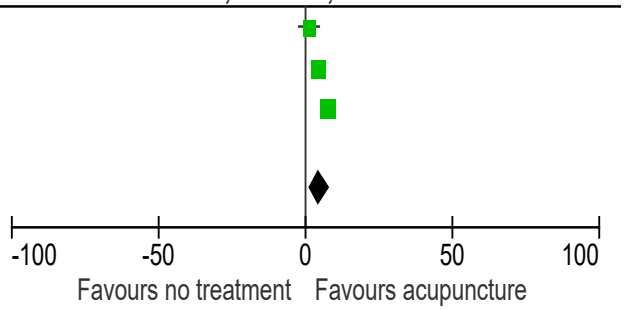


Figure 22: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months

	Acupuncture			No treatment				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Hinman 2014	51.5	11	64	55.8	9.1	69	28.2%	-4.30 [-7.75, -0.85]
Witt 2005	53.6	8.6	150	50.7	8.6	74	33.6%	2.90 [0.51, 5.29]
Witt 2006	51.1	9	322	49.4	8.8	355	38.3%	1.70 [0.36, 3.04]
Total (95% CI)			**536**			**498**	**100.0%**	**0.41 [-2.86, 3.69]**

Heterogeneity: Tau² = 6.83; Chi² = 12.20, df = 2 (P = 0.002); I² = 84%

Test for overall effect: Z = 0.25 (P = 0.80)

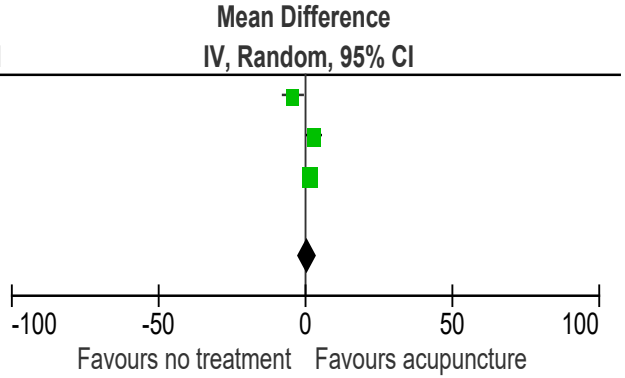


Figure 23: Quality of life (AQoL-SF-36 physical functioning, scale range unclear, high is good, final value) at ≤3 months

	Acupuncture			No treatment			Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	IV, Fixed, 95% CI
Zhang 2019	693	116	30	640	96	30	53.00 [-0.88, 106.88]

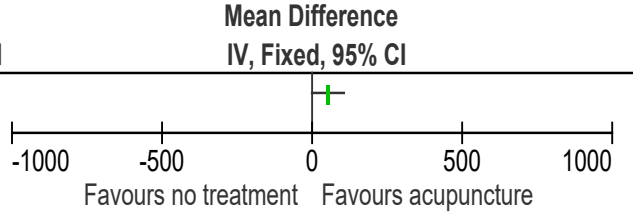


Figure 24: Quality of life (AQoL-SF-36 bodily pain, scale range unclear, high is good, final value) at ≤3 months

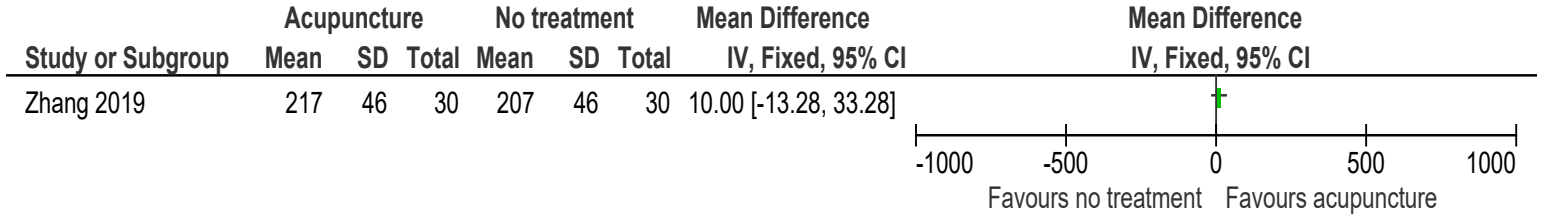


Figure 25: Quality of life (AQoL-SF-36 role physical, scale range unclear, high is good, final value) at ≤3 months

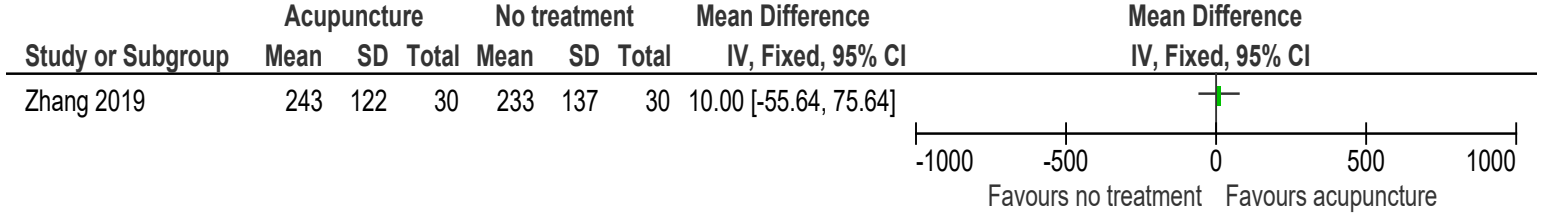


Figure 26: Quality of life (AQoL-SF-36 vitality, scale range unclear, high is good, final value) at ≤3 months

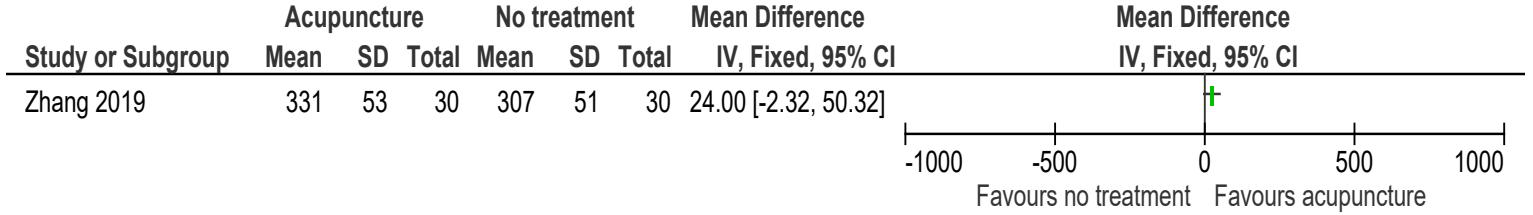


Figure 27: Quality of life (AQoL-SF-36 general health, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	387	94	30	312	64	30	75.00 [34.31, 115.69]

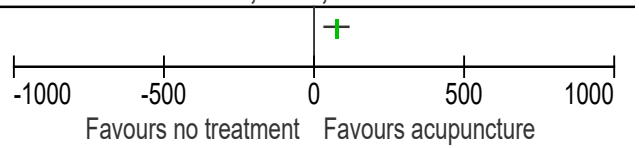


Figure 28: Quality of life (AQoL-SF-36 mental health, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	355	69	30	331	69	30	24.00 [-10.92, 58.92]

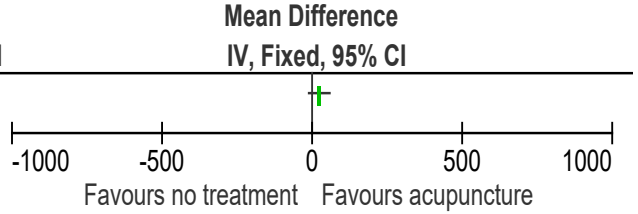


Figure 29: Quality of life (AQoL-SF-36 role emotional, scale range unclear, high is good, final value) at ≤3 months

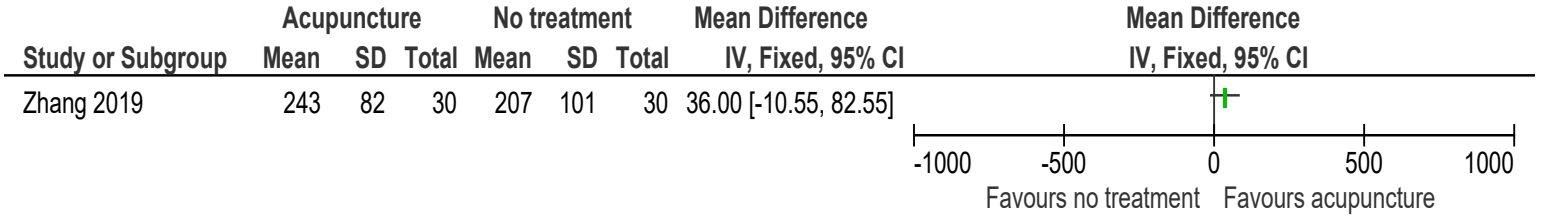


Figure 30: Quality of life (AQoL-SF-36 social functioning, scale range unclear, high is good, final value) at ≤3 months

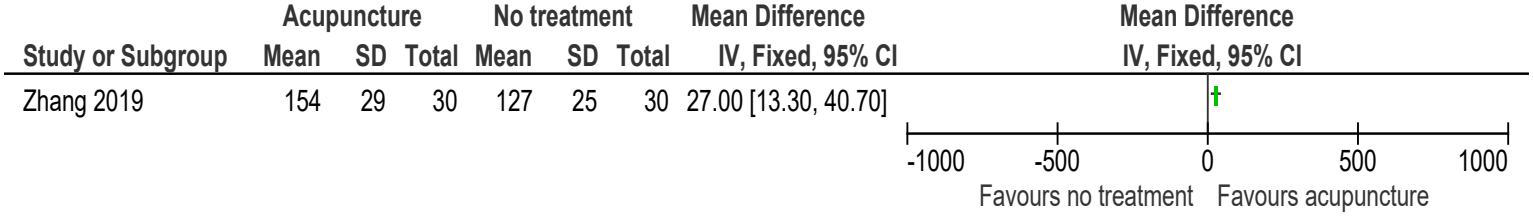


Figure 31: Quality of life (EQ-5D, -0.11-1, high is good, final values) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Foster 2007	0.621	0.3	117	0.616	0.3	116	80.2%	0.01 [-0.07, 0.08]
Lansdown 2009	0.66	0.24	15	0.63	0.19	15	19.8%	0.03 [-0.12, 0.18]
Total (95% CI)			**132**			**131**	**100.0%**	**0.01 [-0.06, 0.08]**

Heterogeneity: Chi² = 0.08, df = 1 (P = 0.78); I² = 0%
Test for overall effect: Z = 0.28 (P = 0.78)

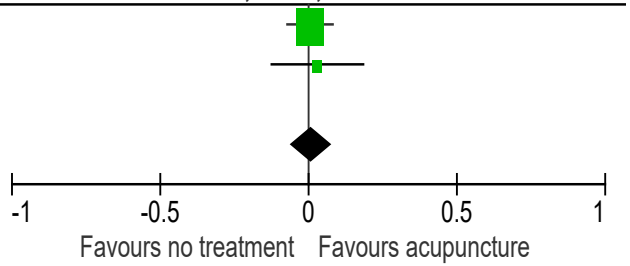


Figure 32: Quality of life (SF-12 physical component, 0-100, high is good, final value) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Hinman 2014	41.7	10.8	59	38.9	11.2	62	2.80 [-1.12, 6.72]

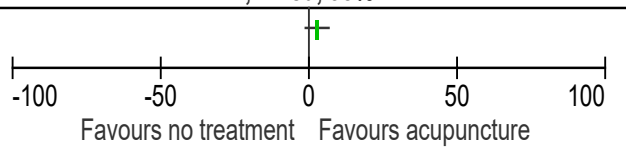


Figure 33: Quality of life (SF-12 mental component, 0-100, high is good, final value) at >3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Hinman 2014	51.5	11	59	54.4	10.2	62	-2.90 [-6.68, 0.88]

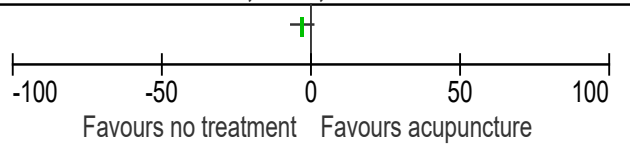


Figure 34: Pain (WOMAC, 0-20, high is poor, change score and final values) at ≤3 months

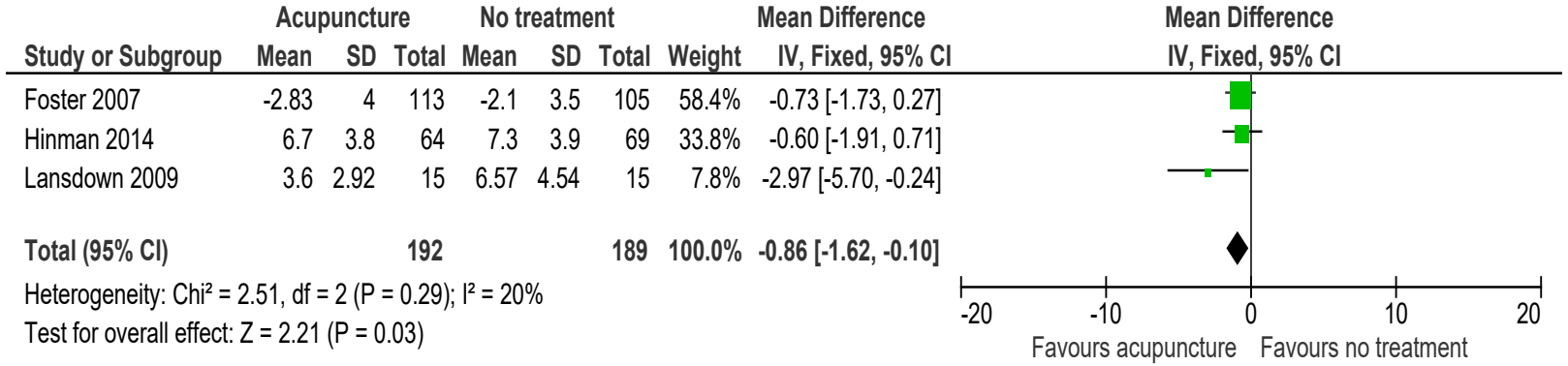


Figure 35: Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at ≤3 months

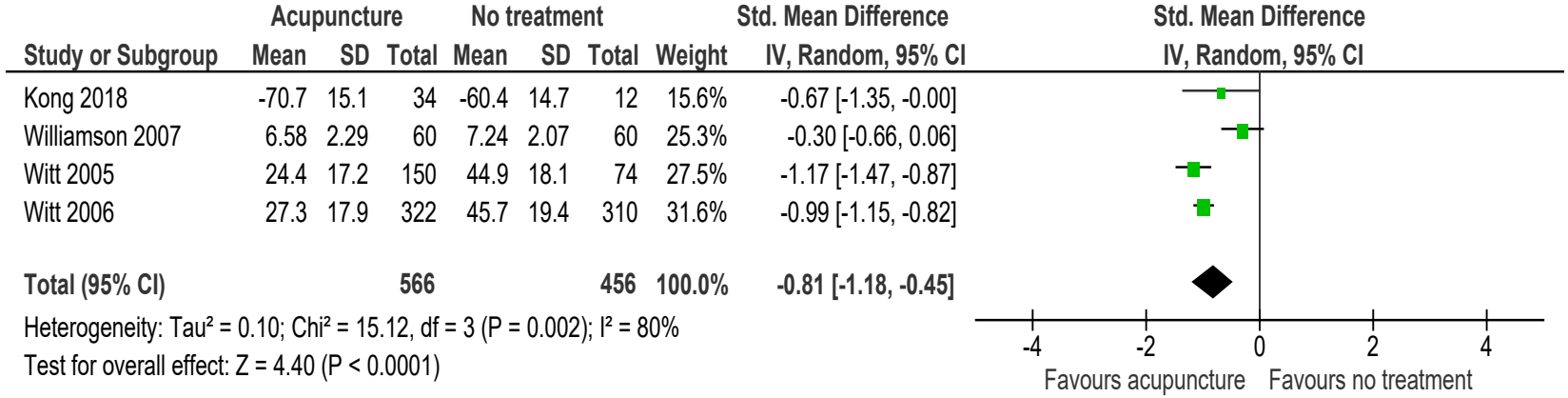


Figure 36: Pain (WOMAC, 0-20, high is poor, change score and final values) at >3 months

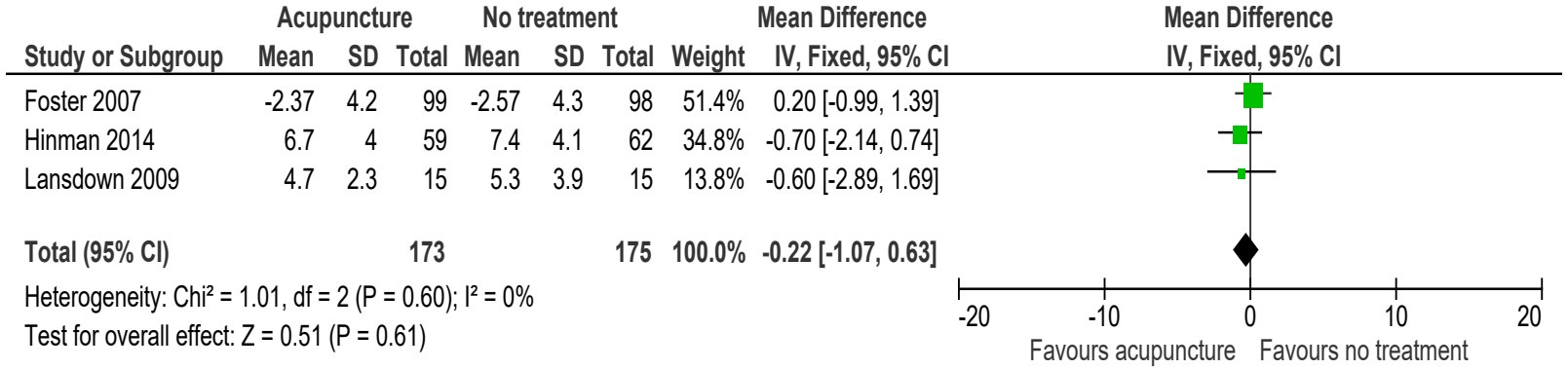


Figure 37: Physical function (WOMAC, 0-68, high is poor, change score and final values) at ≤3 months

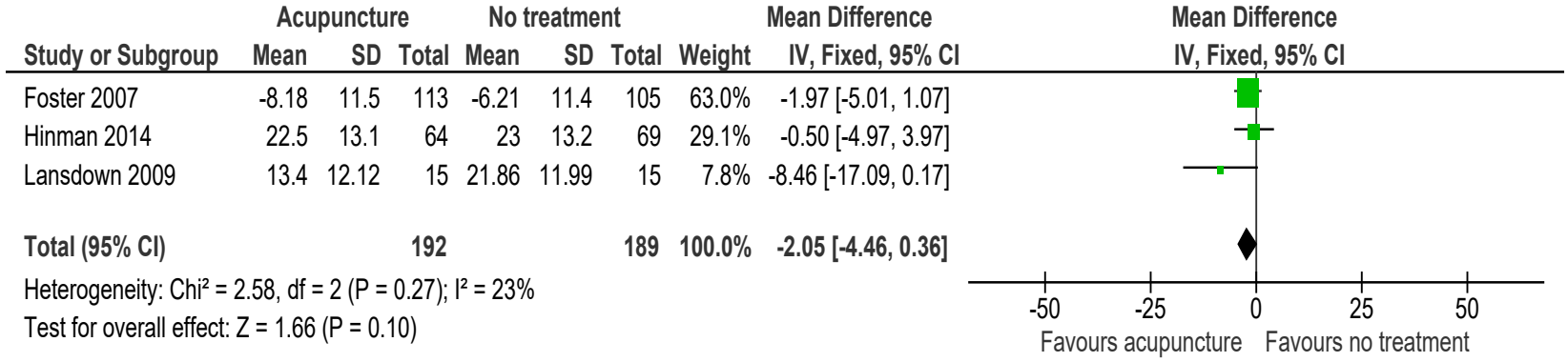


Figure 38: Physical function (KOOS, WOMAC, 0-100, high is poor, final values) at ≤3 months

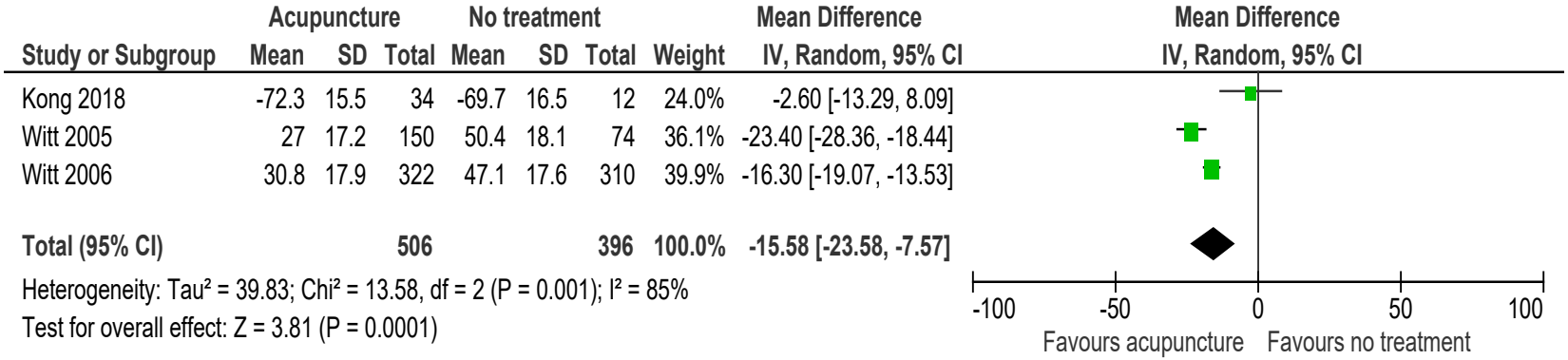


Figure 39: Physical function (WOMAC, 0-68, high is poor, change score and final values) at >3 months

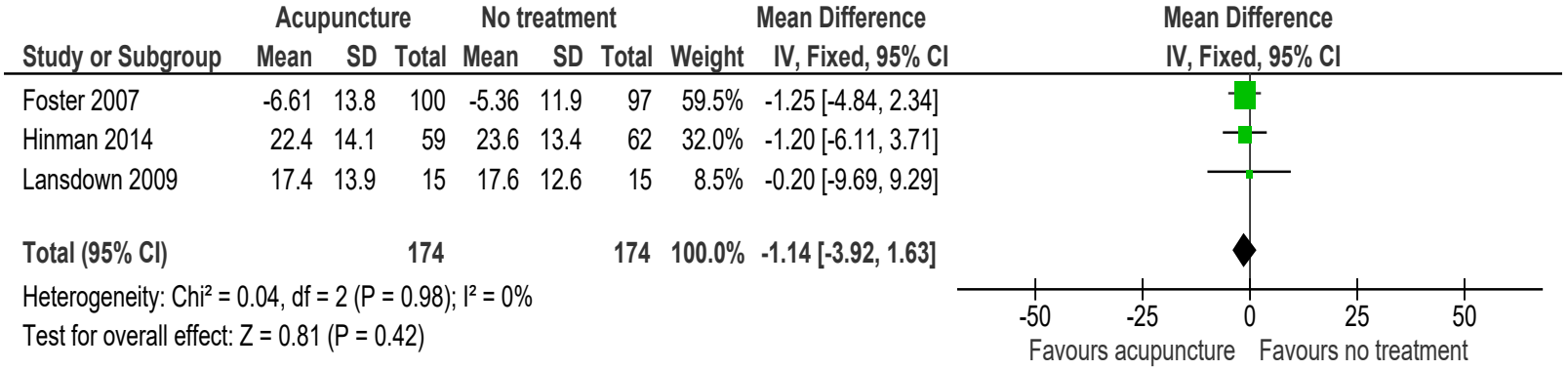


Figure 40: Psychological distress (HADS anxiety, 0-21, high is poor, final value) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Williamson 2007	6.88	4.15	60	6.54	3.93	60	0.34 [-1.11, 1.79]

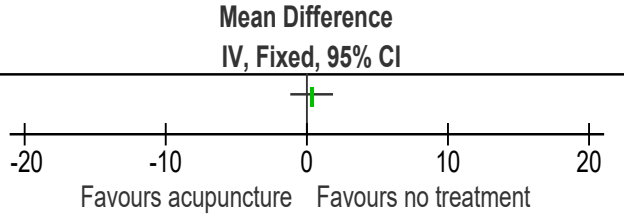


Figure 41: Psychological distress (HADS depression, depression ADS [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Acupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Williamson 2007	6.72	3.18	60	7.13	3.54	60	37.7%	-0.12 [-0.48, 0.24]
Witt 2005	47.9	9.8	150	49.4	9.5	74	62.3%	-0.15 [-0.43, 0.12]
Total (95% CI)			**210**			**134**	**100.0%**	**-0.14 [-0.36, 0.08]**

Heterogeneity: Chi² = 0.02, df = 1 (P = 0.89); I² = 0%

Test for overall effect: Z = 1.26 (P = 0.21)

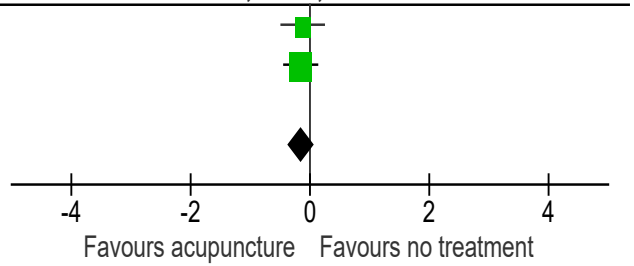


Figure 42: Serious adverse events at ≤3 months

Study or Subgroup	Acupuncture Events	Total	No treatment Events	Total	Risk Difference M-H, Fixed, 95% CI
Williamson 2007	0	60	0	60	0.00 [-0.03, 0.03]

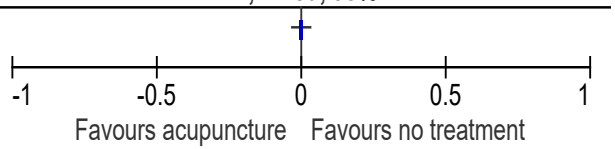
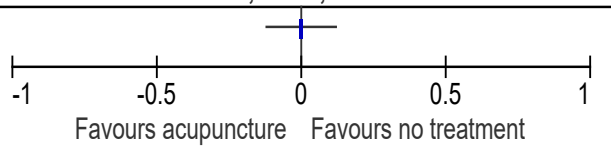


Figure 43: Serious adverse events at >3 months

Study or Subgroup	Acupuncture Events	Total	No treatment Events	Total	Risk Difference M-H, Fixed, 95% CI
Lansdown 2009	0	15	0	15	0.00 [-0.12, 0.12]



E.3 Electroacupuncture compared to acupuncture/dry needling

Figure 44: Quality of life (SF-12, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Wang 2020	62.21	16.9	28	60.87	18.42	30	1.34 [-7.75, 10.43]

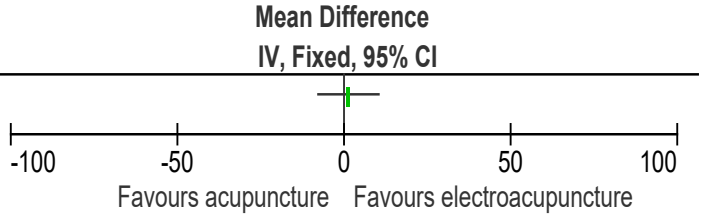


Figure 45: Quality of life (SF-12 physical health, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Tu 2020	38.97	8.33	151	39.22	8.26	145	-0.25 [-2.14, 1.64]

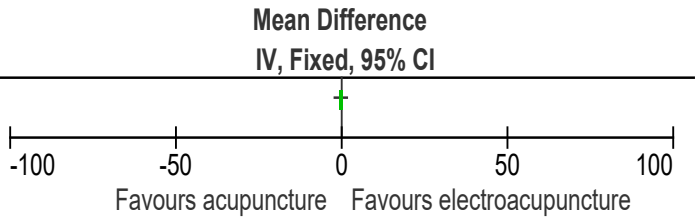


Figure 46: Quality of life (SF-12 mental health, 0-100, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Tu 2020	53.73	9.23	151	54.67	8.51	145	-0.94 [-2.96, 1.08]

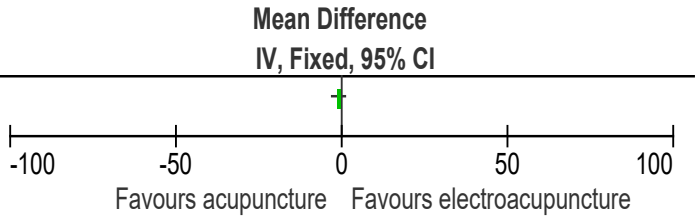


Figure 47: Quality of life (AQoL-SF 36 physical functioning, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	718	124	30	693	116	30	25.00 [-35.76, 85.76]

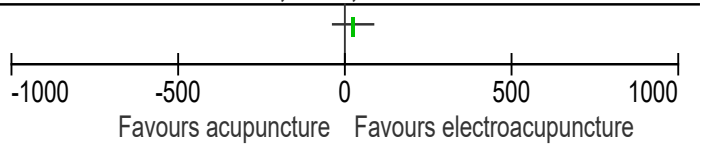


Figure 48: Quality of life (AQoL-SF 36 bodily pain, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	223	43	30	217	46	30	6.00 [-16.53, 28.53]

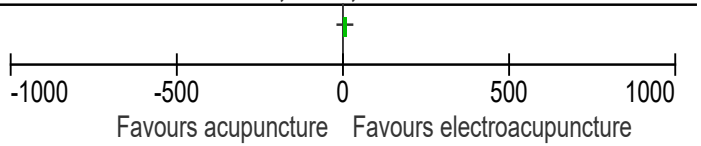


Figure 49: Quality of life (AQoL-SF 36 role physical, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	247	117	30	243	122	30	4.00 [-56.49, 64.49]

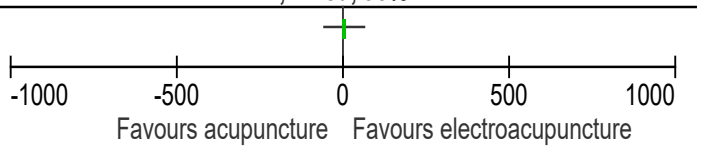


Figure 50: Quality of life (AQoL-SF 36 vitality, scale range unclear, high is good, final value) at ≤3 months

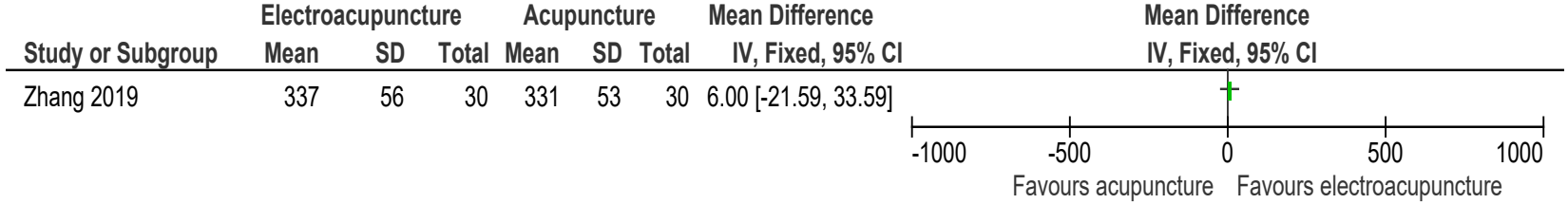


Figure 51: Quality of life (AQoL-SF 36 general health, scale range unclear, high is good, final value) at ≤3 months

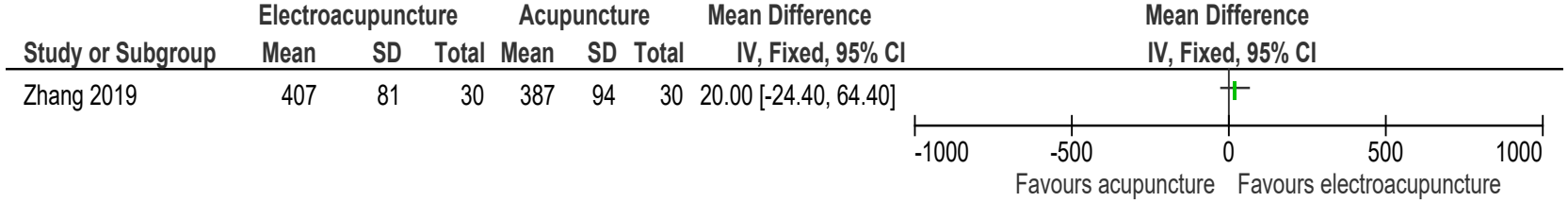


Figure 52: Quality of life (AQoL-SF 36 mental health, scale range unclear, high is good, final value) at ≤3 months

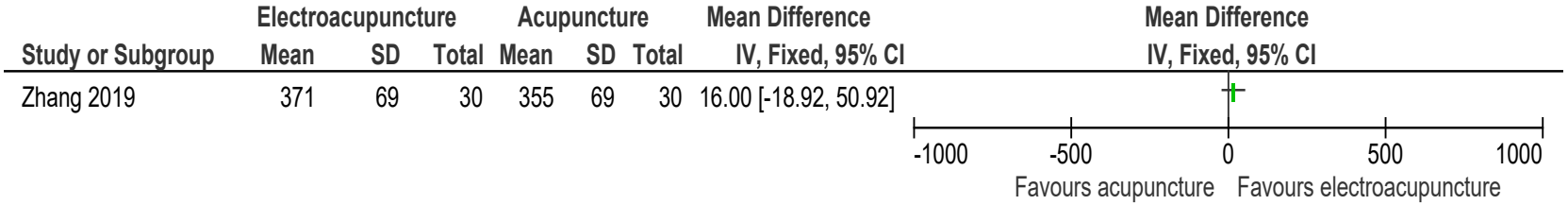


Figure 53: Quality of life (AQoL-SF 36 role emotional, scale range unclear, high is good, final value) at ≤3 months

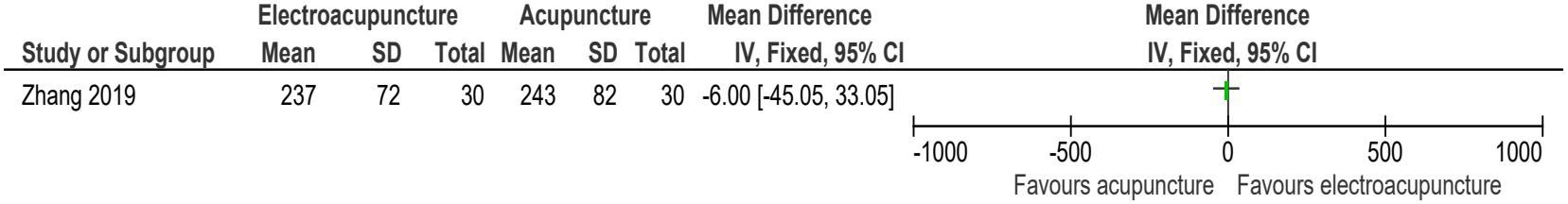


Figure 54: Quality of life (AQoL-SF 36 social functioning, scale range unclear, high is good, final value) at ≤3 months

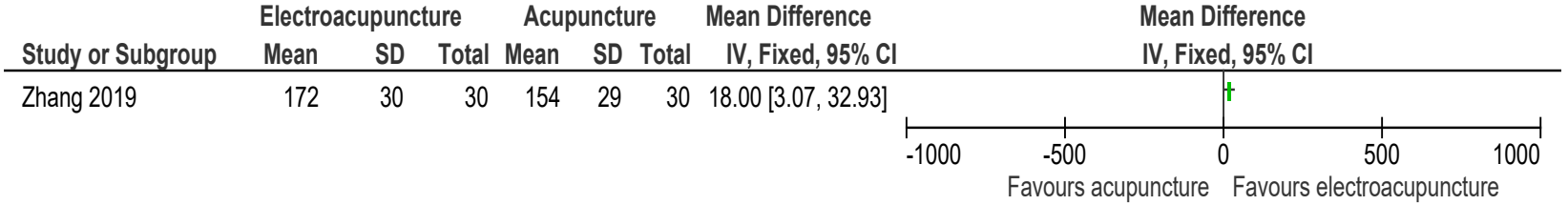


Figure 55: Quality of life (SF-12, 0-100, high is good, final value) at >3 months

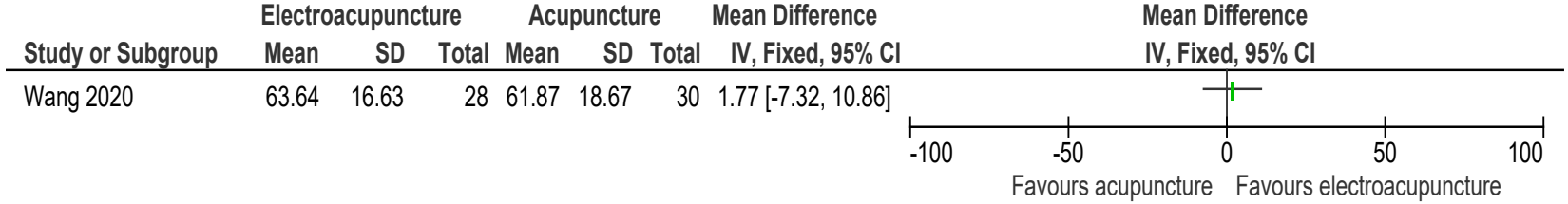


Figure 56: Quality of life (SF-12 physical health, 0-100, high is good, final value) at >3 months

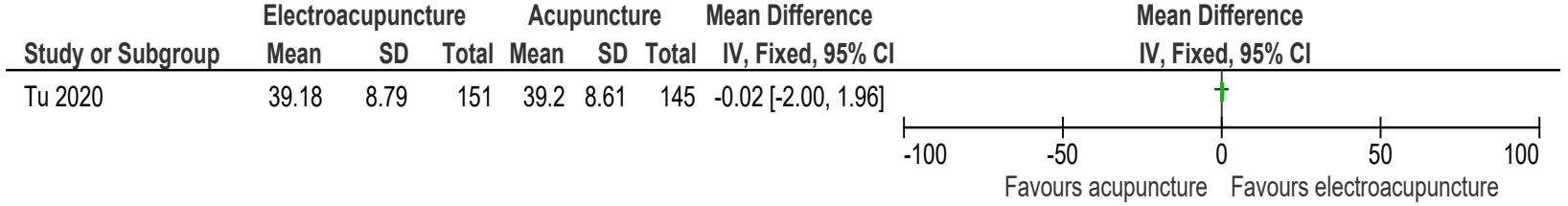


Figure 57: Quality of life (SF-12 mental health, 0-100, high is good, final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Tu 2020	54.11	8.04	151	54.92	8.22	145	-0.81 [-2.66, 1.04]

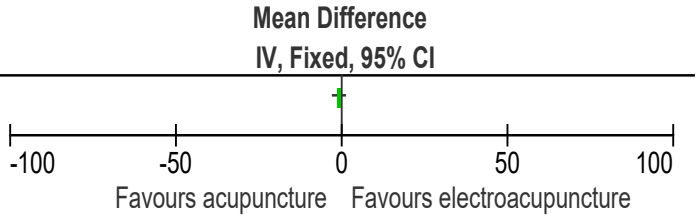


Figure 58: Pain (WOMAC, 0-20, high is poor, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Tu 2020	2.5	1.95	151	2.79	1.91	145	86.8%	-0.29 [-0.73, 0.15]
Wang 2020	2.54	2.08	28	3.43	2.31	30	13.2%	-0.89 [-2.02, 0.24]
Total (95% CI)			**179**			**175**	**100.0%**	**-0.37 [-0.78, 0.04]**

Heterogeneity: Chi² = 0.94, df = 1 (P = 0.33); I² = 0%
Test for overall effect: Z = 1.76 (P = 0.08)

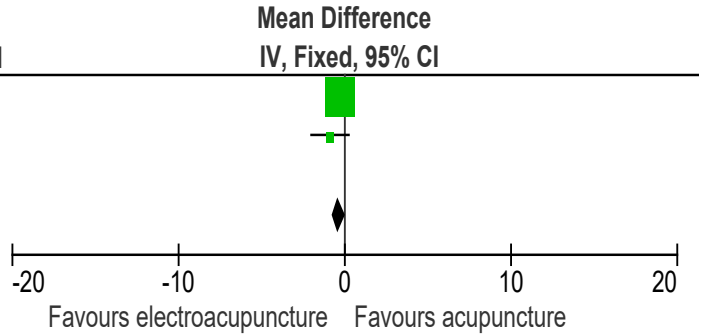


Figure 59: Pain (WOMAC, 0-20, high is poor, final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Tu 2020	2.79	2.41	151	3.16	2.03	145	86.9%	-0.37 [-0.88, 0.14]
Wang 2020	2.79	2.21	28	3.6	2.84	30	13.1%	-0.81 [-2.11, 0.49]
Total (95% CI)			**179**			**175**	**100.0%**	**-0.43 [-0.90, 0.04]**

Heterogeneity: Chi² = 0.38, df = 1 (P = 0.54); I² = 0%
Test for overall effect: Z = 1.77 (P = 0.08)

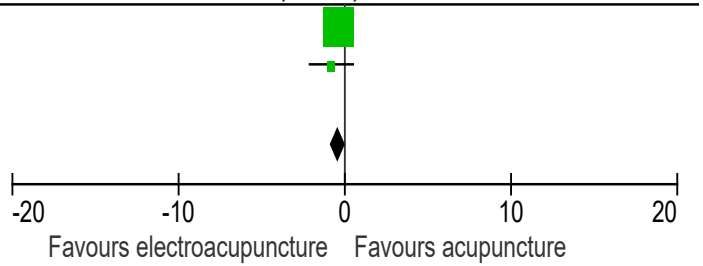


Figure 60: Physical function (WOMAC, 0-68, high is poor, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Tu 2020	9.26	7.03	151	10.82	7.32	145	82.9%	-1.56 [-3.20, 0.08]
Wang 2020	10.04	7.07	28	11.07	6.93	30	17.1%	-1.03 [-4.64, 2.58]
Total (95% CI)			**179**			**175**	**100.0%**	**-1.47 [-2.96, 0.02]**

Heterogeneity: Chi² = 0.07, df = 1 (P = 0.79); I² = 0%
Test for overall effect: Z = 1.93 (P = 0.05)

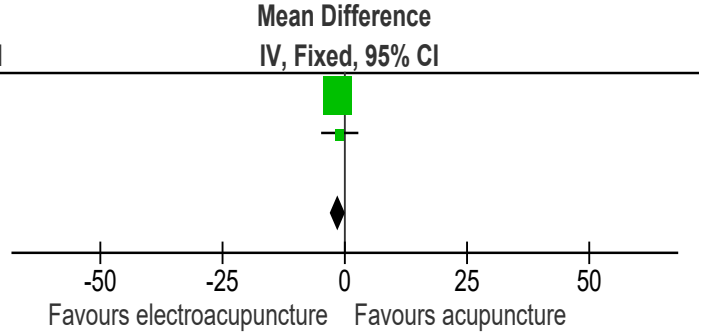


Figure 61: Physical function (WOMAC, 0-68, high is poor, final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Tu 2020	9.87	7.04	151	11.29	7.7	145	86.4%	-1.42 [-3.10, 0.26]
Wang 2020	9.39	7.83	28	12.33	8.62	30	13.6%	-2.94 [-7.17, 1.29]
Total (95% CI)			**179**			**175**	**100.0%**	**-1.63 [-3.19, -0.06]**

Heterogeneity: Chi² = 0.43, df = 1 (P = 0.51); I² = 0%
Test for overall effect: Z = 2.04 (P = 0.04)

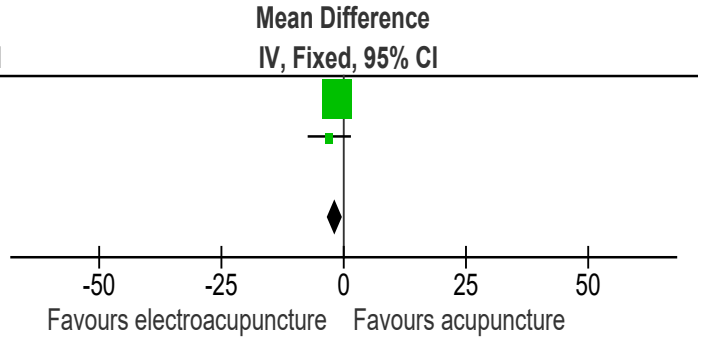
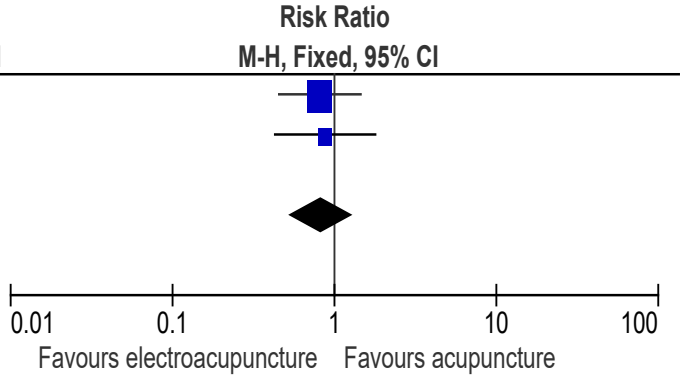


Figure 62: Serious adverse events at >3 months

Study or Subgroup	Electroacupuncture Events	Total	Acupuncture Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Tu 2020	18	156	22	155	67.5%	0.81 [0.45, 1.45]
Wang 2020	9	28	11	30	32.5%	0.88 [0.43, 1.79]
Total (95% CI)		**184**		**185**	**100.0%**	**0.83 [0.53, 1.31]**
Total events	27		33			

Heterogeneity: Chi² = 0.03, df = 1 (P = 0.87); I² = 0%
Test for overall effect: Z = 0.78 (P = 0.43)



E.4 Electroacupuncture compared to sham acupuncture

Figure 63: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score and final values) at ≤3 months

Study or Subgroup	Mean Difference	SE	Electroacupuncture Total	Sham acupuncture Total	Weight	Mean Difference IV, Random, 95% CI
Berman 2004	1.6	1.8439	169	169	18.4%	1.60 [-2.01, 5.21]
Mavrommatis 2012	10.6	1.3854	40	40	20.0%	10.60 [7.88, 13.32]
Suarez-almazor 2010	0.8	0.9761	153	302	21.2%	0.80 [-1.11, 2.71]
Tu 2020	0.55	1.0151	151	146	21.1%	0.55 [-1.44, 2.54]
Wang 2021	1.91	1.5367	30	30	19.5%	1.91 [-1.10, 4.92]
Total (95% CI)			**543**	**687**	**100.0%**	**3.07 [-0.55, 6.68]**

Heterogeneity: Tau² = 15.12; Chi² = 41.29, df = 4 (P < 0.00001); I² = 90%

Test for overall effect: Z = 1.66 (P = 0.10)

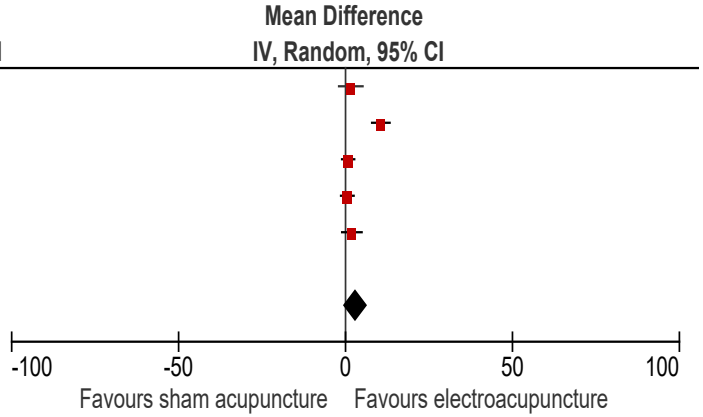


Figure 64: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Mavrommatis 2012	52.2	8	40	51.5	6.1	40	12.8%	0.70 [-2.42, 3.82]
Suarez-almazor 2010	54.1	8.2	153	53.2	8.9	302	46.2%	0.90 [-0.74, 2.54]
Tu 2020	38.97	8.33	151	38.42	9.13	146	31.5%	0.55 [-1.44, 2.54]
Wang 2021	43.86	7.6	30	43.5	6.74	30	9.4%	0.36 [-3.27, 3.99]
Total (95% CI)			**374**			**518**	**100.0%**	**0.71 [-0.40, 1.83]**

Heterogeneity: Chi² = 0.11, df = 3 (P = 0.99); I² = 0%
Test for overall effect: Z = 1.25 (P = 0.21)

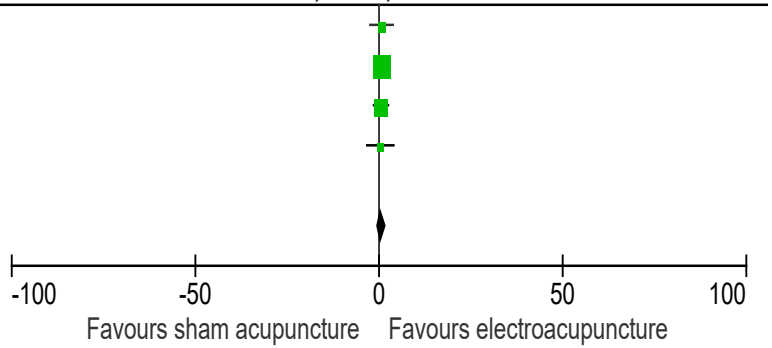


Figure 65: Quality of life (PLQC physical capability, 0-4, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Vas 2004	2.8	0.7	48	2.5	0.8	49	0.30 [0.00, 0.60]

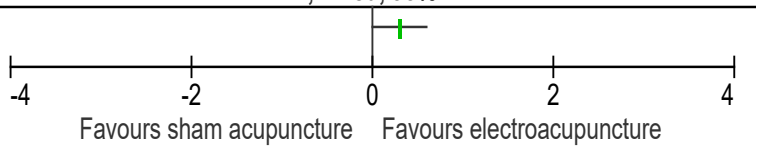


Figure 66: Quality of life (PLQC psychological functioning, 0-4, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Vas 2004	2.7	0.4	48	2.5	0.6	49	0.20 [-0.00, 0.40]

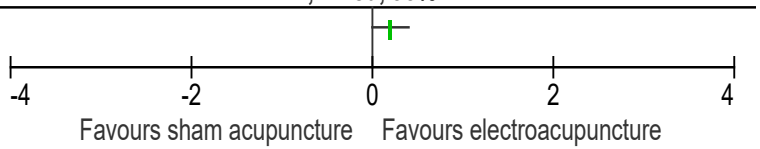


Figure 67: Quality of life (PLQC negative mood, 0-4, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Vas 2004	3.2	0.7	48	3.1	0.7	49	0.10 [-0.18, 0.38]

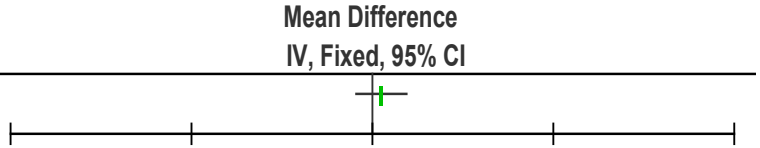


Figure 68: Quality of life (PLQC social functioning, 0-4, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Vas 2004	2.8	0.5	48	2.7	0.7	49	0.10 [-0.14, 0.34]

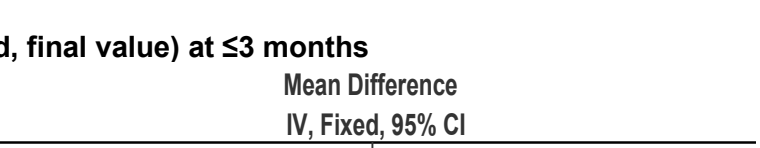


Figure 69: Quality of life (PLQC social wellbeing, 0-4, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Vas 2004	3.2	0.5	48	3.2	0.5	49	0.00 [-0.20, 0.20]

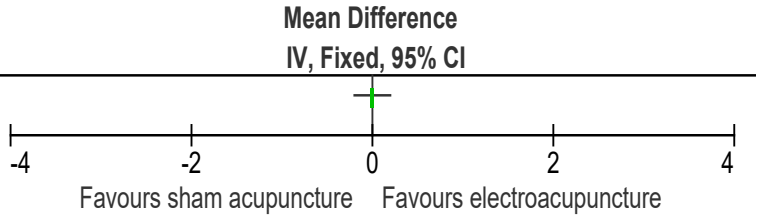


Figure 70: Quality of life (SF-36 physical component, SF-12 physical health, 0-100, high is good, change score and final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 2004	10.7	19.1	142	8.2	17.8	141	12.3%	2.50 [-1.80, 6.80]
Tu 2020	39.18	8.79	151	38.05	8.11	146	61.8%	1.13 [-0.79, 3.05]
Wang 2021	36.05	4.31	30	35.33	7.09	30	25.9%	0.72 [-2.25, 3.69]
Total (95% CI)			**323**			**317**	**100.0%**	**1.19 [-0.32, 2.70]**

Heterogeneity: Chi² = 0.46, df = 2 (P = 0.80); I² = 0%
Test for overall effect: Z = 1.55 (P = 0.12)

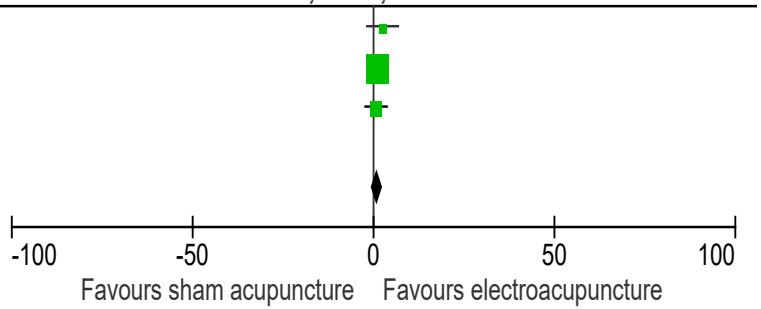


Figure 71: Quality of life (SF-36 mental component, SF-12 mental health, 0-100, high is good, final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Tu 2020	54.11	8.04	151	51.34	9.37	146	77.0%	2.77 [0.78, 4.76]
Wang 2021	43.86	7.6	30	43.5	6.74	30	23.0%	0.36 [-3.27, 3.99]
Total (95% CI)			**181**			**176**	**100.0%**	**2.21 [0.47, 3.96]**

Heterogeneity: Chi² = 1.30, df = 1 (P = 0.25); I² = 23%
Test for overall effect: Z = 2.49 (P = 0.01)

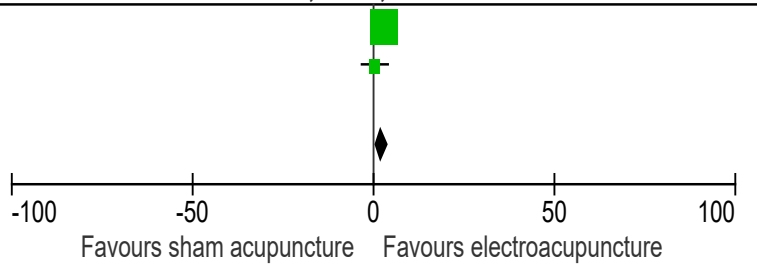


Figure 72: Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Berman 2004	-3.63	3.9	158	-2.68	4.14	157	33.5%	-0.24 [-0.46, -0.01]
Lv 2019	-2.9	1.23	217	1.19	1.21	75	33.1%	-3.33 [-3.71, -2.95]
Sangdee 2002	-5.97	4.66	97	-4.07	4.3	94	33.4%	-0.42 [-0.71, -0.13]
Total (95% CI)			**472**			**326**	**100.0%**	**-1.32 [-3.00, 0.36]**

Heterogeneity: Tau² = 2.18; Chi² = 203.47, df = 2 (P < 0.00001); I² = 99%
Test for overall effect: Z = 1.54 (P = 0.12)

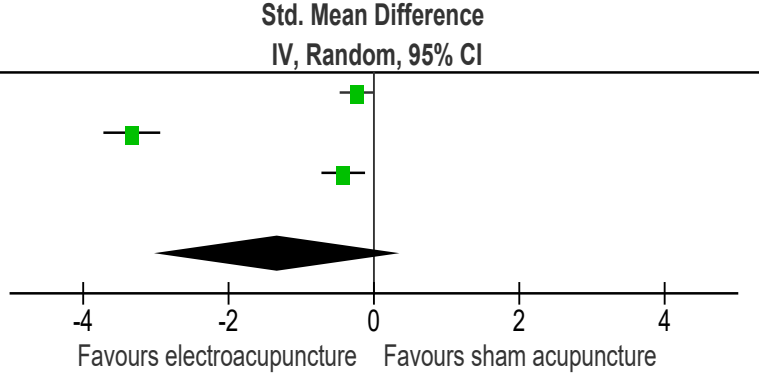


Figure 73: Pain (WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Electroacupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Random, 95% CI
Ju 2015	-0.37	0.2301	40	37	13.9%	-0.37 [-0.82, 0.08]
Mavrommatis 2012	-1.9677	0.2747	40	40	13.0%	-1.97 [-2.51, -1.43]
Suarez-almazor 2010	-0.0107	0.0992	153	302	15.9%	-0.01 [-0.21, 0.18]
Tu 2020	-0.4667	0.1177	151	146	15.7%	-0.47 [-0.70, -0.24]
Vas 2004	-1.0338	0.2168	48	49	14.1%	-1.03 [-1.46, -0.61]
Wang 2021	-0.5842	0.264	30	30	13.2%	-0.58 [-1.10, -0.07]
Weiner 2007	0.0882	0.2133	44	44	14.2%	0.09 [-0.33, 0.51]
Total (95% CI)			**506**	**648**	**100.0%**	**-0.59 [-1.02, -0.17]**

Heterogeneity: Tau² = 0.28; Chi² = 62.26, df = 6 (P < 0.00001); I² = 90%

Test for overall effect: Z = 2.74 (P = 0.006)

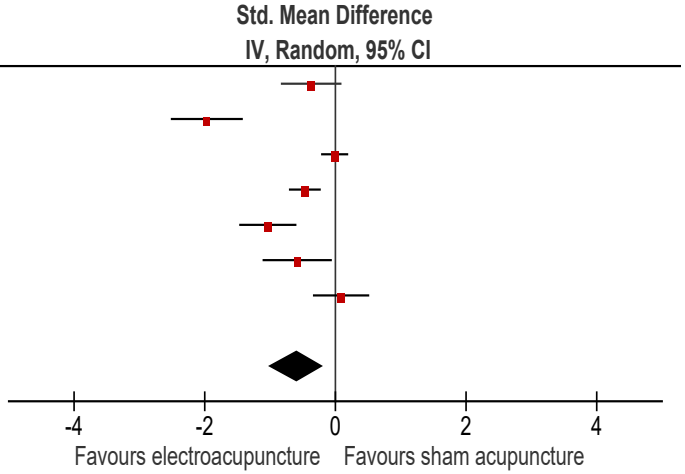


Figure 74: Pain (WOMAC, 0-20, high is poor, change score and final value) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 2004	-3.79	3.93	142	-2.92	3.56	141	30.8%	-0.87 [-1.74, 0.00]
Tu 2020	2.79	2.41	151	3.94	2.7	146	69.2%	-1.15 [-1.73, -0.57]
Total (95% CI)			**293**			**287**	**100.0%**	**-1.06 [-1.55, -0.58]**

Heterogeneity: Chi² = 0.27, df = 1 (P = 0.60); I² = 0%

Test for overall effect: Z = 4.30 (P < 0.0001)

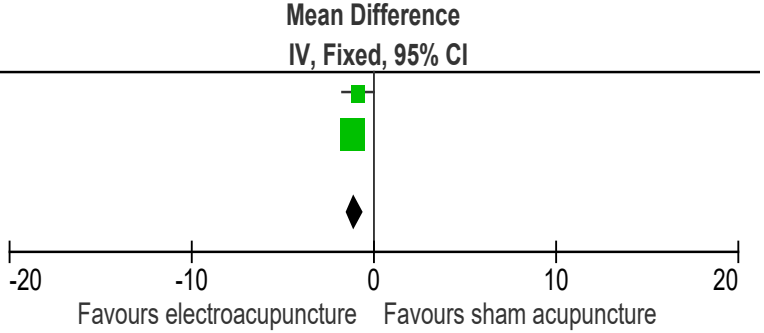


Figure 75: Physical function (WOMAC, 0-68, high is poor, change scores) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 2004	-12.18	12.07	158	-9.4	11.78	157	66.9%	-2.78 [-5.41, -0.15]
Sangdee 2002	-19.08	13.47	92	-13.4	12.54	94	33.1%	-5.68 [-9.42, -1.94]
Total (95% CI)			**250**			**251**	**100.0%**	**-3.74 [-5.89, -1.59]**

Heterogeneity: Chi² = 1.54, df = 1 (P = 0.21); I² = 35%
Test for overall effect: Z = 3.40 (P = 0.0007)

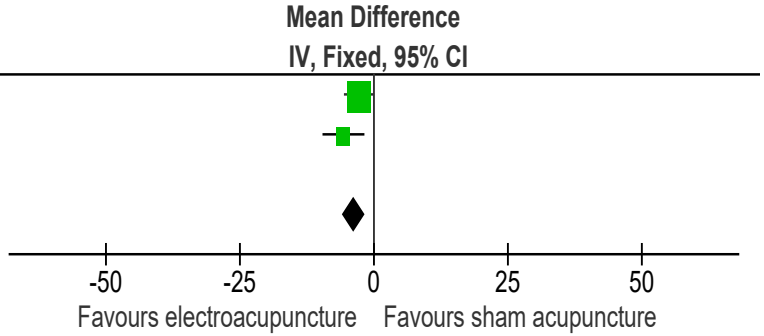


Figure 76: Physical function (WOMAC [different scale ranges], high is poor, final values) at ≤3 months

Study or Subgroup	Std. Mean Difference	SE	Electroacupuncture Total	Sham acupuncture Total	Weight	Std. Mean Difference IV, Random, 95% CI
Ju 2015	-0.5989	0.2334	40	37	13.8%	-0.60 [-1.06, -0.14]
Mavrommatis 2012	-2.0857	0.2804	40	40	12.9%	-2.09 [-2.64, -1.54]
Suarez-almazor 2010	-0.0495	0.0992	153	302	16.0%	-0.05 [-0.24, 0.14]
Tu 2020	-0.3302	0.1169	151	146	15.8%	-0.33 [-0.56, -0.10]
Vas 2004	-1.071	0.2177	48	49	14.1%	-1.07 [-1.50, -0.64]
Wang 2021	-0.441	0.2615	30	30	13.3%	-0.44 [-0.95, 0.07]
Weiner 2007	-0.1057	0.2134	44	44	14.2%	-0.11 [-0.52, 0.31]
Total (95% CI)			**506**	**648**	**100.0%**	**-0.64 [-1.06, -0.22]**

Heterogeneity: Tau² = 0.28; Chi² = 60.93, df = 6 (P < 0.00001); I² = 90%
Test for overall effect: Z = 2.97 (P = 0.003)

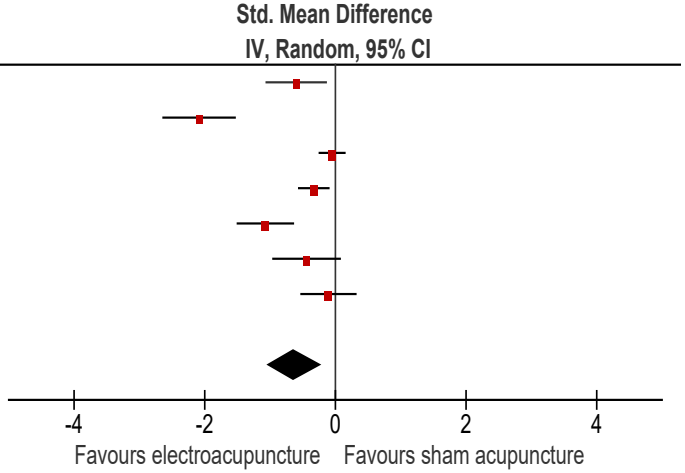


Figure 77: Physical function (WOMAC, 0-68, high is poor, change score) at >3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 2004	-12.42	13.4	142	-9.88	11	141	29.6%	-2.54 [-5.40, 0.32]
Tu 2020	9.87	7.04	151	13.21	9.09	146	70.4%	-3.34 [-5.19, -1.49]
Total (95% CI)			**293**			**287**	**100.0%**	**-3.10 [-4.66, -1.55]**

Heterogeneity: Chi² = 0.21, df = 1 (P = 0.65); I² = 0%

Test for overall effect: Z = 3.91 (P < 0.0001)

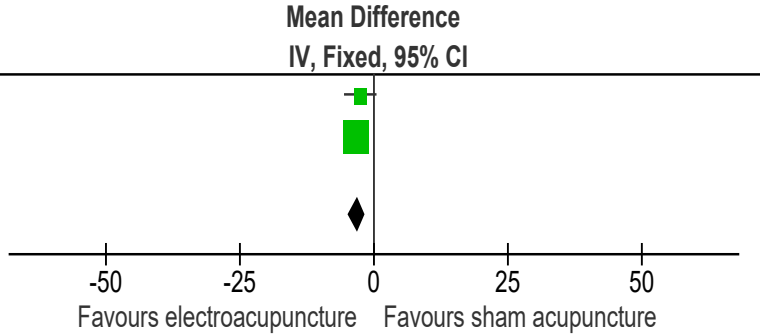


Figure 78: Psychological distress (Geriatric depression scale, 0-20, high is poor, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	Sham acupuncture Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Weiner 2007	3.78	4.82	44	3.36	3.38	44	0.42 [-1.32, 2.16]

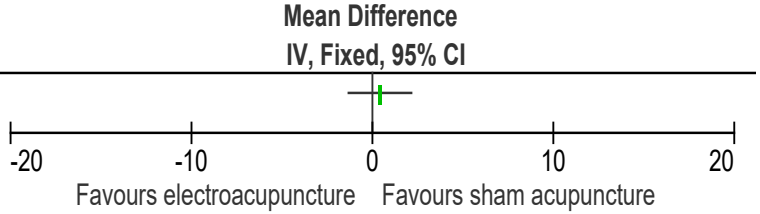


Figure 79: Osteoarthritis flares at ≤3 months

Study or Subgroup	Electroacupuncture Events	Total	Sham acupuncture Events	Total	Risk Ratio M-H, Fixed, 95% CI
Sangdee 2002	2	97	2	95	0.98 [0.14, 6.81]

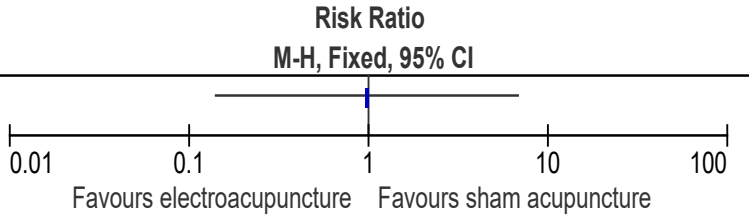


Figure 80: Serious adverse events at ≤3 months

Study or Subgroup	Electroacupuncture Events	Total	Sham acupuncture Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Lv 2019	32	217	11	75	44.3%	0.00 [-0.09, 0.09]
Sangdee 2002	3	97	0	95	38.2%	0.03 [-0.01, 0.07]
Weiner 2007	0	44	0	44	17.5%	0.00 [-0.04, 0.04]
Total (95% CI)		**358**		**214**	**100.0%**	**0.01 [-0.03, 0.06]**
Total events	35		11			

Heterogeneity: Chi² = 1.23, df = 2 (P = 0.54); I² = 0%

Test for overall effect: Z = 0.53 (P = 0.59)

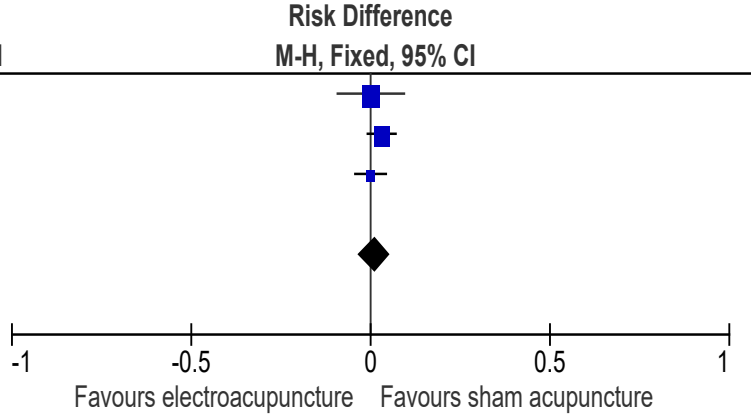
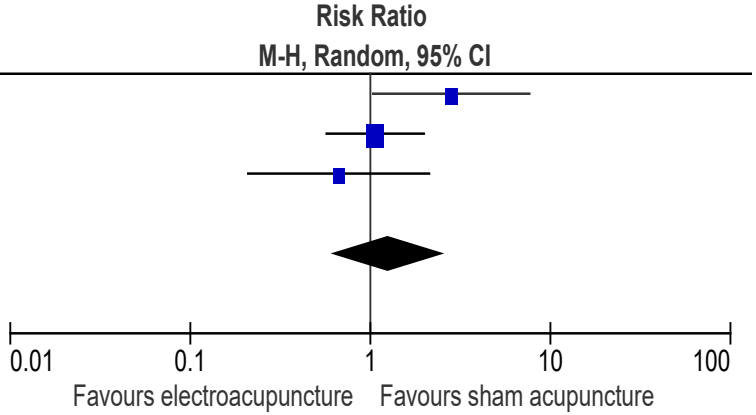


Figure 81: Serious adverse events at >3 months

Study or Subgroup	Electroacupuncture Events	Total	Sham acupuncture Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Berman 2004	14	190	5	191	29.8%	2.81 [1.03, 7.66]
Tu 2020	18	156	17	157	45.1%	1.07 [0.57, 1.99]
Wang 2021	4	30	6	30	25.1%	0.67 [0.21, 2.13]
Total (95% CI)		**376**		**378**	**100.0%**	**1.27 [0.61, 2.64]**
Total events	36		28			

Heterogeneity: Tau² = 0.21; Chi² = 3.94, df = 2 (P = 0.14); I² = 49%

Test for overall effect: Z = 0.63 (P = 0.53)



E.5 Electroacupuncture compared to no treatment

Figure 82: Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Mavrommatis 2012	45.8	6.9	40	35.3	4.5	40	50.0%	10.50 [7.95, 13.05]
Suarez-almazor 2010	39.5	9.7	153	35.8	8.9	72	50.0%	3.70 [1.13, 6.27]
Total (95% CI)			**193**			**112**	**100.0%**	**7.10 [0.44, 13.77]**

Heterogeneity: Tau² = 21.41; Chi² = 13.55, df = 1 (P = 0.0002); I² = 93%
Test for overall effect: Z = 2.09 (P = 0.04)

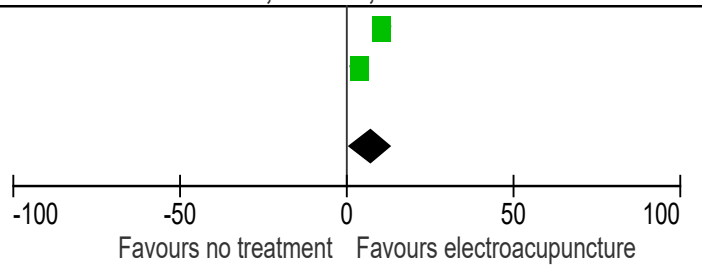


Figure 83: Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, final values) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Mavrommatis 2012	52.2	8	40	50.7	7.4	40	37.4%	1.50 [-1.88, 4.88]
Suarez-almazor 2010	54.1	8.2	153	51.6	9.8	72	62.6%	2.50 [-0.11, 5.11]
Total (95% CI)			**193**			**112**	**100.0%**	**2.13 [0.06, 4.19]**

Heterogeneity: Chi² = 0.21, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 2.02 (P = 0.04)

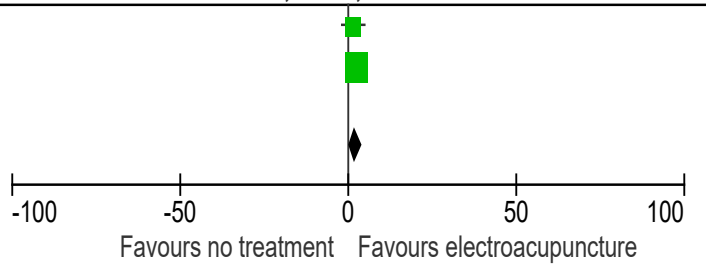


Figure 84: Quality of life (AQoL-SF 36 physical functioning, scale range unclear, high is good, final value) at ≤3 months

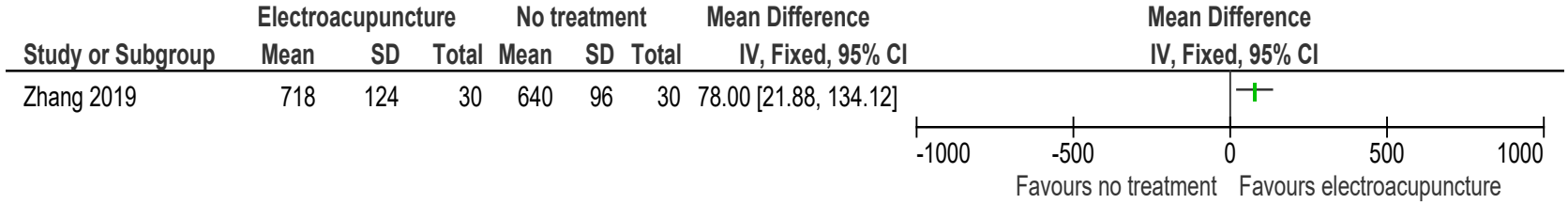


Figure 85: Quality of life (AQoL-SF 36 bodily pain, scale range unclear, high is good, final value) at ≤3 months

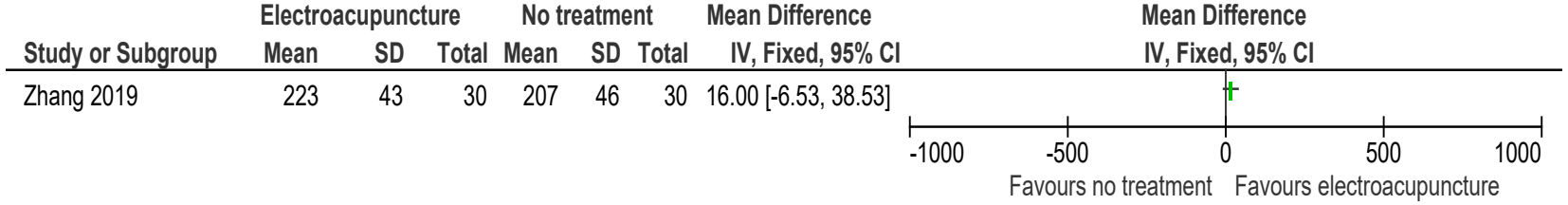


Figure 86: Quality of life (AQoL-SF 36 role physical, scale range unclear, high is good, final value) at ≤3 months

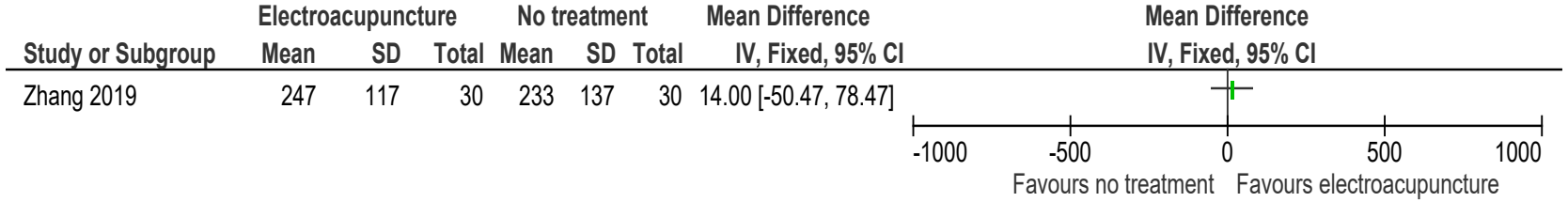


Figure 87: Quality of life (AQoL-SF 36 vitality, scale range unclear, high is good, final value) at ≤3 months

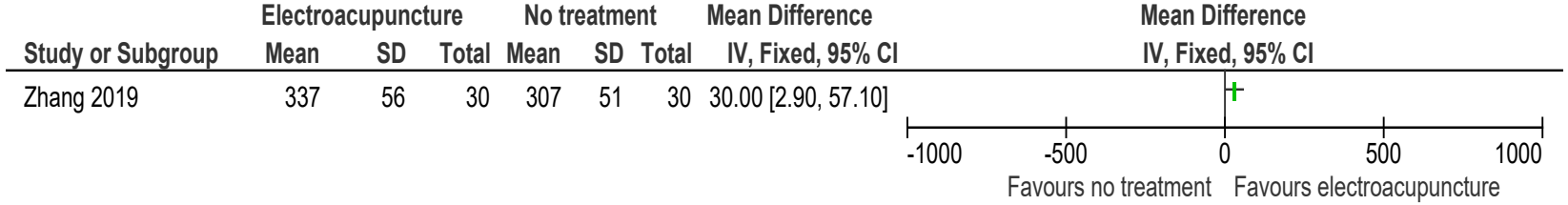


Figure 88: Quality of life (AQoL-SF 36 general health, scale range unclear, high is good, final value) at ≤3 months

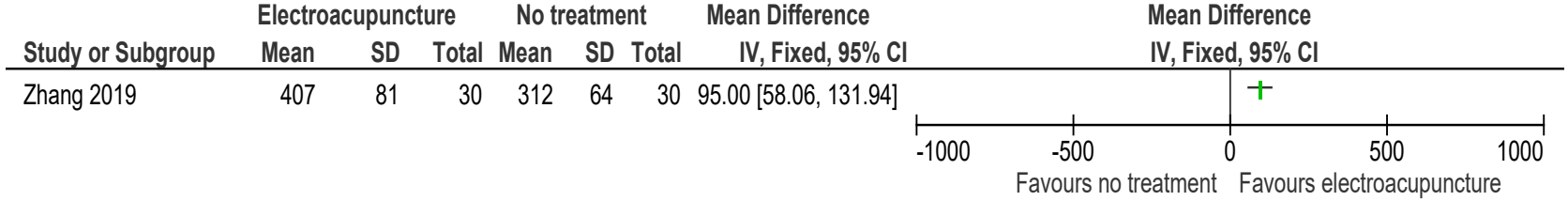


Figure 89: Quality of life (AQoL-SF 36 mental health, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	371	69	30	331	69	30	40.00 [5.08, 74.92]

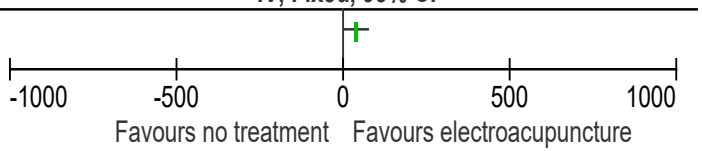


Figure 90: Quality of life (AQoL-SF 36 role emotional, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	237	72	30	207	101	30	30.00 [-14.38, 74.38]

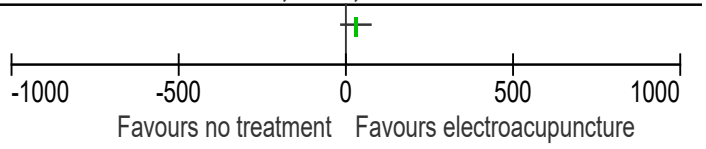


Figure 91: Quality of life (AQoL-SF 36 social functioning, scale range unclear, high is good, final value) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Zhang 2019	172	30	30	127	25	30	45.00 [31.03, 58.97]

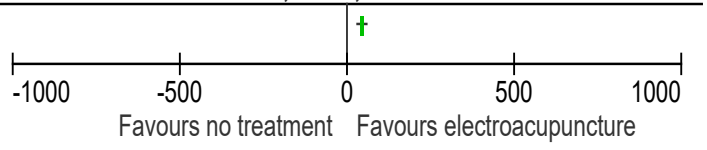


Figure 92: Pain (WOMAC, 0-20, high is poor, change score and final value) at ≤3 months

Study or Subgroup	Mean Difference	SE	Electroacupuncture Total	No treatment Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 1999	-3.95	0.756	37	36	25.0%	-3.95 [-5.43, -2.47]
Dunning 2018	-3.1	0.4363	111	111	75.0%	-3.10 [-3.96, -2.24]
Total (95% CI)			**148**	**147**	**100.0%**	**-3.31 [-4.05, -2.57]**

Heterogeneity: Chi² = 0.95, df = 1 (P = 0.33); I² = 0%
Test for overall effect: Z = 8.77 (P < 0.00001)

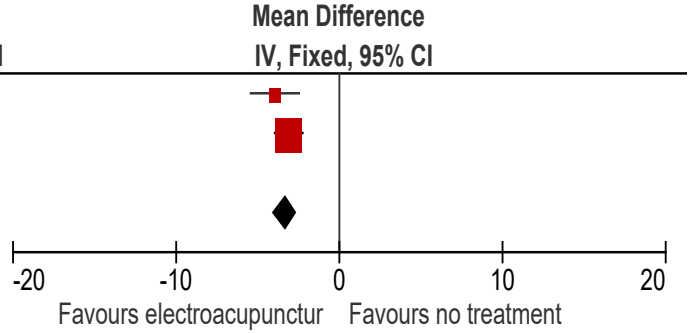


Figure 93: Pain (WOMAC [different scale ranges], high as poor, final values) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Mavrommatis 2012	77.9	32.9	40	153.8	41.2	40	48.5%	-2.02 [-2.56, -1.47]
Suarez-almazor 2010	30.8	17.9	153	42.4	16.8	72	51.5%	-0.66 [-0.95, -0.37]
Total (95% CI)			**193**			**112**	**100.0%**	**-1.32 [-2.65, 0.01]**

Heterogeneity: Tau² = 0.87; Chi² = 18.79, df = 1 (P < 0.0001); I² = 95%
Test for overall effect: Z = 1.94 (P = 0.05)

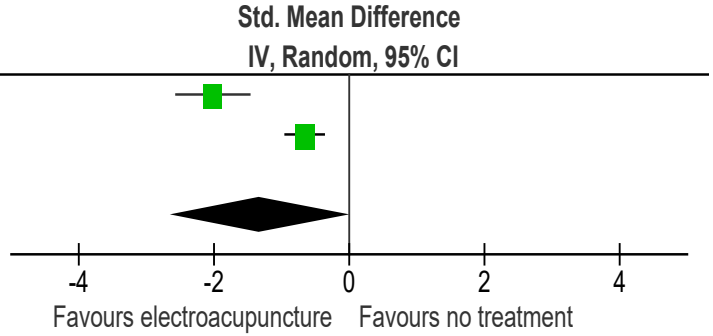


Figure 94: Physical function (WOMAC, 0-68, high is poor, change score and final value) at ≤3 months

Study or Subgroup	Mean Difference	SE	Electroacupuncture Total	No treatment Total	Weight	Mean Difference IV, Fixed, 95% CI
Berman 1999	-13.61	2.9023	37	36	18.2%	-13.61 [-19.30, -7.92]
Dunning 2018	-9.4	1.3702	111	111	81.8%	-9.40 [-12.09, -6.71]
Total (95% CI)			**148**	**147**	**100.0%**	**-10.17 [-12.60, -7.74]**

Heterogeneity: Chi² = 1.72, df = 1 (P = 0.19); I² = 42%

Test for overall effect: Z = 8.21 (P < 0.00001)

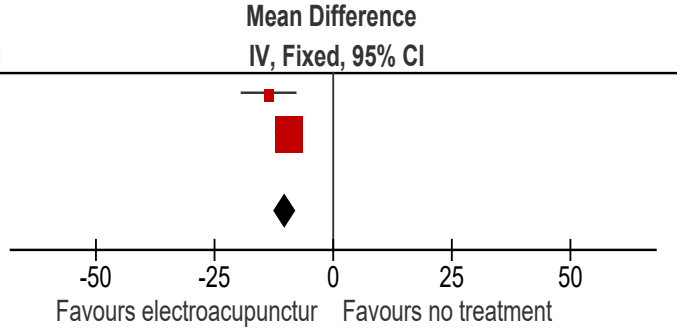


Figure 95: Physical function (WOMAC [different scale ranges], high as poor, final values) at ≤3 months

Study or Subgroup	Electroacupuncture Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Mavrommatis 2012	321.6	141	40	653.4	157	40	48.8%	-2.20 [-2.76, -1.64]
Suarez-almazor 2010	31.2	17.9	153	41.7	18	72	51.2%	-0.58 [-0.87, -0.30]
Total (95% CI)			**193**			**112**	**100.0%**	**-1.37 [-2.96, 0.21]**

Heterogeneity: Tau² = 1.26; Chi² = 25.42, df = 1 (P < 0.00001); I² = 96%

Test for overall effect: Z = 1.70 (P = 0.09)

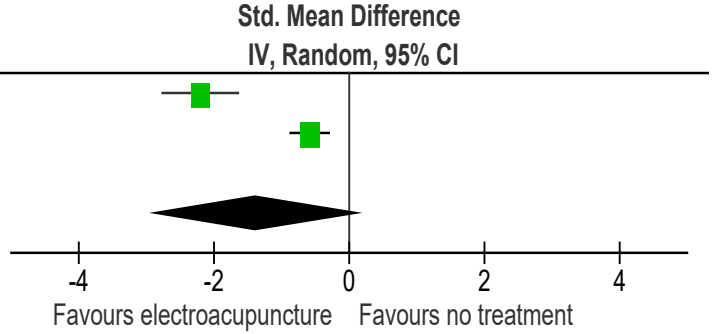
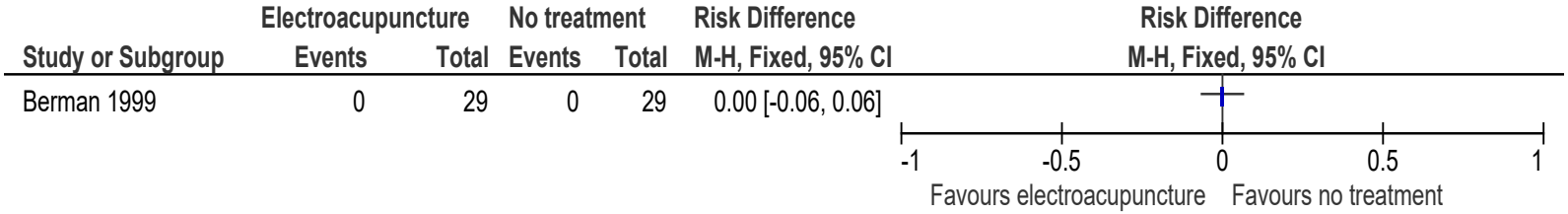


Figure 96: Serious adverse events at ≤3 months



Appendix F – GRADE tables

Table 19: Clinical evidence profile: acupuncture/dry needling compared to sham acupuncture

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
Quality of life (EQ-5D, 5-15, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: EQ-5D; Scale from: 5 to 15)												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	31	31	-	MD **0.13 higher** (0.56 lower to 0.82 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	not serious	not serious	not serious	not serious	none	786	755	-	MD **1.26 higher** (0.21 higher to 2.32 higher)	⨁⨁⨁⨁ High	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	not serious	not serious	not serious	not serious	none	786	755	-	MD **0.56 higher** (0.48 lower to 1.6 higher)	⨁⨁⨁⨁ High	CRITICAL
Quality of life (EQ-5D, 5-15, high is good, final value) at >3 months (follow-up: 12 months; assessed with: EQ-5D; Scale from: 5 to 15)												
1	randomised trials	serious[a]	not serious	serious[b]	serious[c]	none	31	31	-	MD **0.15 higher** (0.58 lower to 0.88 higher)	⨁◯◯◯ Very low	CRITICAL
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score and final values) at >3 months (follow-up: mean 33 weeks; assessed with: SF-36 physical component, SF-12 physical 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	acupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
4	randomised trials	not serious	not serious	not serious	not serious	none	642	608	-	MD **1.31 higher** (0.13 higher to 2.49 higher)	⨁⨁⨁⨁ High	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental component, 0-100, high is good, change score and final values) at >3 months (follow-up: mean 33 weeks; assessed with: SF-36 mental component, SF-12 mental component; Scale from: 0 to 100)												
4	randomised trials	not serious	very serious[d]	not serious	not serious	none	642	608	-	MD **0.92 higher** (2.11 lower to 3.95 higher)	⨁⨁◯◯ Low	CRITICAL
Pain (WOMAC [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC)												
5	randomised trials	not serious	not serious	not serious	not serious	none	604	648	-	SMD **0.08 SD lower** (0.19 lower to 0.03 higher)	⨁⨁⨁⨁ High	CRITICAL
Pain (WOMAC, KSS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC, KSS)												
8	randomised trials	not serious	not serious	not serious	not serious	none	480	400	-	SMD **0.3 SD lower** (0.44 lower to 0.17 lower)	⨁⨁⨁⨁ High	CRITICAL
Pain (WOMAC [different scale ranges], high is poor, change scores) at >3 months (follow-up: mean 35 weeks; assessed with: WOMAC)												
3	randomised trials	not serious	not serious	not serious	not serious	none	529	579	-	SMD **0.06 SD lower** (0.18 lower to 0.06 higher)	⨁⨁⨁⨁ High	CRITICAL
Pain (WOMAC [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 29 weeks; assessed with: WOMAC)												
4	randomised trials	not serious	not serious	not serious	not serious	none	347	274	-	SMD **0.21 SD lower** (0.38 lower to 0.05 lower)	⨁⨁⨁⨁ High	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
Physical function (WOMAC [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC)												
5	randomised trials	not serious	not serious	not serious	not serious	none	604	643	-	SMD **0.06 SD lower** (0.17 lower to 0.05 higher)	⨁⨁⨁⨁ High	CRITICAL
Physical function (WOMAC, KSS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC, KSS)												
7	randomised trials	not serious	not serious	not serious	not serious	none	406	329	-	SMD **0.28 SD lower** (0.43 lower to 0.13 lower)	⨁⨁⨁⨁ High	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, change scores) at >3 months (follow-up: mean 35 weeks; assessed with: WOMAC)												
3	randomised trials	not serious	not serious	not serious	not serious	none	530	578	-	SMD **0.03 SD lower** (0.15 lower to 0.09 higher)	⨁⨁⨁⨁ High	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, final values) at >3 months (follow-up: mean 29 weeks; assessed with: WOMAC)												
4	randomised trials	not serious	not serious	not serious	not serious	none	347	274	-	SMD **0.22 SD lower** (0.38 lower to 0.06 lower)	⨁⨁⨁⨁ High	CRITICAL
Psychological distress (Depression ADS, 0-100, high is poor, final value) at <3 months (follow-up: 8 weeks; assessed with: Depression ADS; Scale from: 0 to 100)												
1	randomised trials	not serious	not serious	not serious	not serious	none	150	76	-	MD **0.4 lower** (3.07 lower to 2.27 higher)	⨁⨁⨁⨁ High	IMPORTANT
Psychological distress (Depression ADS, 0-100, high is poor, final value) at >3 months (follow-up: 52 weeks; assessed with: Depression ADS; Scale from: 0 to 100)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	not serious	none	150	76	-	MD **1.2 lower** (4 lower to 1.6 higher)	⨁⨁⨁⨁ High	IMPORTANT
Serious adverse events at <3 months (follow-up: mean 4 weeks)												
2	randomised trials	serious[a]	very serious[e]	not serious	not serious	none	22/107 (20.6%)	4/105 (3.8%)	**RD 0.12** (-0.26 to 0.50)	**120 more per 1,000** (from 260 fewer to 500 more)[f]	⨁◯◯◯ Very low	IMPORTANT
Serious adverse events at >3 months (follow-up: mean 26 weeks)												
5	randomised trials	not serious	serious[e]	not serious	very serious[g]	none	251/757 (33.2%)	227/728 (31.2%)	**RD 0.04** (-0.02 to 0.10)	**40 more per 1,000** (from 20 fewer to 100 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 or 2 increments because of population indirectness

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 by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

e. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

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

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

Table 20: Clinical evidence profile: acupuncture/dry needling compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	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 7 weeks; assessed with: EQ-5D, KOOS)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	166	143	-	SMD **0.11 higher** (0.11 lower to 0.34 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 11 weeks; assessed with: SF-36 physical component, SF-12 physical component; Scale from: 0 to 100)												
3	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	536	453	-	MD **4.69 higher** (1.27 higher to 8.11 higher)	⨁◯◯◯ VERY 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 11 weeks; assessed with: SF-36 mental component, SF-12 mental component; Scale from: 0 to 100)												
3	randomised trials	serious[a]	very serious[a]	not serious	not serious	none	536	498	-	MD **0.41 higher** (2.86 lower to 3.69 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 physical functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 physical functioning)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **53 higher** (0.88 lower to 106.88 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 bodily pain, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 bodily pain; Scale from: 0 to 1000)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **10 higher** (13.28 lower to 33.28 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 role physical, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 role physical)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **10 higher** (55.64 lower to 75.64 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (AQoL-SF-36 vitality, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 vitality)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **24 higher** (2.32 lower to 50.32 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (AQoL-SF-36 general health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 general health)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **75 higher** (34.31 higher to 115.69 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (AQoL-SF-36 mental health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 mental health)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **24 higher** (10.92 lower to 58.92 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (AQoL-SF-36 role emotional, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 role emotional)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **36 higher** (10.55 lower to 82.55 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (AQoL-SF-36 social functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 social functioning)

1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **27 higher** (13.3 higher to 40.7 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Quality of life (EQ-5D, -0.11-1, high is good, final values) at >3 months

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	very serious[a]	not serious	not serious	very serious[c]	none	132	131	-	MD **0.01 higher** (0.06 lower to 0.08 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-12 physical component, 0-100, high is good, final value) at >3 months (follow-up: 12 months; assessed with: SF-12 physical component; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	59	62	-	MD **2.8 higher** (1.12 lower to 6.72 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (SF-12 mental component, 0-100, high is good, final value) at >3 months (follow-up: 12 months; assessed with: SF-12 mental component; Scale from: 0 to 100)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	59	62	-	MD **2.9 lower** (6.68 lower to 0.88 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score and final values) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
3	randomised trials	serious[a]	not serious	not serious	not serious	none	192	189	-	MD **0.86 lower** (1.62 lower to 0.1 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Pain (KOOS, WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 9 weeks; assessed with: KOOS, WOMAC)												
4	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	566	456	-	SMD **0.81 lower** (1.18 lower to 0.45 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score and final values) at >3 months (follow-up: mean 12 months; assessed with: WOMAC; Scale from: 0 to 20)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	173	175	-	MD **0.22 lower** (1.07 lower to 0.63 higher)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score and final values) at <3 months (follow-up: mean 10 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	acupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
3	randomised trials	serious[a]	not serious	not serious	not serious	none	192	189	-	MD **2.05 lower** (4.46 lower to 0.36 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (KOOS, WOMAC, 0-100, high is poor, final values) at <3 months (follow-up: mean 8 weeks; assessed with: KOOS, WOMAC; Scale from: 0 to 100)												
3	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	506	396	-	MD **15.58 lower** (23.58 lower to 7.57 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score and final values) at >3 months (follow-up: 12 months; assessed with: WOMAC; Scale from: 0 to 68)												
3	randomised trials	very serious[a]	not serious	not serious	not serious	none	174	174	-	MD **1.14 lower** (3.92 lower to 1.63 higher)	⨁⨁◯◯ 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	60	-	MD **0.34 higher** (1.11 lower to 1.79 higher)	⨁⨁◯◯ LOW	IMPORTANT
Psychological distress (HADS depression, depression ADS [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: HADS depression, depression ADS)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	210	134	-	SMD **0.14 SD lower** (0.36 lower to 0.08 higher)	⨁⨁⨁◯ MODERATE	IMPORTANT
Serious adverse events at <3 months (follow-up: 12 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	serious[d]	none	0/60 (0.0%)	0/60 (0.0%)	not estimable	**0 fewer per 1,000** (from 30 fewer to 30 more)[e]	⨁◯◯◯ VERY LOW	IMPORTANT
Serious adverse events at >3 months (follow-up: 12 months)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	acupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	very serious[d]	none	0/15 (0.0%)	0/15 (0.0%)	not estimable	**0 fewer per 1,000** (from 120 fewer to 120 more)	⨁◯◯◯ 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 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 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 21: Clinical evidence profile: electroacupuncture compared to acupuncture/dry needling

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	acupuncture	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-12, 0-100, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: SF-12; Scale from: 0 to 100)												
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	28	30	-	MD **1.34 higher** (7.75 lower to 10.43 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-12 physical health, 0-100, high is good, final value) at <3 months (follow-up: 8 weeks; assessed with: SF-12 physical 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	electroacupuncture	acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	not serious	none	151	145	-	MD **0.25 lower** (2.14 lower to 1.64 higher)	⨁⨁⨁⨁ HIGH	CRITICAL

Quality of life (SF-12 mental health, 0-100, high is good, final value) at <3 months (follow-up: 8 weeks; assessed with: SF-12 mental health; Scale from: 0 to 100)

1	randomised trials	not serious	not serious	not serious	not serious	none	151	145	-	MD **0.94 lower** (2.96 lower to 1.08 higher)	⨁⨁⨁⨁ HIGH	CRITICAL

Quality of life (AQoL-SF 36 physical functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 physical functioning, scale range unclear)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **25 higher** (35.76 lower to 85.76 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 bodily pain, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 bodily pain)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **6 higher** (16.53 lower to 28.53 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 role physical, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 role physical)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **4 higher** (56.49 lower to 64.49 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 vitality, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 vitality)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **6 higher** (21.59 lower to 33.59 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 general health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 general health)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **20 higher** (24.4 lower to 64.4 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 mental health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 mental health)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **16 higher** (18.92 lower to 50.92 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 role emotional, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 role emotional)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **6 lower** (45.05 lower to 33.05 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (AQoL-SF 36 social functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF 36 social functioning)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	30	30	-	MD **18 higher** (3.07 higher to 32.93 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (SF-12, 0-100, high is good, final value) at >3 months (follow-up: 16 weeks; assessed with: SF-12; Scale from: 0 to 100)

1	randomised trials	serious[a]	not serious	not serious	serious[b]	none	28	30	-	MD **1.77 higher** (7.32 lower to 10.86 higher)	⨁⨁◯◯ LOW	CRITICAL

Quality of life (SF-12 physical health, 0-100, high is good, final value) at >3 months (follow-up: 26 weeks; assessed with: SF-12 physical health; Scale from: 0 to 100)

1	randomised trials	not serious	not serious	not serious	not serious	none	151	145	-	MD **0.02 lower** (2 lower to 1.96 higher)	⨁⨁⨁⨁ HIGH	CRITICAL

Quality of life (SF-12 mental health, 0-100, high is good, final value) at >3 months (follow-up: 26 weeks; assessed with: SF-12 mental 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	electroacupuncture	acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	not serious	not serious	not serious	not serious	none	151	145	-	MD **0.81 lower** (2.66 lower to 1.04 higher)	⨁⨁⨁⨁ HIGH	CRITICAL
Pain (WOMAC, 0-20, high is poor, final value) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
2	randomised trials	not serious	not serious	not serious	not serious	none	179	175	-	MD **0.37 lower** (0.78 lower to 0.04 higher)	⨁⨁⨁⨁ HIGH	CRITICAL
Pain (WOMAC, 0-20, high is poor, final value) at >3 months (follow-up: mean 21 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
2	randomised trials	not serious	not serious	not serious	not serious	none	179	175	-	MD **0.43 lower** (0.9 lower to 0.04 higher)	⨁⨁⨁⨁ HIGH	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	not serious	not serious	not serious	serious[b]	none	179	175	-	MD **1.47 lower** (2.96 lower to 0.02 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (WOMAC, 0-68, high is poor, final value) at >3 months (follow-up: mean 21 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	not serious	not serious	not serious	serious[b]	none	179	175	-	MD **1.63 lower** (3.19 lower to 0.06 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Serious adverse events at >3 months (follow-up: mean 21 weeks)												
2	randomised trials	not serious	not serious	not serious	very serious[b]	none	27/184 (14.7%)	33/185 (17.8%)	**RR 0.83** (0.53 to 1.31)	**30 fewer per 1,000** (from 84 fewer to 55 more)	⨁⨁◯◯ 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 22: Clinical evidence profile: electroacupuncture compared to sham acupuncture

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, change score 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)												
5	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	543	687	-	MD **3.07 higher** (0.55 lower to 6.68 higher)	⨁◯◯◯ VERY 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)												
4	randomised trials	serious[a]	not serious	not serious	not serious	none	374	518	-	MD **0.71 higher** (0.4 lower to 1.83 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Quality of life (PLQC physical capability, 0-4, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: PLQC physical capability; Scale from: 0 to 4)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	48	49	-	MD **0.3 higher** (0 to 0.6 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (PLQC psychological functioning, 0-4, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: PLQC psychological functioning; Scale from: 0 to 4)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	48	49	-	MD **0.2 higher** (0 to 0.4 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (PLQC negative mood, 0-4, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: PLQC negative mood; Scale from: 0 to 4)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	48	49	-	MD **0.1 higher** (0.18 lower to 0.38 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (PLQC social functioning, 0-4, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: PLQC social functioning; Scale from: 0 to 4)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	48	49	-	MD **0.1 higher** (0.14 lower to 0.34 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (PLQC social wellbeing, 0-4, high is good, final value) at <3 months (follow-up: 12 weeks; assessed with: PLQC social wellbeing; Scale from: 0 to 4)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	48	49	-	MD **0** (0.2 lower to 0.2 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (SF-36 physical component, SF-12 physical health, 0-100, high is good, change score and final value) at >3 months (follow-up: mean 26 weeks; assessed with: SF-36 physical component; Scale from: 0 to 100)												
3	randomised trials	not serious	not serious	not serious	serious[c]	none	323	317	-	MD **1.19 higher** (0.32 lower to 2.7 higher)	⨁⨁⨁◯ MODERATE	CRITICAL
Quality of life (SF-36 mental component, SF-12 mental health, 0-100, high is good, final value) at >3 months (follow-up: 26 weeks; assessed with: SF-12 mental health; Scale from: 0 to 100)												
1	randomised trials	not serious	not serious	not serious	not serious	none	181	176	-	MD **2.21 higher** (0.47 higher to 3.96 higher)	⨁⨁⨁⨁ HIGH	CRITICAL
Pain (WOMAC, VAS [different scale ranges], high is poor, change scores) at <3 months (follow-up: mean 7 weeks; assessed with: WOMAC, VAS)												
3	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	472	326	-	SMD **1.32 SD lower** (3 lower to 0.36 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
7	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	506	648	-	SMD **0.59 SD lower** (1.02 lower to 0.17 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score and final value) at >3 months (follow-up: mean 26 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
2	randomised trials	not serious	not serious	not serious	not serious	none	293	287	-	MD **1.06 lower** (1.55 lower to 0.58 lower)	⨁⨁⨁⨁ HIGH	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change scores) at <3 months (follow-up: mean 9 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	250	251	-	MD **3.74 lower** (5.89 lower to 1.59 lower)	⨁⨁⨁◯ MODERATE	CRITICAL
Physical function (WOMAC [different scale ranges], high is poor, final values) at <3 months (follow-up: mean 10 weeks; assessed with: WOMAC)												
7	randomised trials	serious[a]	very serious[b]	not serious	serious[c]	none	506	648	-	SMD **0.64 SD lower** (1.06 lower to 0.22 lower)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score) at >3 months (follow-up: mean 26 weeks; assessed with: WOMAC; Scale from: 0 to 68)												
2	randomised trials	not serious	not serious	not serious	not serious	none	293	287	-	MD **3.1 lower** (4.66 lower to 1.55 lower)	⨁⨁⨁⨁ HIGH	CRITICAL
Psychological distress (Geriatric depression scale, 0-20, high is poor, final value) at <3 months (follow-up: 12 weeks; assessed with: Geriatric depression scale; Scale from: 0 to 20)												
1	randomised trials	serious[a]	not serious	not serious	serious[c]	none	44	44	-	MD **0.42 higher** (1.32 lower to 2.16 higher)	⨁⨁◯◯ LOW	IMPORTANT
Osteoarthritis flares at <3 months (follow-up: 4 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	sham acupuncture	Relative (95% CI)	Absolute (95% CI)		
1	randomised trials	serious[a]	not serious	serious[d]	very serious[c]	none	2/97 (2.1%)	2/95 (2.1%)	**RR 0.98** (0.14 to 6.81)	**0 fewer per 1,000** (from 18 fewer to 122 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Serious adverse events at <3 months (follow-up: mean 6 weeks)												
3	randomised trials	serious[a]	serious[e]	not serious	very serious[f]	none	35/358 (9.8%)	11/214 (5.1%)	**RD 0.01** (-0.03 to 0.06)	**10 more per 1,000** (from 30 fewer to 60 more)[g]	⨁◯◯◯ VERY LOW	IMPORTANT
Serious adverse events at >3 months (follow-up: mean 26 weeks)												
3	randomised trials	serious[a]	serious[b]	not serious	very serious[c]	none	36/376 (9.6%)	28/378 (7.4%)	**RR 1.27** (0.61 to 2.64)	**20 more per 1,000** (from 29 fewer to 121 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 by 1 or 2 increments because of outcome indirectness

e. Downgraded for heterogeneity due to conflicting number of events in different studies (zero events in one or more studies)

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

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

Table 23: Clinical evidence profile: electroacupuncture compared to no treatment

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (SF-36 physical component, SF-12 physical component, 0-100, high is good, final values) at <3 months (follow-up: 10 weeks; assessed with: SF-36 physical component, SF-12 physical component; Scale from: 0 to 100)												
2	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	193	112	-	MD **7.1 higher** (0.44 higher to 13.77 higher)	⨁◯◯◯ VERY 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 10 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	not serious	none	193	112	-	MD **2.13 higher** (0.06 higher to 4.19 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (AQoL-SF-36 physical functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 physical functioning)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **78 higher** (21.88 higher to 134.12 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 bodily pain, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 bodily pain)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **16 higher** (6.53 lower to 38.53 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 role physical, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 role physical)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **14 higher** (50.47 lower to 78.47 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 vitality, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 vitality)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **30 higher** (2.9 higher to 57.1 higher)	⨁◯◯◯ VERY LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	electroacupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
Quality of life (AQoL-SF-36 general health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 general health)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	30	30	-	MD **95 higher** (58.06 higher to 131.94 higher)	⨁⨁◯◯ LOW	CRITICAL
Quality of life (AQoL-SF-36 mental health, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 mental health)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **40 higher** (5.08 higher to 74.92 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 role emotional, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 role emotional)												
1	randomised trials	very serious[a]	not serious	not serious	serious[c]	none	30	30	-	MD **30 higher** (14.38 lower to 74.38 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Quality of life (AQoL-SF-36 social functioning, scale range unclear, high is good, final value) at <3 months (follow-up: 2 weeks; assessed with: AQoL-SF-36 social functioning)												
1	randomised trials	very serious[a]	not serious	not serious	not serious	none	30	30	-	MD **45 higher** (31.03 higher to 58.97 higher)	⨁⨁◯◯ LOW	CRITICAL
Pain (WOMAC, 0-20, high is poor, change score and final value) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC; Scale from: 0 to 20)												
2	randomised trials	serious[a]	not serious	not serious	not serious	none	148	147	-	MD **3.31 lower** (4.05 lower to 2.57 lower)	⨁⨁⨁◯ Moderate	CRITICAL
Pain (WOMAC [different scale ranges], high as poor, final values) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC)												
2	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	193	112	-	SMD **1.32 SD lower** (2.65 lower to 0.01 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Physical function (WOMAC, 0-68, high is poor, change score and final value) at <3 months (follow-up: mean 12 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	electroacupuncture	no treatment	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious[a]	not serious	not serious	serious[c]	none	148	147	-	MD **10.17 lower** (12.6 lower to 7.74 lower)	⨁⨁◯◯ LOW	CRITICAL
Physical function (WOMAC [different scale ranges], high as poor, final values) at <3 months (follow-up: mean 12 weeks; assessed with: WOMAC)												
2	randomised trials	very serious[a]	very serious[b]	not serious	serious[c]	none	193	112	-	SMD **1.37 SD lower** (2.96 lower to 0.21 higher)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse events at <3 months (follow-up: 12 weeks)												
1	randomised trials	very serious[a]	not serious	not serious	very serious[d]	none	0/29 (0.0%)	0/29 (0.0%)	**RD 0.00** (-0.06 to 0.06)	**0 fewer per 1,000** (from 60 fewer to 60 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

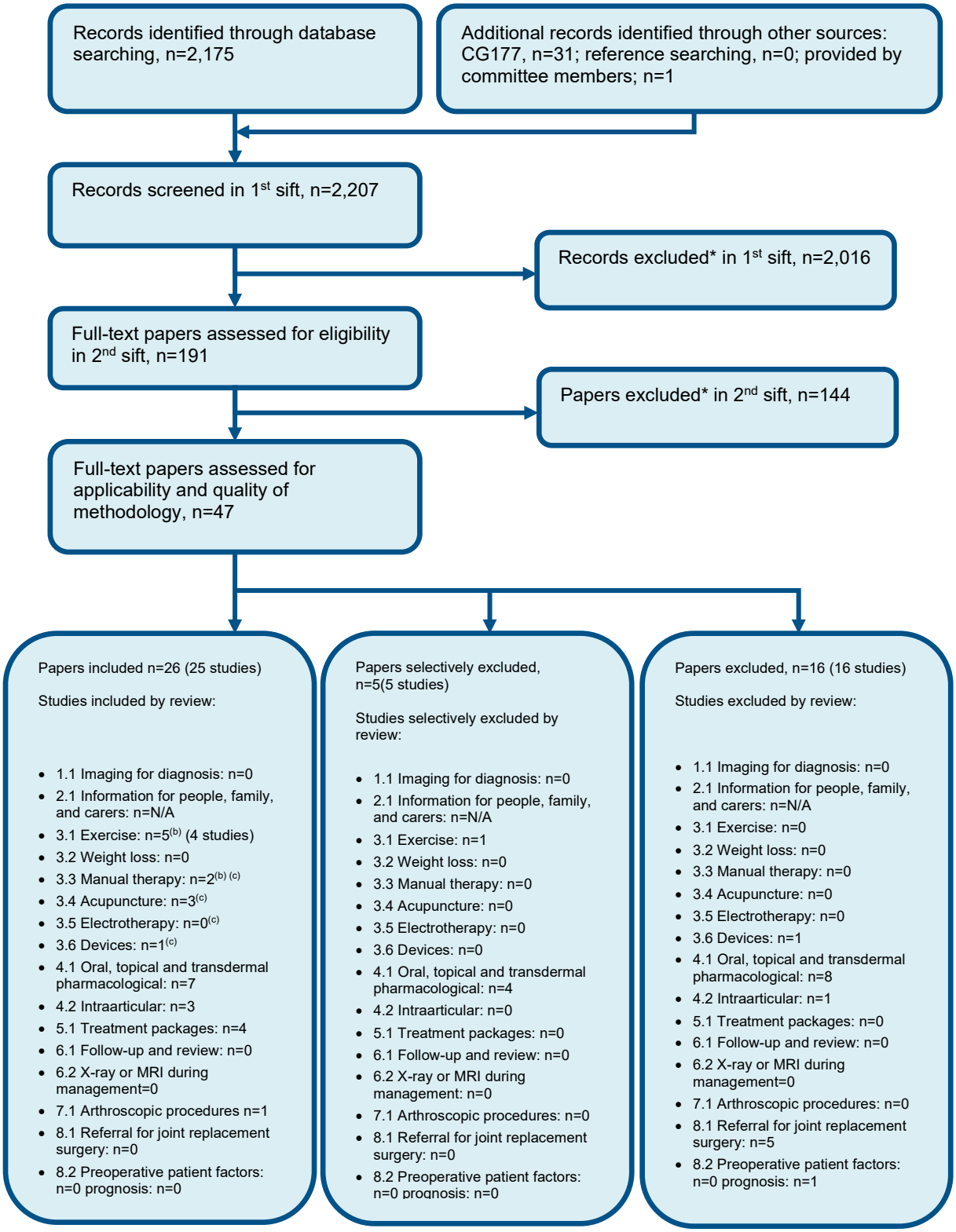
b. Downgraded by 1 or 2 increments because heterogeneity, unexplained by subgroup analysis

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 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

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	Latimer (2012)[65]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome = QALYs) **Study design:** Separate analyses of three trials **Approach to analysis:** Reanalysis of the NICE osteoarthritis guideline CG59 wherein acupuncture was compared against usual care and placebo (sham acupuncture). **Perspective:** UK NHS **Time horizon:** Berman 2004: 6 months Scharf 2006: 6 months Witt 2005: 8 weeks **Treatment duration:** Berman 2004: 23 sessions over 26 weeks Scharf 2006: 10 sessions over 6 weeks Witt (2005): 12 sessions over 8 weeks **Discounting:** n/a	**Population:** Adults aged 16 years or older who have a diagnosis of OA. **Patient characteristics:** **Berman 2004** Age = 65 Male = 36% N = 570 **Scharf 2006** Age = 63 Male = 31% N = 1007 **Witt 2005** Age = 64 Male = 34% N = 294 **Intervention 1:** Usual care (specific treatment not described)	**Total costs (mean per patient):** Incremental (2-1): Berman 2004: £414 (CI = NR; p=NR) Scharf 2006: £225 (CI = NR; p=NR) Witt 2005: £216 (CI = NR; p=NR) **Currency & cost year:** 2010 UK pounds. **Cost components incorporated:** Intervention costs only: physiotherapists time and the cost of acupuncture needles.	**QALYs gained versus baseline (mean per patient):** Intervention 1: Berman 2004: 0.033 Scharf 2006: 0.038 Witt 2005: 0.002 Intervention 2: Berman 2004: 0.056 Scharf 2006: 0.071 Witt 2005: 0.016 Incremental (2-1): Berman 2004: 0.024 (CI = NR; p=NR) Scharf 2006: £0.033 (CI = NR; p=NR) Witt 2005: 0.014 (CI = NR; p=NR)	**Cost per QALY gained:** Incremental (2-1): Berman 2004: £17,381 Scharf 2006: £6,911 Witt 2005: £15,621 Incremental (2-1): Using comparison with Sham for effects and comparison with usual care for costs Berman 2004: £40,039 Scharf 2006: £68,284 Witt 2005: £70,519 **Analysis of uncertainty:** Results of sensitivity analysis were not published.

	Intervention 2: Acupuncture			
Data sources				
Health outcomes: This is cost utility analysis reported alongside a randomised controlled trial, the results of the trial are reported elsewhere.[9, 113, 155] **Quality-of-life weights:** Utilities were transformed into generic EQ-5D quality-of-life scores from the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). The quality-of-life calculation used in Witt et al. differed from the methodology used in the OA guidelines and may result in an overestimation of the total QALYs. **Cost sources:** The cost to the NHS (physiotherapists time) were obtained from the Personal Social Services Research Unit. Acupuncture sessions were assigned a unit cost associated with NHS community physical therapy. The cost of traditional acupuncture needles was taken from the NICE 2008 OA clinical guidelines. **Resource use:** It was assumed that 30 minutes is needed for a physiotherapist to deliver a session.				
Comments				
Source of funding: Not reported. **Limitations:** 2010 resource use and unit costs may not reflect current UK NHS practice. The time horizon varies across studies and the results did not include any extrapolations outside the trial settings. The costs of sham acupuncture were assumed to be zero, which is not reflective of 'real-world' costs. Adverse events and their downstream consequences were not considered. Sensitivity analysis of the results were not conducted. **Other:** The comparison between true acupuncture and sham acupuncture goes against the standard practice adopted by the NGC wherein all interventions are compared with placebo to ensure a standard methodological practice across all disease areas. The applicability of the study results to the guideline development of osteoarthritis should therefore be carefully considered.				
Overall applicability:[(a)] Directly applicable		**Overall quality:**[(b)] Potentially serious limitations		

Abbreviations: CI = 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; NHS = National Health Service; NGC = National Guideline Centre; NR = not reported; OA = Osteoarthritis; QALYs = quality-adjusted life years; UK = United Kingdom; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

(a) Directly applicable / Partially applicable / Not applicable

(b) Minor limitations /Potentially serious limitations / Very serious limitations

<table>
<tr><th>Study</th><th colspan="4">MacPherson 2017 [85]</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: Network meta-analysis based on a systematic review of 88 trials. Three different networks were used:
1. All trials
2. Subset of trials that were graded with a low risk of bias for allocation concealment
3. Same as point 2 but further restricting trials to those that reported outcomes between 3 and 13 weeks.

Approach to analysis: QALY changes from the different networks</td>
<td>Population:
Patients reporting pain resulting from OA of the knee

Patient characteristics:
Mean age across all trials = 53-85
Male = NR

Intervention 1: Usual care (specific treatment not described)
Intervention 2: Static magnets
Intervention 3: Insoles
Intervention 4: TENS
Intervention 5: Braces
Intervention 6: Acupuncture
Intervention 7: Heat treatment
Intervention 8: Manual therapy</td>
<td>Total costs (mean per patient):
<u>All trials</u>
Intervention 1: £0
Intervention 2: £5
Intervention 3: £13
Intervention 4: £31
Intervention 5: £40
Intervention 6: £179
Intervention 7: £297
Intervention 8: £304
Intervention 9: £396
Intervention 10: £481
Intervention 11: £503
Intervention 12: £770
Intervention 13: £1,453

<u>Trials with adequate allocation concealment</u>
Intervention 1: £0
Intervention 2: £5
Intervention 3: £13
Intervention 4: £30
Intervention 5: NR
Intervention 6: £192
Intervention 7: £214</td>
<td>QALYs gained versus baseline (mean per patient):
<u>All trials</u>
Intervention 1: 0.000
Intervention 2: 0.001
Intervention 3: 0.001
Intervention 4: 0.011
Intervention 5: 0.001
Intervention 6: 0.014
Intervention 7: 0.005
Intervention 8: 0.008
Intervention 9: 0.011
Intervention 10: 0.005
Intervention 11: 0.007
Intervention 12: 0.033
Intervention 13: 0.007

<u>Trials with adequate allocation concealment</u>
Intervention 1: 0.000
Intervention 2: 0.000
Intervention 3: 0.002
Intervention 4: 0.005
Intervention 5: NR
Intervention 6: 0.017
Intervention 7: 0.003</td>
<td>Full incremental analysis[(c)] [(d)]:
<u>All trials</u>
<table>
<tr><th></th><th>Cost</th><th>QALYs</th><th>Inc. Cost</th><th>Inc. QALY</th><th>Cost per QALY</th><th>% most CE at £20 K</th></tr>
<tr><td>1</td><td>£0</td><td>0.000</td><td colspan="3">Baseline</td><td>0%</td></tr>
<tr><td>2</td><td>£5</td><td>0.001</td><td>£5</td><td>0.001</td><td>ED</td><td>22%</td></tr>
<tr><td>3</td><td>£13</td><td>0.001</td><td>£8</td><td>0.000</td><td>ED</td><td>0%</td></tr>
<tr><td>4</td><td>£31</td><td>0.011</td><td>£31</td><td>0.011</td><td>£2,690</td><td>49%</td></tr>
<tr><td>5</td><td>£40</td><td>0.001</td><td>£9</td><td>-0.01</td><td>D</td><td>6%</td></tr>
<tr><td>6</td><td>£179</td><td>0.014</td><td>£148</td><td>0.003</td><td>ED</td><td>6%</td></tr>
<tr><td>7</td><td>£297</td><td>0.005</td><td>£266</td><td>-0.006</td><td>D</td><td>0%</td></tr>
<tr><td>8</td><td>£304</td><td>0.008</td><td>£273</td><td>-0.003</td><td>D</td><td>0%</td></tr>
<tr><td>9</td><td>£396</td><td>0.011</td><td>£365</td><td>0.000</td><td>D</td><td>0%</td></tr>
<tr><td>10</td><td>£481</td><td>0.005</td><td>£450</td><td>-0.006</td><td>D</td><td>16%</td></tr>
<tr><td>11</td><td>£503</td><td>0.007</td><td>£472</td><td>-0.004</td><td>D</td><td>0%</td></tr>
<tr><td>12</td><td>£770</td><td>0.033</td><td>£739</td><td>0.022</td><td>£33,866</td><td>0%</td></tr>
<tr><td>13</td><td>£1,453</td><td>0.007</td><td>£683</td><td>-0.026</td><td>D</td><td>0%</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr>
<td>of analysis were combined with treatment and non-treatment-related costs.

Perspective: UK NHS

Time horizon/ treatment duration: 8 weeks

Discounting: n/a</td>
<td>Intervention 9: PES
Intervention 10: NMES
Intervention 11: Laser light therapy
Intervention 12: Interferential therapy
Intervention 13: PEMF</td>
<td>Intervention 8: £276
Intervention 9: £410
Intervention 10: NR
Intervention 11: £288
Intervention 12: £1,179
Intervention 13: £577

<u>Trials with adequate allocation concealment and an end point reported at 3-13 weeks</u>
Intervention 1: £0
Intervention 2: £5
Intervention 3: £14
Intervention 4: £30
Intervention 5: NR
Intervention 6: £192
Intervention 7: £213
Intervention 8: £277
Intervention 9: £410
Intervention 10: NR
Intervention 11: £288
Intervention 12: £1,179
Intervention 13: £277

For incremental analyses see cost effectiveness column</td>
<td>Intervention 8: 0.013
Intervention 9: 0.010
Intervention 10: NR
Intervention 11: 0.003
Intervention 12: 0.016
Intervention 13: 0.008

<u>Trials with adequate allocation concealment and an end point reported at 3-13 weeks</u>

Intervention 1: 0.000
Intervention 2: -0.001
Intervention 3: 0.004
Intervention 4: 0.006
Intervention 5: NR
Intervention 6: 0.017
Intervention 7: 0.002
Intervention 8: 0.018
Intervention 9: 0.010
Intervention 10: NR
Intervention 11: 0.003
Intervention 12: 0.017
Intervention 13: 0.007

For incremental analyses see cost effectiveness column</td>
<td><u>Trials with adequate allocation concealment</u>[e]
<table><tr><th></th><th>Cost</th><th>QALYs</th><th>Inc. Cost</th><th>Inc. QALY</th><th>Cost per QALY</th><th>% most CE at £20 K</th></tr><tr><td>1</td><td>£0</td><td>0.000</td><td colspan="3">Baseline</td><td>0%</td></tr><tr><td>2</td><td>£5</td><td>0.000</td><td>£5</td><td>0.000</td><td>D</td><td>26%</td></tr><tr><td>3</td><td>£13</td><td>0.002</td><td>£13</td><td>0.002</td><td>ED</td><td>4%</td></tr><tr><td>4</td><td>£30</td><td>0.005</td><td>£30</td><td>0.005</td><td>£6,142</td><td>15%</td></tr><tr><td>6</td><td>£192</td><td>0.017</td><td>£162</td><td>0.012</td><td>£13,502</td><td>47%</td></tr><tr><td>7</td><td>£214</td><td>0.003</td><td>£22</td><td>-0.014</td><td>D</td><td>0%</td></tr><tr><td>8</td><td>£276</td><td>0.013</td><td>£84</td><td>-0.004</td><td>D</td><td>7%</td></tr><tr><td>11</td><td>£288</td><td>0.003</td><td>£96</td><td>-0.014</td><td>D</td><td>0%</td></tr><tr><td>9</td><td>£410</td><td>0.010</td><td>£218</td><td>-0.007</td><td>D</td><td>0%</td></tr><tr><td>13</td><td>£577</td><td>0.008</td><td>£385</td><td>-0.009</td><td>D</td><td>0%</td></tr><tr><td>12</td><td>£1,179</td><td>0.016</td><td>£987</td><td>-0.001</td><td>D</td><td>0%</td></tr></table>
<u>Trials with adequate allocation concealment and an end point reported at 3-13 weeks</u>[e]
<table><tr><th></th><th>Cost</th><th>QALYs</th><th>Inc. Cost</th><th>Inc. QALY</th><th>Cost per QALY</th><th>% most CE at £20 K</th></tr><tr><td>1</td><td>£0</td><td>0.000</td><td colspan="3">Baseline</td><td>0%</td></tr><tr><td>2</td><td>£5</td><td>-0.001</td><td>£5</td><td>-0.001</td><td>D</td><td>17%</td></tr><tr><td>3</td><td>£14</td><td>0.004</td><td>£14</td><td>0.004</td><td>£3,540</td><td>13%</td></tr><tr><td>4</td><td>£30</td><td>0.006</td><td>£16</td><td>0.002</td><td>£9,750</td><td>25%</td></tr><tr><td>6</td><td>£192</td><td>0.017</td><td>£162</td><td>0.011</td><td>£14,275</td><td>25%</td></tr></table></td>
</tr>
</table>

		Currency & cost year: 2011/12 UK pounds. **Cost components incorporated:** Physiotherapist's time to conduct weekly sessions, except for TENS, where patients self-administered after an initial physiotherapist visit. Changes in non-treatment-related visits to GPs and specialists arising from changes in EQ-5D score.		

7	£213	0.002	£21	-0.015	D	0%
8	£277	0.018	£85	0.001	£86,964	20%
13	£277	0.007	£0	-0.011	D	0%
11	£288	0.003	£11	-0.015	D	0%
9	£410	0.010	£133	-0.008	D	0%
12	£1,179	0.017	£902	-0.001	D	0%

Analysis of uncertainty:

TENS was the most cost-effective alternative at a £20K threshold when a linear relationship were assumed between EQ-5D treatment effect and session duration. When all the treatment benefit were assumed in the first 20/30 minutes of the session, interferential therapy was the most cost-effective option.

In an analysis of all trials, TENS remained the most cost-effective option when the duration of treatment benefit were extended by 50%.

Data sources

Health outcomes: Study-level reported mean differences in pain as a measure of treatment effectiveness were standardised to the EQ-5D measure for each of the three network meta analyses. **Quality-of-life weights:** Generic EQ-5D quality-of-life scores were mapped from the SF-12 & SF-36 surveys, pain NRD, pain VAS and WOMAC scales. **Cost sources:** The cost to the NHS (physiotherapists time, GP and specialists' consultations) was obtained from the Personal Social Services Research Unit 2012. Equipment administered by physiotherapists (e.g., devices) were not included as the per-patient costs as these were expected to be small. **Resource use:** Estimates of resource use were based on consultations with clinical experts and published literature including trial data and NHS data. Treatment duration was based on a weighted average of the clinical trial data.

Comments

Source of funding: National Institute for Health Research (NIHR). **Limitations:** Unit costs taken from 2011/12 may not reflect current UK NHS practice. The time horizon was only 8 weeks. Adverse events and their downstream consequences were not considered. **Other:** Non-treatment-specific healthcare resource use was assumed to be a function of change in EQ-5D and was taken from the TOIB trial. TENS machine assumed to last for 1 year.

Overall applicability:[(a)] Partially applicable **Overall quality:**[(b)] Potentially serious limitations

Abbreviations: CE= cost effective; CI = confidence interval; CUA = cost-utility analysis; D= dominated; ED= extendedly dominated; EQ-5D = Euroqol 5 dimensions (scale: 0.0 [death] to 1.0 [full health], negative values mean worse than death); GP= general practitioner; ICER = incremental cost-effectiveness ratio; Inc.= incremental; K= thousand; n/a = not applicable; NHS = National Health Service; NMES= neuromuscular electrical stimulation; NR = not reported; NRS = numeric rating scale; OA = Osteoarthritis; PEMF= pulsed

electromagnetic field; PES= pulsed electrical stimulation; QALYs = quality-adjusted life years; SF-12 = short-form health survey 12 items; SF-36= short-form health survey 36 items; TENS= transcutaneous electrical nerve stimulation; UK= United Kingdom; VAS = visual analogue scale; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index.

(a) Directly applicable / Partially applicable / Not applicable

(b) Minor limitations /Potentially serious limitations / Very serious limitations

(c) Intervention number in order of least to most costly (in terms of cost)

(d) Full incremental analysis of available strategies: first strategies are ruled out that are dominated (another strategy is more effective and has lower costs) or subject to extended dominance (the strategy is more effective and more costly but the incremental cost effectiveness ratio is higher than the next most effective option and so it would never be the most cost effective option); incremental costs, incremental effects and incremental cost effectiveness ratios are calculated for the remaining strategies by comparing each to the next most effective option.

(e) Interventions 5 and 10 not available because these intervention did not provide information to network meta analysis

Study	Reinhold 2008[107]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome = QALYs) **Study design:** Within-trial analysis **Approach to analysis:** Analysis of individual level quality of life and resource use data adjusted for age, gender, diagnosis, utility at study initiation, costs before study initiation. Unit costs applied. **Perspective:** German societal perspective – healthcare costs reported separately and reported here. **Time horizon:** 12 months **Treatment effect duration:** 12 months (treatment was for 3 months after which utility gradually began declining, to baseline at 12 months.) **Discounting:** n/a	**Population:** Patients over 40 years of age with chronic pain (defined as more than 6 months) due to osteoarthritis of the knee or hip. **Patient characteristics:** Age = 61 Male = 40% N = 489* **Intervention 1:** Delayed acupuncture (3 months) **Intervention 2:** Acupuncture (between 10-15 sessions) *cost effectiveness analysis only included 219 acupuncture patients and 202 non acupuncture patients due to missing QoL data.	**Total costs (mean per patient):** Intervention 1: NR Intervention 2: NR Incremental (2-1): £353 (CI = £58 - £648; p=NR) **Currency & cost year:** 2006 Euros (presented here as 2006 UK pounds[(a)]) **Cost components incorporated:** Acupuncture costs, physician visits, medication, hospital stays (and indirect costs of lost workdays).	**QALYs (mean per patient):** Intervention 1: NR Intervention 2: NR Incremental (2-1): 0.0241 (CI = NR; p=NR)	**Cost per QALY gained (Intervention 2 vs. Intervention 1):** £13,944 CI: NR Probability cost-effective (at £20,000 per QALY gained)[(b)]: 85% **Cost per QALY gained (Intervention 2 vs. Intervention 1); Osteoarthritis- specific costs only:** £13,238 **Subgroup analyses:** *Cost per QALY gained (Intervention 2 vs. Intervention 1);* Knee = £17,019; Hip = £6,582 Male (all diagnoses) = £46,767; Female (all diagnoses)= £9,228 *Osteoarthritis-specific costs only :* Knee = £14,332; Hip = £10,196 Male = £26,828; Female = £9,863 **Analysis of uncertainty:** Sensitivity analysis showed that the parameters which had the largest effect were the cost of acupuncture, and the effect duration.

Data sources
Health outcomes: This investigation was part of the Acupuncture in Routine Care (ARC) studies.[469] The Reinhold paper itself was excluded from the clinical review on the grounds that clinical outcomes do not differentiate between patients with hip or knee OA (although patients were differentiated by joint type for the economic analysis). Additionally, it is the same trial used in another study [470] (included in the clinical review), therefore its inclusion would also have been a duplication of data. **Quality-of-life weights:** For the cost effectiveness analysis, quality of life data from the SF-36 questionnaire were converted to SF-6D using the Brazier algorithm. **Cost sources:** Costs were taken from health insurance funds data. These included the direct health related costs of acupuncture; cost of an acupuncture session (€35, wasn't reimbursed by social health insurance at the time of the study), physician visits and hospital stays, drugs prescribed (including patient co-payments) and indirect costs caused by lost workdays. Data on resource use was based on data maintained by the social health insurance funds.
Comments
Source of funding: Several German social health insurance companies. **Limitations:** Short time horizon. Health outcome mapped onto a utility measure. Cost of lost workdays were included in the cost of intervention. **Other:** Assumptions - A linear decrease in acupuncture effects after the intervention period of 3 months, returning to baseline 12 months after study onset. Note that patients were free to use conventional routine medical care as offered by the German social insurance funds. Total costs for each arm were reported for the total number of patients in the trial, but as not all the patients took part in the cost effectiveness analysis, then costs for these differs and only the increments were reported. The mean cost difference between the two groups was primarily due to the acupuncture costs. The diagnosis-specific cost effectiveness analysis was used ICD-10 codes to identify costs due only to OA pain and related conditions.
Overall applicability:(c) Partially applicable **Overall quality:**(d) Potentially serious limitations

Abbreviations: CI = 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); ICD-10 = World Health Organisation's International Classification of Diseases version 10; ICER = incremental cost-effectiveness ratio; NR = not reported; n/a = not applicable; OA = Osteoarthritis; QALYs =quality-adjusted life years; QoL = quality of life; SF-6D = short form questionnaire six dimensions; SF-36 = short form 36 questionnaire; UK= United Kingdom.

(a) Converted using 2006 purchasing power parities. Organisation for Economic Co-operation and Development (OECD). [101]

(b) Read off the graph (at approx. €24,000)

(c) Directly applicable / Partially applicable / Not applicable;

(d) Minor limitations /Potentially serious limitations / Very serious limitations

Study	Whitehurst (2011)[153]			
Study details	**Population & interventions**	**Costs**	**Health outcomes**	**Cost effectiveness**
Economic analysis: CUA (health outcome = QALYs) **Study design:** Within-trial analysis **Approach to analysis:** **Perspective:** UK NHS **Time horizon:** 12 months **Treatment effect duration:** 12 months (treatment was for 6 weeks after which utility gradually began declining, but still above baseline at 12 months.) **Discounting:** n/a	**Population:** Patients aged 50 years or older who had been referred to NHS physical therapy centres with a clinical diagnosis of knee OA. **Patient characteristics:** Age = NR Male = NR N = 352 **Intervention 1:** Advice + exercise 6 sessions over 6 weeks plus a leaflet and a home exercise program. **Intervention 2:** Advice, exercise + acupuncture (AE+A) Same as intervention 1 plus acupuncture at traditional Chinese acupuncture points.	**Total costs (mean per patient):** Intervention 1: £229 Intervention 2: £314 Incremental (2-1): £85 (CI = 41 to 129; p=NR) **Currency & cost year:** 2004-5 UK pounds. **Cost components incorporated:** Questionnaires collected data on consultations with primary care-based practitioners, hospital consultants (outpatient attendance) or any other health care provider. Participants were also asked to report any prescribed medications and over the counter purchases.	**QALYs (mean per patient):** Intervention 1: NR Intervention 2: NR Incremental (2-1): 0.022 (CI = -0.03 to 0.07; p= 0.37)	**Cost per QALY gained (Intervention 2 vs. Intervention 1):** £3,889 Probability cost-effective at a threshold of £20,000 per QALY gained: 77% **Analysis of uncertainty:** The following sensitivity analyses were carried out (using data from the sample of participants who had complete resource use and EQ-5D data): 1. A complete case analysis explored the implication of missing data. This had a cost per QALY of £2,278. 2. A cost perspective that incorporated non-NHS health care resource use (had little effect on the results). 3. An analysis of the AE+NPA group within the base case. AE+NPA was associated with an additional cost of £36 and an incremental QALY of 0.001 compared with AE+A. Thus ICER (Intervention 3 vs. Intervention 2): £36/ 0.001 = £36,000. This implies that needle penetration is not essential for the effect given

	Intervention 3: Advice, exercise + non-penetrating acupuncture (AE+NPA) Same as intervention 1 plus acupuncture with non-penetrating blunt-tip needles. (Intervention 3 was only included in the sensitivity analysis.)			the small differences observed between the two interventions.

Data sources

Health outcomes: This is cost utility analysis reported alongside a randomised controlled trial, the results of the trial are reported elsewhere: [151] **Quality-of-life weights:** Utility was measured using the EQ-5D questionnaire at baseline, 6 weeks, 6 months and 12 months. **Cost sources:** NHS care was costed as standardised national averages, using data obtained from the Personal Social Services Research Unit, NHS reference costs, and the British National Formulary. Acupuncture sessions were assigned a unit cost associated with NHS community physical therapy. The cost of traditional acupuncture needles is negligible and was not accounted for in the analysis. Resource use data was collected from study therapists (the number and content of treatment sessions attended by participants), and knee OA related resource use data were collected from self-report questionnaires. At each follow up, patients were asked to recall the time period since their last questionnaire, and the total cost estimate resulted from the summation of data from the distinct time periods.

Comments

Source of funding: The clinical trial was supported by Project Grant H0640 from Arthritis Research UK and Support for Science funding secured by North Staffordshire Primary Care Research Consortium for NHS service support costs. **Limitations:** 2004/05 resource use and unit costs may not reflect current UK NHS practice. Time horizon could be longer. **Other:** Treatments were delivered over a 6-week period. 'Traditional Chinese acupuncture points' were used, and needles were manipulated to achieve needle sensations. Patients were informed that they 'may receive acupuncture, using 1 of 2 different types of acupuncture needle'.

Note that no inpatient data were collected since, given the lengthy waiting list for orthopaedic consultations it was unlikely that participants would have proceeded to receive hospital-based treatment during the follow-up period. Costing assumptions for the study interventions and consultations with other health care professionals were based on average session times for the study interventions and clinical judgement of current practice. Costs and QALYs were calculated for both the observed sample, and the imputed sample (which includes all participants; even those for which there was missing data; and this was imputed based on the last observed value). The ICER reported in this table is based on the imputed sample, as this was the ICER reported in the paper.

Overall applicability:[(a)] Directly applicable **Overall quality:**[(b)] Potentially serious limitations

Abbreviations: AE+A = Advice, exercise + acupuncture; AE+NPA = Advice, exercise + non-penetrating acupuncture; CI = 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; NHS = National Health Service; n/a = not applicable; NR = not reported; OA = Osteoarthritis; QALYs = quality-adjusted life years; UK = United Kingdom.

(a) Directly applicable / Partially applicable / Not applicable

(b) Minor limitations /Potentially serious limitations / Very serious limitations

Appendix I – Excluded studies

Clinical studies

Table 24: Studies excluded from the clinical review

Study	Exclusion reason
Ahsin 2009[2]	No usable outcomes (no standard deviation reported and no way to calculate this from the information available)
Ahsin 2010[1]	Not available
Al rashoud 2014[3]	Incorrect interventions (laser acupuncture which is included in the electrotherapy review)
Ammer 1988[4]	Non-English language study
Appleyard 2016[5]	Inappropriate comparison (compares two types of acupuncture techniques)
Ashraf 2014[6]	Incorrect interventions (included lateral wedge insoles which are not included in the protocol)
Atalay 2021[7]	Inappropriate comparison (acupuncture vs physiotherapy)
Bao 2007[8]	Incorrect interventions (included diclofenac which is not included in the protocol)
Brinkhaus 2006[13]	Non-English language study
Byun 2007[14]	Inappropriate comparison (compares two types of acupuncture techniques)
Cao 2012[15]	Systematic review; references checked (inadequate quality assessment)
Chen 2017[20]	Systematic review: study designs inappropriate
Chen 2020[21]	Non-English language study
Christensen 1992[22]	No usable outcomes (no standard deviation reported and no way to calculate this from the information available)
Corbett 2013[23]	Systematic review is not relevant to review question or unclear PICO (different definition of outcomes, different definition of no treatment)
Dickens 1989[24]	Inappropriate comparison (compares acupuncture to sham transcutaneous electrical nerve stimulation [TENS] which was not included in the protocol)
Ding 2016[25]	Incorrect interventions (included acupotomy which was not included in the protocol)
Elbadawy 2017[27]	Incorrect interventions (included TENS which was not included in the protocol).
Endres 2007[28]	Not review population (included people with chronic back pain). Studies with an unclear population (e.g. type of arthritis, proportion of participants with osteoarthritis).
Ernst 1997[29]	Spinal osteoarthritis
Fargas-babjak 1992[31]	Spinal osteoarthritis. Not review population (included people with chronic pain syndrome).
Fink 1996[33]	Non-English language study
Fink 2000[35]	Non-English language study
Fink 2000[36]	Non-English language study

Fink 2001[32]	Non-English language study
Foster 2021[38]	Individual patient data systematic review that reports osteoarthritis studies separately but uses a different definition on the outcomes then that used in the review. To maintain consistency this was excluded but the references were checked.
Fu 2012[39]	Incorrect interventions (included western medicine which was not included in the protocol)
Gaw 1975[40]	Spinal osteoarthritis
Gollub 2018[41]	Inappropriate comparison (compared a high expectancy group to a low expectancy group using functional neuroimaging to stratify, the outcomes reported are not relevant or are not based on patient validated scales)
Gong 2019[42]	Incorrect interventions (included acupressure which was considered in the manual therapy review)
Green 2005[43]	Not review population (included people with shoulder pain, which could be caused by other non-osteoarthritis conditions)
Grotle 2011[44]	Commentary only
Haslam 2001[45]	Incorrect interventions (included exercise and advice which was not included in the protocol)
Helianthi 2016[46]	Incorrect interventions (laser acupuncture which is included in the electrotherapy review)
Hou 2020[48]	No usable outcomes
Huang 2012[49]	Spinal osteoarthritis
Itoh 2008[50]	Spinal osteoarthritis
Itoh 2008[51]	No usable outcomes (no standard deviation reported and no way to calculate this from the information available)
Jia 2005[52]	Non-English language study
Jia 2020[53]	No usable outcomes (outcomes reported as median values with interquartile ranges or confidence intervals only)
Jubb 2008[55]	Incorrect interventions (combined electroacupuncture and non-electroacupuncture)
Jun 2018[56]	Incorrect interventions (included miniscalpal acupuncture which was not included in the protocol)
Karner 2013[57]	No usable outcomes (no standard deviation reported and no way to calculate this from the information available)
Kim 2010[58]	Incorrect interventions (included pharmacopuncture which was not included in the protocol)
Kim 2012[60]	Not review population (included people with low back pain and headaches)
Kim 2013[59]	Protocol only
Kwon 2001[62]	Inappropriate comparison (included bee venom acupuncture which was not included in the protocol)
Lam 2021[63]	Duplicate reference
Lee 2009[66]	Not review population (included people with various different conditions, including Parkinson's disease and low back pain). Incorrect interventions (included constitutional medicine which is not included in the protocol)
Li 2015[70]	Non-English language study
Li 2018[68]	Incorrect interventions (included acupressure which is considered in the manual therapy review)

Li 2018[69]	Inappropriate comparison (compared acupuncture to other types of acupuncture)
Liang 2019[71]	Non-English language study
Lim 2006[72]	Non-English language study
Lin 2016[75]	Systematic review; references checked (inadequate quality assessment)
Lin 2020[74]	Inappropriate comparison (compared acupuncture delivered three times a week to acupuncture delivered once a week)
Linde 2007[76]	Secondary analysis of RCTs
Lorenc 2018[77]	Not review population (included any musculoskeletal condition). Incorrect interventions (included all forms of complementary medicine)
Lu 2010[78]	Less than minimum duration (<1 week)
Lue 2017[79]	Incorrect interventions (included a variety of non-surgical therapies)
Lundeberg 1991[80]	Spinal osteoarthritis
Luo 2019[81]	Non-English language study
Lv 2019[82]	Duplicate reference (Lv 2019 [83])
Maa 2008[84]	Incorrect study design (non-randomised)
Manheimer 2007[89]	Systematic review; references checked (inadequate quality assessment)
Manheimer 2010[86]	Inappropriate comparison (different definition of no treatment/usual care). Cochrane review; references checked
Manheimer 2018[87]	Inappropriate comparison comparison (different definition of no treatment/usual care). Cochrane review; references checked
Manyanga 2014[90]	Systematic review; references checked (inadequate quality assessment)
Mcindoe 1995[93]	Incorrect interventions (included intra-articular steroids which are not included in the protocol)
Meng 2009[94]	Inappropriate comparison (compared different types of acupuncture)
Min 2009[96]	Non-English language study
Molsberger 1993[97]	Non-English language study
Ng 2003[99]	No usable outcomes (no standard deviation reported and no way to calculate this from the information available)
Nie 2015[100]	Non-English language study
Penagos-martinez 2021[102]	Non-English language study
Petrou 1988[103]	No usable outcomes (reports outcomes using non-validated patient scales)
Plaster 2014[104]	Less than minimum duration (<1 week)
Rahou-el-bachiri 2020[105]	Systematic review; references checked
Rodriguez-merchan 2016[108]	Systematic review is not relevant to review question or unclear PICO (review of Cochrane reviews)
Saleki 2013[109]	Incorrect interventions (included exercise which was not included in the protocol)
Selfe 2008[114]	Systematic review; references checked (inadequate quality assessment)
Shafshak 1995[115]	No usable outcomes (no outcomes reported which are usable in the protocol)

Shen 2009[116]	Incorrect interventions (no standard deviation reported and no way to calculate this from the information available)
Sheng 2015[117]	Non-English language study
Shim 2016[118]	Systematic review; references checked (inadequate quality assessment)
Singh 2001[119]	No usable outcomes (crossover study that only reports outcomes after all participants have received acupuncture)
Song 2004[120]	No usable outcomes (no outcomes reported which are usable in the protocol)
Soni 2012[121]	Inappropriate comparison (compared acupuncture and physiotherrapy to standardised exercise and education leaflet)
Stener-victorin 2004[122]	Incorrect interventions (included a treatment package which is considered in the treatment package review)
Suen 2016[125]	Erratum only
Taechaarpornkul 2009[126]	Inappropriate comparison (compares two different methods for giving acupuncture)
Tang 2018[128]	Protocol only
Thomas 1991[129]	Spinal osteoarthritis
Tillu 2002[130]	Incorrect study design (non-randomised)
Trinh 2003[131]	Not review population. Systematic review is not relevant to review question or unclear PICO (the aim of the study was to look at the blinding of acupuncture being delivered in the trials)
Tu 2019[133]	Erratum only
Tu 2021[132]	Non-English language study
Tukmachi 2004[136]	Inappropriate comparison (compared acupuncture and medicine which is not included in the protocol)
Ughreja 2021[137]	Not available (order cancelled systematic review unlikely to included in the review)
Vickers 2012[140]	Not review population (included people with back and neck pain, chronic headache and shoulder pain)
Vickers 2018[141]	Not review population (included people with back and neck pain, chronic headache and shoulder pain)
Wang 2020[142]	Non-English language study
Wang 2020[143]	No usable outcomes
Wang 2020[144]	Inappropriate comparison (thumb-tack needling vs medication)
Weiner 2013[148]	Wrong unit of randomisation (knee)
White 2007[150]	Systematic review; references checked (inadequate quality assessment)
White 2010[152]	Systematic review; references checked (inadequate quality assessment)
White 2016[11]	Incorrect interventions (included a mixture of electroacupuncture and non-electroacupuncture)
Witt 2019[157]	Not review population (included people with chronic headache, migraine, back, neck and shoulder pain)
Woods 2017[106]	Systematic review is not relevant to review question or unclear PICO (different definition of outcomes, different definition of no treatment)
Wu 2008[158]	Non-English language study
Xi 2008[159]	Non-English language study
Xu 2007[160]	Non-English language study

Yurtkuran 2007[161]	Incorrect interventions (laser acupuncture which is included in the electrotherapy review)
Zhang 2016[165]	Inappropriate comparison (compares acupuncture to physiotherapy which was not included in the protocol)
Zhang 2017[164]	Not review population (included people with rheumatoid arthritis)
Zhang 2020[163]	Protocol only
Zhen 2004[166]	Non-English language study
Zhou 2008[167]	Not available

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.

None.

Appendix J – Research recommendations – full details

J.1.1 Research recommendation

Is electroacupuncture a clinically and cost-effective treatment for any subgroup of people with osteoarthritis?

J.1.2 Why this is important

Evidence for electroacupuncture was inconsistent, with some potential benefits being identified in this review. However, the benefits were not consistent for all people with osteoarthritis across all of the trials. Therefore, to find the people who would benefit from electroacupuncture, a study to investigate the clinical and cost-effectiveness of electroacupuncture for different treatment groups would be beneficial.

J.1.3 Rationale for research recommendation

Importance to 'patients' or the population	Electroacupuncture may provide clinically important benefits for some people with osteoarthritis or be an effective treatment option in some groups for whom other treatments are not possible. However, it is unclear for which subgroups of people electroacupuncture is effective and indicated. Therefore, to identify who would respond best further research would be required.
Relevance to NICE guidance	The committee concluded that they could not make a recommendation discussing the use of electroacupuncture due to the low quality of evidence. Therefore, additional evidence is required before further recommendation can be made. Correctly identifying populations who may benefit from electroacupuncture may provide evidence to support this.
Relevance to the NHS	Providing electroacupuncture widely in the NHS would be associated with significant costs therefore, identifying the people who would optimise use of healthcare resources and cost effectiveness.
National priorities	This is not an area of national priority.
Current evidence base	Current evidence for electroacupuncture is mostly in the short-term and is heterogenous in nature often being made up of trials with a small number of participants. Therefore, research including larger adequately powered randomised controlled trials that allow the investigation of a-priori subgroup effects would be beneficial.
Equality considerations	This research will consider specific groups that may have increased adverse effects from other treatments, such as people with comorbidities and older people (over the age of 75 years). The committee noted that the research identified in this review 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.

J.1.4 Modified PICO table

Population	Inclusion: • Adults (age ≥16 years) with osteoarthritis affecting any joint Stratifications such as: • People in whom pharmacological treatment has been ineffective • People who are receiving large amounts of psychotropic analgesia • People in whom other treatments (for example: pharmacological, surgery) have been assess to have a significant risk 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
Intervention	Electroacupuncture
Comparator	Acupuncture (same positions and needling characteristics but without the electrical stimulation) Attention control (equivalent time spent with a healthcare professional including counselling about condition) Usual care (no acupuncture, no additional healthcare professional contact related to the study)
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 data] • Serious adverse events [dichotomous data]
Study design	Randomised control trial
Timeframe	Long term
Additional information	Adequately powered high quality randomised controlled trials. Trials with sufficient blinding, adequate randomisation methods and allocation concealment. Subgroup analyses: • Presence of multimorbidity (high versus low morbidity score) • Age (≤/> 75 years)