Centre for Reviews and Dissemination

Acupuncture and other physical treatments for the relief of chronic pain due to osteoarthritis of the knee: a systematic review and network meta-analysis

REPORT 40

THE UNIVERSITY of York

Acupuncture and other physical treatments for the relief of chronic pain due to osteoarthritis of the knee: a systematic review and network meta-analysis

Mark Corbett[1]
Stephen Rice[1]
Russell Slack[1]
Melissa Harden[1]
Vichithranie Madurasinghe[1]
Alex Sutton[2]
Hugh MacPherson[3]
Philip Conaghan[4]
Nerys Woolacott[1]

[1] Centre for Reviews and Dissemination (CRD), University of York
[2] Department of Health Sciences, University of Leicester
[3] Department of Health Sciences, University of York
[4] NIHR Leeds Musculoskeletal Biomedical Research Unit, University of Leeds

March 2012

The Centre for Reviews and Dissemination is part of the National Institute for Health Research and is a department of the University of York. The Centre undertakes high-quality systematic reviews that evaluate the effects of health and social care interventions and the delivery and organisation of health care.

ACKNOWLEDGEMENTS

This report presents independent research commissioned by the National Institute for Health Research (NIHR) under Programme Grants for Applied Research (Grant No. RPPG-0707-10186). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. All researchers were independent of the funders. The funders had no role in study design, data collection, data synthesis, data interpretation, or writing the report. Any errors are the responsibility of the authors.

Thanks to Mark Roman, Peter Hall, Andrea Manca, Cynthia Iglesias, and Ann Hopton for their help at various stages of the review, particularly during protocol development and to Tony Danso-Appiah for his assistance with data-extraction.

Table of Contents

Contents of Tables

Contents of Figures

DEFINITION OF TERMS AND LIST OF ABBREVIATIONS

List of Abbreviations

Acu	Acupuncture
Ae Ex	Aerobic exercise
AIMS	Arthritis Impact Measurement Scale
Bal, BAL	Balneotherapy
BMI	Body mass index
Bra, BRA	Braces
CI	Confidence interval
CrI	95% credible interval
EQ-5D	A self-reported generic preference-based measure of health
Ex	Home exercise and/or education (as an adjunct treatment)
ExAe	Exercise - Aerobic (weight bearing)
ExMu	Exercise - Muscle strengthening (non-weight bearing)
GP	General practitioner (primary care physician)
HADS	Hospital Anxiety and Depression Scale
Hea, HEA	Heat treatment
Ice, ICE	Ice/cooling treatment
Ins, INS	Insoles
Int, INT	Interferential therapy
KOOS	Knee injury and Osteoarthritis Outcome Score
Las, LAS	Laser/light therapy
MACTAR	McMaster Toronto Arthritis
Mag, MAG	Static magnets
Man, MAN	Manual therapy
MCIC	minimal clinically important change
MCII	minimal clinically important improvement
Muscle str. exercise	Exercise - Muscle strengthening (non-weight bearing)
NHS	National Health Service (UK)
NICE	National Institute for health and Clinicla excellence
NMA	Network meta-analysis
NMES	neuromuscular electrical stimulation
NoMed	No medication
NoMed+EX	No medication + exercise/ education
NoTr	No treatment
NSAIDs	Non-steroidal anti-inflammatory drugs
OA	osteoarthritis
P	Placebo
PEMF	pulsed electromagnetic fields
PES	pulsed electrical stimulation
PPI (pain scale)	Present pain intensity (pain scale)
RCT	randomised controlled trial
SC	Standard care
SD	Standard deviation
ShAcu	Sham acupuncture
SMD	standardised mean differences
Tai, TAI	Tai Chi
TENS	transcutaneous electrical nerve stimulation
UK	United Kingdom
UT	Treatment as usual/ Unclear
UT+ EX	Treatment as usual/unclear + exercise/education
UT+AN	Treatment as usual/unclear + analgesia
VAS	visual analogue scale
Wei, WEI	Weight loss
WOMAC	Western Ontario and MacMaster Universities

GLOSSARY

Grouped Interventions set
The Grouped interventions set of comparators was defined according to the alternatives available to a general practitioner in the NHS after the initial prescription of medication and advice. This was motivated by the requirements of the economic evaluation part of the project. Many of the main interventions were considered to be utilised within a physiotherapy session as part of the repertoire of a physiotherapist, and the GP's choice was whether or not to refer to a physiotherapist. For this reason, these main treatments were collectively defined as physiotherapy**.**

Kellgren and Lawrence scores
Kellgren and Lawrence scores are generated by the classification for osteoarthritis (OA) described by Kellgren and Lawrence. This classification is the most widely used radiological classification to identify and grade OA. Kellgren and Lawrence defined OA in five grades (0, normal to 4, severe).

Likert scale
An interval-based multiple-choice style of question used in questionnaires.

Therapy-only Interventions set
The Therapy-only intervention set was defined such that any adjunct treatments for an intervention were grouped together. So there were 22 possible individual interventions in the analysis.

Therapy-plus-adjunct Interventions set
The Therapy-plus-adjunct interventions set was defined to evaluate plausible differences in treatment effect between the competing interventions as defined at the main intervention plus adjunctive therapy level, e.g. acupuncture plus standard care versus acupuncture plus home exercise.

WOMAC
WOMAC is a widely used self-administered health status measure that assesses the dimensions of pain, stiffness, and function in patients with OA of the hip or knee; it is available in 5-point Likert, 11-point numerical rating, and 100 mm visual analogue scale formats. Under each dimension there are number of questions designed to assess the clinical severity of the disease (5 questions for pain, 2 questions for stiffness and 17 questions for physical function). The patient's response to each question produces a score which is then added up to derive an aggregated score for each dimension. It produces three subscale scores (pain, stiffness, and physical function) and a total score (WOMAC index), which reflects disability overall.

WOMAC index
A score that reflects disability overall.

1. EXECUTIVE SUMMARY

1.1 Background

To control the pain of osteoarthritis of the knee general practitioners might consider acupuncture as an alternative to drug treatments and/or as an adjunct to advice about exercise and weight loss, as recommended by NICE. Alternative physical treatments to acupuncture include the many types of treatment administered by a physiotherapist, exercise programmes, and footwear insoles. Many reviews have evaluated individual types of physical treatments for osteoarthritis of the knee, but no review using network meta-analysis methods has attempted to address the question of how effective such treatments are relative to each other.

1.2 Objectives

The purpose of this systematic review therefore, was to synthesise both the indirect and direct evidence, using network meta-analysis methods, in order to compare the effectiveness of acupuncture with other relevant physical treatments for alleviating pain due to osteoarthritis in patients requiring additional or alternative therapy to pharmacological analgesics.

1.3 Methods

The review processes and methods of analysis were specified in advance and documented in a protocol.

Literature search

We searched 17 electronic databases (15 for primary studies) from inception to June 2010. A combination of relevant free text terms, synonyms and subject headings relating to osteoarthritis of the knee and named physical therapies were included in the strategy. A search filter was used to limit retrieval of studies to randomised controlled trials (RCTs). No language or date restrictions were applied. Bibliographies of all relevant reviews and guidelines were checked for further potentially relevant studies, and internet searches were made of websites relating to osteoarthritis.

Study selection

We included RCTs which assessed pain in adults with osteoarthritis of the knee (where the mean age of the population was ≥55 years) after treatment with any of the following: acupuncture, balneotherapy, braces, aerobic exercise, muscle strengthening exercise, heat treatment, ice/cooling treatment, insoles, interferential therapy, laser/light therapy, manual therapy, neuromuscular electrical stimulation (NMES), pulsed electrical stimulation (PES), pulsed electromagnetic fields (PEMF), static magnets, Tai Chi, transcutaneous electrical nerve stimulation (TENS), and weight loss. These interventions could be given in addition to standard care. Eligible comparators included: standard care (which could incorporate one or more of analgesics, education, and exercise/advice), placebo interventions, no intervention, and sham acupuncture. It was anticipated that pain would be measured using a variety of measures: all scales were eligible.

Studies comparing only different regimens/durations/modalities of the same type of intervention were excluded, as were interventions which combined two or more physical treatments. Exercise interventions which were predominantly home-based, and unsupervised, were excluded. Two reviewers independently screened all abstracts, and then all relevant full papers, with disagreements resolved by discussion, or by a third reviewer when necessary.

Data extraction and assessment of trial quality

Using a standardised data extraction form data were extracted on: population characteristics (population type, method of diagnosis, age, sex, weight, BMI, Kellgren & Lawrence score), intervention parameters and study quality. Interventions were categorised both with (e.g. ‘acupuncture plus treatment as usual, with specified analgesics’) and without (e.g. ‘acupuncture’) the recording of any adjunct treatments. The five adjunct categories used were: ‘treatment as usual’, ‘treatment as

usual' plus specified home exercise or education, 'treatment as usual' plus specified analgesics, no medication, and no medication plus specified home exercise or education. Data from the end of treatment pain assessment, and from all subsequent time points, were extracted onto an Excel spreadsheet.

Trial quality was assessed and based upon the number of criteria satisfied, studies were then graded as excellent, good, satisfactory or poor. Data extraction and quality assessments were performed by one reviewer and checked by a second reviewer.

Outcomes and data transformations

Since a variety of pain scales were used, Hedges-g standardised mean differences (SMDs) were calculated for the meta-analyses (studies reporting medians were excluded from our analyses). Standard deviations and patient numbers were imputed where possible. Different doses/regimens of the same type of treatment within a study were pooled. Final values were used in the analysis in order to maximise the evidence available, and to avoid the need to make assumptions about within-patient correlation between baseline and final values, which the use of change from baseline data would have necessitated.

Synthesis

Pair-wise meta-analyses were conducted using outcomes recorded at the end of treatment only. They were not intended as a comprehensive stand-alone synthesis, but as a means of informing and complementing the network meta-analysis. In particular, they investigated the within-intervention clinical and statistical heterogeneity. Where enough studies were available, a funnel plot was used to assess for possible publication bias.

A network meta-analysis, which can draw on both direct evidence and indirect evidence, was used to analyse the relative treatment effects. Analyses were planned for three different time points: end of treatment (primary time point); three months from the start of treatment; and three months after the end of treatment. However, for around two-thirds of trials, the three months from the start of treatment and the end of treatment time points were the same. The three months after the end of treatment time point was evaluated in few trials and no connected network incorporating acupuncture existed; only two very small other networks existed, each comprising three interventions.

The interventions were grouped and defined to give three sets of interventions which were analysed separately: 'Therapy-plus-adjunct interventions', 'Therapy-only interventions' and 'Grouped interventions'. The Therapy-plus-adjunct interventions set was defined to evaluate plausible differences in treatment between main interventions plus adjunctive therapy. The Therapy-only intervention set grouped all the adjunct treatments for an intervention together. The Grouped interventions set was defined according to the alternatives available to a general practitioner in the NHS such that certain main treatments were collectively defined as physiotherapy.

Network meta-analyses were conducted using WinBUGS software (version 1.4; MRC Biostatistics Unit 2007, Cambridge, UK) which uses Markov Chain Monte Carlo (MCMC) simulation to estimate model parameters and follows a Bayesian approach where prior probabilities are specified for parameters (these were specified to be vague throughout the analysis). The treatment difference was assumed to be normally distributed and a random effects network meta-analysis model was selected since clinical and methodological heterogeneity within the treatment definitions appeared likely. The model fit was evaluated using the residual deviance where this should be approximately equal to the number of data points if the fit is good. Inconsistency in the treatment effect estimates derived separately from direct and indirect evidence was assessed for many of the comparisons distributed across the networks. Uncertainty in all estimates is presented using the upper and lower limits of the 95% credible intervals (CrI) of these estimates.

In order to present more clinically meaningful network meta-analysis results, we present - for the end of treatment, Therapy-only intervention set - both SMDs, and the SMDs converted to the WOMAC pain VAS 0-100 scale (although it is acknowledged that back-transformation can be of limited value in heterogeneous populations).

To evaluate the impact of study quality on the results, two sets of analyses were performed: the main analysis including all studies regardless of quality ('any-quality'), a sensitivity analysis including only

studies of satisfactory, or better, quality ('higher-quality'). A further sensitivity analysis was performed to investigate the impact of excluding studies with atypical populations, interventions, or results.

1.4 Results

The searches retrieved 3,820 references. Of these, following screening of titles and abstracts, 553 full papers were considered potentially relevant to the review and a total of 138 trials were eligible; 134 original trials formed the basis of the review since four papers could not be translated. Thirty two studies (24%) were classified as being of either good or satisfactory study quality whilst the remainder were poor.

There were 22 main interventions and comparators in the included studies. Muscle-strengthening exercise, acupuncture, TENS, and balneotherapy were the most commonly studied interventions. Most studies were classified as having recruited a general population with osteoarthritis of the knee, although weight loss trials (as expected) recruited only overweight or obese participants. The mean BMIs of some studies recruiting a general population fell into the overweight or obese classification, although most studies did not report BMI. The majority of participants were women (range 26-100%, median 72%) and the median of the mean population ages was 64 years. Standard care and placebo were the most frequently studied comparators, with 'no intervention' being used rarely. There was considerable variation in the average treatment duration across the interventions.

1.4.1 Standard meta-analysis results (direct comparisons only)

There was some evidence, when all studies were considered, of a reduction in pain with acupuncture, muscle strengthening exercise, aerobic exercise, balneotherapy, TENS, static magnets, braces, NMES, and interferential therapy. However, the quality of most trials was poor and sample sizes small. When only higher-quality trials are considered, a benefit was demonstrated for acupuncture and muscle strengthening exercise only. Evidence from higher-quality trials indicated that insoles (without ankle support) did not have a beneficial effect. These analyses identified four trials as potential sources of significant heterogeneity in the network meta-analyses.

1.4.2 Network meta-analysis results (direct and indirect comparisons)

1.4.2.1 *Pain*

The main results are those for the end of treatment time point. There was no great difference in the results between the end of treatment analysis and the 3 months from start of treatment analysis. There was no network incorporating acupuncture for the 3 months from the end of treatment analysis.

Therapy-plus-adjunct intervention definition

Of the potential 110 Therapy-plus-adjunct interventions, 35 interventions formed part of a connected network with acupuncture and the evidence was informed by 79 trials. The results provided no indication of a treatment effect difference between the adjuncts for the majority of interventions. This suggests a lack of power in distinguishing between these treatment effects. Aerobic exercise with no medication was more effective than aerobic exercise with treatment as usual, but this lacks face validity. The analysis of any-quality studies for this set, found that PES, acupuncture, balneotherapy, sham acupuncture, laser/light treatment, static magnets and Tai Chi all showed a statistically significant treatment benefit over standard care, regardless of the adjunctive treatment. The sensitivity analysis of just higher-quality trials, showed a statistically significant treatment benefit over standard care for PES, acupuncture, balneotherapy, sham acupuncture, and muscle-strengthening regardless of the adjunctive treatment.

Therapy-only intervention definition

The results for the Therapy-only set of any-quality trials (87 trials, 22 types of intervention, 6753 patients), found interferential therapy, acupuncture, PES, TENS, aerobic exercise, and muscle-strengthening exercise to have a statistically significant treatment benefit over standard care. The results of the sensitivity analysis of higher-quality trials only (19 trials, 10 types of intervention, 2394 patients), reflected those of the main analysis except the credible intervals for aerobic exercise now crossed the line of no effect, the effect of balneotherapy became significant, and there were no

higher-quality trials of interferential therapy. When acupuncture was compared with the other interventions in both the main and sensitivity analysis it was found to be statistically significantly better at a 95% level of credibility than sham acupuncture and muscle-strengthening exercise. Acupuncture's median rank was 2 (95% credible intervals 1-4).

Grouped intervention definition
The results for the Grouped interventions set of any-quality trials (13 interventions informed by 86 trials) found that acupuncture, muscle-strengthening exercise, aerobic exercise, physiotherapy treatments and Tai Chi all showed a significant treatment benefit over standard care, and acupuncture showed a significant treatment benefit over muscle-strengthening exercise, insoles, and sham acupuncture, as well as placebo and no intervention.

The result for acupuncture compared with standard care was consistent across all the network meta-analyses (SMD of around -1.00) indicating reliable evidence of a beneficial effect of acupuncture on the pain of knee OA. We also compared our pair-wise analyses with those from the network meta-analysis, and found consistency for those interventions with a reasonable number of trials, in particular for acupuncture, aerobic exercise and muscle-strengthening exercise.

Publication bias could only be assessed for the muscle-strengthening exercise versus standard care comparison; no evidence was found for publication bias.

1.4.2.2 *WOMAC Index*

There were few studies and few interventions included in a connected network for the analyses with a WOMAC index outcome. Across the analyses of all available trials the results consistently indicate that acupuncture compared to standard care has a beneficial effect on the WOMAC index (mean SMD around -1.0) which is statistically significant (at the 95% level of credibility). For other treatments a statistically significant beneficial effect compared to standard care could be demonstrated only in the Therapy-plus-adjunct interventions analyses: weight loss with usual care, muscle strengthening exercise with usual care and Tai Chi with usual care (Tai Chi was only statistically significant at the end of treatment time point). Standard care + home exercise was also found to have a beneficial effect on the WOMAC index compared to standard care alone.

The sensitivity analysis including only higher-quality trials included only acupuncture, muscle strengthening exercise, sham acupuncture and Tai Chi. The results were consistent across the Therapy-only Intervention and Grouped Intervention networks and across the time points and mean estimates of effect were similar to those from the all trials analyses, but the results were no longer statistically significant (at the 95% level of credibility).

Again the result for acupuncture compared with standard care was consistent across all the network analyses (SMD of around -1.00) indicating consistent evidence of a beneficial effect of acupuncture on the level of overall disability associated with knee OA.

The mean estimate of effectiveness consistently favoured acupuncture over muscle-strengthening exercise, Tai Chi, heat treatment and sham acupuncture, but there was no consistent evidence that acupuncture was statistically significantly more effective than any of the other main interventions at a 95% level of credibility.

1.5 Discussion

Principal findings
Our analyses indicate that acupuncture is a worthwhile treatment option in the short term for treating knee pain due to osteoarthritis. The ability to distinguish between the effectiveness of the various physical treatments is subject to considerable uncertainty but analysis of higher-quality trials found acupuncture to be significantly better than standard care, sham acupuncture, muscle-strengthening exercise, weight loss and aerobic exercise. We also found reliable evidence that muscle-strengthening exercise also has pain-alleviating effects significantly better than standard care.

Acupuncture was found to have a significantly beneficial effect in improving the level of overall disability, as measured by the WOMAC index.

Strengths and limitations
Our review incorporates the first network meta-analysis comparing the relative efficacies of all relevant physical treatments for osteoarthritis of the knee. A network meta-analysis provides a basis of synthesising all the available evidence in a consistent framework, rather than making such decisions by subjective inferences from disparate data. We believe our study is the first network meta-analysis of physical treatments for knee OA. As such we encountered significant methodological challenges.

Our comprehensive and rigorous search strategy minimised the risk of missing eligible trials. However, although our review included 134 studies, limitations and differences in the reporting of data restricted the data available for our analyses, such that only 87 trials were included in the standard and network meta-analyses. Furthermore, the lack of long term data limits the interpretation of the results to only the short term effects of therapy.

The eligibility criteria of our review, encompassing a large number of interventions, with various adjunct therapies, placebos and populations, meant a certain amount of clinical heterogeneity was inevitable. Most studies recruited general populations, although it was acknowledged that within this categorisation there will have been variation in characteristics. Heterogeneity was explored using standard meta-analysis, and trials which were clearly a source of heterogeneity were removed in sensitivity analyses. A more general concern relates to the poor quality of a large majority of the studies. This should be borne in mind when interpreting our results. However, a major strength of our review is that trials of a diverse range of interventions have been evaluated equally, using the same quality assessment tool; this allows for a fair comparison in terms of assessing the strength of the evidence base for each intervention. Although we were unable to assess the impact of publication bias/small study effects on most of our interventions, it is possible that our results may well be subject to such biases.

A high level of inconsistency across the direct and indirect evidence for the effect on pain was found for the treatment comparisons involving PES and balneotherapy in both the Therapy-plus-adjunct and Therapy-only intervention set analyses, which suggests that there is bias or lack of exchangeability across the associated comparisons, and therefore the credibility interval estimates for both PES and balneotherapy may be underestimated. This implication may hold true for analyses including higher-quality trials only even though inconsistency for comparisons involving PES could not be evaluated due to a lack of triangles of evidence. Therefore, the results for the effect of both these interventions on pain may be unduely favourable.

Suggested research priorities
Larger, more robust RCTs with longer treatment periods, which also examine the effectiveness of re-treatment following treatment cessation (to evaluate durability and attenuation effects) are needed in order to comprehensively assess the value of many of these interventions. This is particularly true for TENS, where the studies conducted so far have been of unreliable quality, and PES and balneotherapy, which although our results highlight them as being promising treatments, were both represented by only one small higher-quality trial. The optimum timing and parameters of treatment for both acupuncture and muscle-strengthening exercise also need to be more clearly defined. Results from higher-quality studies suggest there would be little value for further research into the efficacy of insoles (without ankle support) or laser/light therapy since there appeared to be evidence these treatments were not effective.

We found that adding or subtracting trials in a network meta-analysis sometimes causes results to change more than expected, given the credible intervals around the estimates. This indicates unquantified uncertainty and unreliable results. Research could be conducted to develop a statistic to measure the stability or instability of the results given change in the evidence base.

Implications for service provision
Acupuncture can be considered as an evidence-based treatment option for relieving pain due to osteoarthritis of the knee. Although our review did not evaluate the cost-effectiveness of the interventions, it is worth noting that our results on effectiveness do not concur with the NICE guidance

for osteoarthritis management which states that TENS, insoles, braces, manual therapy, and heat or cold (thermotherapy) should be considered as adjuncts to core treatment. For these interventions our analyses found no evidence (of significant differences from standard care) to support this guidance, other than for TENS where the evidence was equivocal: all the TENS studies in our analyses were of poor quality, raising concerns about the reliability of the evidence. We have provided evidence on the effectiveness of acupuncture that NICE may want to consider when revising their guidance.

Conclusions

The first network meta-analysis of physical interventions for knee pain due to osteoarthritis, indicates that acupuncture is one of a number of physical treatments that produces a clinically-relevant effect in alleviating pain in the short-term. Moreover, acupuncture compared favourably with the other treatments. Acupuncture also significantly improved levels of overall disability. Although further research is needed to substantiate these conclusions, acupuncture should nevertheless be considered as an evidence-based treatment option for relieving pain due to osteoarthritis of the knee.

2. BACKGROUND

2.1 Acupuncture

Acupuncture – the insertion of fine needles into the skin – is used to treat a broad variety of illness and conditions. Although NICE currently recommends use of acupuncture only for lower back pain, it has also been commonly used to treat many other types of pain (both acute and chronic), including post-operative pain, headache and migraine, and neck, back, joint, and dental pain. Furthermore, acupuncture has been widely used to treat post-operative nausea and vomiting, allergies, infertility, menstrual disorders, digestive disorders, depression, anxiety, fatigue, and insomnia.[1] Proposed pain-relieving mechanisms and mediators for acupuncture include activation of the endogenous pain inhibitory system, release of endogenous opioids including β-endorphins, enkephalins, dynorphins, and non-opioid substances such as serotonin, noradrenaline, and GABA.[2]

2.2 Osteoarthritis of the knee

Osteoarthritis is a degenerative condition involving the progressing wearing-down of (joint) bone and cartilage, normally resulting in pain, stiffness, and functional disability. These symptoms usually worsen, according to how much the affected joint is used. In adults aged 45 years or more, the knee represents the most common site of peripheral joint pain, and the prevalence of painful, disabling knee OA in people over 55 years is 10%.[3] Risk factors for knee OA include age, gender, obesity, bone density, genetic factors, and injury.

Diagnosis is usually made using clinical features of knee OA, by radiological assessment of the knee, or by a combination of the two. Radiographic features - the severity of which are commonly summarised using the Kellgren & Lawrence score - have been significantly associated with knee pain.[4]

The Western Ontario and MacMaster Universities Osteoarthritis (WOMAC) index is a self-administered disability status measure for knee (or hip) osteoarthritis; it was developed in 1982, and has been translated into over 80 languages. Its individual components assess pain, stiffness, and function, with the summed scores producing an overall measure of disability (WOMAC index). By producing a standardised and comprehensive assessment of disability, and its components, WOMAC lends itself to increased transparency and comparability within clinical research.

2.3 Management of knee OA and current service provision

The treatment of knee OA should be tailored according to knee risk factors (obesity, adverse mechanical factors, physical activity), general risk factors (age, comorbidity, polypharmacy), level of pain intensity and disability, sign of inflammation, and location and degree of structural damage.[5] The main objective of a general practitioner treating a patient with knee OA is normally alleviation of pain; failure to control pain may result in reduced mobility and daily activities, leading to a reduction in quality of life.[5] The more sedentary lifestyle which might follow may, in turn, exacerbate the symptoms of knee OA through lack of exercise and joint movement, and weight gain.

In clinical practice, treatment often begins with analgesia (paracetamol and/or topical NSAIDs) and, where these are ineffective, a NSAID or COX-2 inhibitor is recommended. General practitioner advice about exercise and weight loss, which the NICE Guideline[6] recommends as part of core therapy, is often given in addition to (rather than instead of) analgesic drugs. The regular use of pharmacological agents for pain, such as NSAIDs, may be associated with side effects like gastrointestinal bleeding, without necessarily resulting in worthwhile pain reduction.[7]

In light of this possibility of adverse effects, the long-term use of oral NSAIDs is not desirable. A UK review of qualitative studies of medicine-taking[8] revealed considerable reluctance to take drugs, and a preference to take as little as possible; knee OA patients want non-pharmacological treatments for pain relief.[9] The use of non-pharmacological (or physical) treatments, such as acupuncture, is therefore likely to be attractive for patients seeking alternatives, particularly for a condition such as OA of the knee, for which there is currently no cure.

In patients where insufficient pain relief has been provided by the aforementioned core interventions (as recommended by NICE), coupled with paracetamol and/or topical NSAIDs, GPs may consider a range of physical treatments as the next step in the treatment pathway. The NICE guideline lists manual therapy, TENS, braces, insoles, and heat and cooling treatments as being among such alternatives; it states that these third tier treatments may have less well-proven efficacy or may provide less symptom relief.[6] Acupuncture, and other types of physiotherapy were not recommended as being part of this third tier, but they occupy a similar place in the treatment pathway. Aids such as walking sticks, or nutritional supplements, which patients may purchase for themselves, generally comprise part of the background therapy, rather than options available to the general practitioner.

Other interventions used for OA of the knee, but which would not be considered as alternatives to acupuncture include surgery, which would be considered at a much later stage in the treatment pathway, and intra-articular injections, which are classed as being pharmacological and are normally only offered to elderly patients, or patients who are on the brink of needing surgery (a population usually Therapy-plus-adjunct from those receiving physical therapies).[6] Similarly, structured psychosocial/educational interventions are generally considered for a different group of patients i.e. when pain-reducing therapies have failed, and the emphasis is on a need for pain-coping skills, rather than pain reduction.[10]

Interventions aimed primarily at reducing anterior knee pain - such as patellar taping - are not potential alternatives to acupuncture, since anterior knee pain is normally seen in patients younger than the general knee OA population, and any associated knee OA may be likely to have a different aetiology.[11]

Many reviews have been undertaken of the varying types of physical therapies for OA of the knee, but none have attempted to address the question of how effective such treatments are relative to each other, and few randomised trials have directly compared physical therapies. The focus of interest within our study was on acupuncture, since one of the reasons for the commissioning of this review as part of a programme of projects on acupuncture and chronic pain, funded by an NIHR Programme Grant for Applied Research, was the uncertainty within the NICE decision-making process with regard to the level of evidence on acupuncture for osteoarthritis relative to other physical treatments.[12] The purpose of this systematic review therefore, is to synthesise the indirect (and any direct) evidence - using mixed treatment comparison (network meta-analysis) methods - in order to compare the effectiveness of different physical therapies for knee OA pain.

3. AIMS AND OBJECTIVES

The objective of this review is to determine how clinically effective acupuncture is in the treatment of OA of the knee in the context of NHS prescribing. Evaluation of a single therapy for a single condition provides only a limited basis for NHS decision making; more relevant is a full evaluation comparing the clinical effectiveness of acupuncture with all relevant comparator treatments using both direct and indirect comparisons. Therefore the decision problem addressed in this report is 'how does the clinical effectiveness of acupuncture for the pain of OA of the knee compare with alternative physical therapies in patients who require additional or alternative therapy to pharmacological analgesia?'

4. METHODS FOR ASSESSMENT OF CLINICAL EFFECTIVENESS

4.1 Methods for identifying clinical effectiveness evidence

The evaluation of the clinical effectiveness of acupuncture in the treatment of OA of the knee comprised a systematic review of physical therapies incorporating a network meta-analysis. The review processes and outline methods of analysis were specified in advance and documented in a protocol. Details of the analysis were finalised once the available data had been identified. The systematic review was conducted following the general principles recommended in the Centre for Reviews and Dissemination's (CRD) guidance[13] and the PRISMA statement.[14]

4.1.1 **Inclusion and exclusion criteria**

4.1.1.1 *Interventions and comparators*

The interventions considered in the review were acupuncture and all interventions that can be considered direct comparators of acupuncture: balneotherapy; braces; exercise - aerobic (weight bearing); exercise – muscle strengthening (non-weight bearing); heat treatment; ice/cooling treatment; insoles; interferential therapy; laser/light therapy; manual therapy; neuromuscular electrical stimulation (NMES); pulsed electrical stimulation (PES); pulsed electromagnetic fields (PEMF); static magnets; Tai Chi; transcutaneous electrical nerve stimulation (TENS); and weight loss. Therapies that comprised a combination of these therapies were not included in the review.

Studies comparing different regimens/durations/modalities of the same type of intervention only were excluded. Comparators included: standard care; placebo interventions; no intervention; and sham acupuncture. It was anticipated that standard care would vary across trials, but could include analgesics, such as NSAIDs, education or advice on fitness, exercise or diet. The components of standard care had to be clearly Therapy-plus-adjunct from more active education or exercise interventions (sessions) to which patients may be referred. Exercise interventions which were predominantly home-based, and unsupervised, were excluded, as they were considered to be too similar to core treatment/standard care.

4.1.1.2 *Population*

Studies of adults with OA of the knee were included. Studies in which patients were diagnosed using either radiological or clinical assessment were eligible. Studies with mixed populations (e.g. including both patients with OA of knee and those with OA of the hip) which presented results by site of OA were eligible for inclusion. Trials of acute knee pain or trials where the mean age of the population is below 55 years were excluded.

4.1.1.3 *Outcomes*

The primary review outcome was pain. Studies that did not report a pain outcome were excluded from the review. It was anticipated that across trials pain would be measured using a variety of measures, e.g. visual analogue scale (VAS); Likert scale; Western Ontario and MacMaster Universities (WOMAC) pain subscale; Arthritis Impact Measurement Scale (AIMS); all scales were accepted. The secondary outcome was the WOMAC Osteoarthritis Index (overall disability).

WOMAC is a widely used self-administered health status measure that assesses the dimensions of pain, stiffness, and function in patients with OA of the hip or knee; it is available in 5-point Likert, 11-point numerical rating, and 100 mm visual analogue scale formats. Under each dimension there are number of questions designed to assess the clinical severity of the disease (5 questions for pain, 2 questions for stiffness and 17 questions for physical function). The patient's response to each question produces a score which is then added up to derive an aggregated score for each dimension. It produces three subscale scores (pain, stiffness, and physical function) and a total score (WOMAC index), which reflects disability overall.

The WOMAC pain score range is variously reported and includes: VAS 0-10 scale (commonly reported across a 0-50 range); VAS 0-100 scale (commonly reported as a 0-500 range); or a Likert scale (commonly reported as a 0-20 range). The overall WOMAC score (index) is determined by summing the scores across the three dimensions and the score range includes: VAS 0-10 scale (commonly reported as a 0-240 range); VAS 0-100 scale (commonly reported as a 0-2,400 range); or a Likert scale (commonly reported as a 0-96 range). A number of various transformations and modifications are reported in the literature.

4.1.1.4 *Study designs*

Only randomised controlled trials (RCTs) were included in the review of clinical effectiveness.

4.1.2 **Search strategy and identification of relevant studies**

An initial search to identify guidelines, syntheses or reviews of non-surgical or non-pharmacological interventions for osteoarthritis of the knee was undertaken to inform the project. A range of resources were searched:

- Clinical Evidence
- NHS Clinical Knowledge Summaries (CKS)
- NHS Evidence- National Library of Guidelines
- National Institute for Health and Clinical Excellence
- National Guideline Clearinghouse
- New Zealand Guidelines Group – Guidelines Library
- Australian National Health and Medical Research Council: Clinical Practice Guidelines
- Canadian Medical Association – Infobase: Clinical Practice Guidelines
- Public Health Agency of Canada
- Cochrane Database of Systematic Reviews
- Database of Abstracts of Reviews of Effects
- Health Technology Assessment database
- Physiotherapy Evidence Database (PEDRO)
- NHS Evidence - Musculoskeletal 2009 Evidence Update on osteoarthritis
- NIHR Health Technology Assessment Programme

To search for primary studies a base search strategy was developed (using Ovid MEDLINE) in consultation with the review team and clinical experts. The strategy was designed to find trials of acupuncture or its relevant comparators, as defined in the protocol (manual therapy, exercise, weight reduction, balneotherapy, thermotherapy, braces and orthoses, TENS, electrical muscle stimulation, electromagnetic fields, low-level laser/light therapy) for osteoarthritis of the knee or chronic knee pain. A combination of relevant free text terms, synonyms and subject headings were included in the strategy. No language or date restrictions were applied. The base search strategy is presented in Appendix 10.1.

The following resources were searched in December 2009/January 2010:

- MEDLINE
- MEDLINE In-Process & Other Non-Indexed Citations
- EMBASE
- AMED (Allied and Complementary Medicine)
- Cochrane Database of Systematic Reviews
- Database of Abstracts of Reviews of Effects
- Health Technology Assessment database
- Cochrane Central Register of Controlled Trials (CENTRAL)
- CINAHL
- Manual, Alternative and Natural Therapy (MANTIS)
- PASCAL
- Inside Conferences
- Conference Proceedings Citation Index-Science (CPCI-S)
- Physiotherapy Evidence Database (PEDRO)
- CAMbase
- Literatura Latinoamericana y del Caribe en Ciencias de la Salud (LILACS)
- ClinicalTrials.gov

The base search strategy developed in MEDLINE was translated to run on the databases listed above. Where available, a search filter was used to limit the retrieval of studies to randomised controlled trials. Adaptations to the search strategy were necessary for certain databases: MANTIS, PASCAL, Inside Conferences, PEDRO, CAMbase, LILACS and ClinicalTrials.gov. It had been planned to search Acubriefs database (http://www.acubriefs.com/) but it was not available when the searches were carried out.

Supplementary internet searches of websites relating to osteoarthritis were undertaken to locate any studies not found from the database searches. The bibliographies of all relevant reviews and guidelines were checked for further potentially relevant studies.

Update searches were carried out in MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, EMBASE, AMED, CDSR, DARE, HTA, CENTRAL, CINAHL, MANTIS, CPCI-S, PEDRO, CAMbase, LILACS and ClinicalTrials.gov in June 2010.

The base search strategy can be found in Appendix 10.1 and full details of databases searched, search strategies and results can be found at:
www.york.ac.uk/inst/crd/Documents/OAKSearchStrategiesWebLink.docx

Two reviewers independently screened all titles and abstracts identified by the searches. Full paper manuscripts of all studies thought to be potentially relevant by either reviewer were obtained. The relevance of each article was assessed by two independent reviewers according to the criteria stated below. Any discrepancies were resolved by consensus or, when consensus could not be reached, a third reviewer was consulted. Non-English language papers were screened by one reviewer with a native speaker.

4.2 Data extraction strategy

Data extraction was conducted by one reviewer and checked by a second reviewer; discrepancies were resolved by discussion, with the involvement of a third reviewer when necessary. Non-English language studies were extracted by one reviewer with a native speaker. Multiple publications of the same study were extracted as one study, using all the information available.

Data relating to both study content and quality were extracted by using a standardised data extraction form entered onto an online database in EPPI-reviewer. Extraction included data on population characteristics (population type, method of diagnosis, age, sex, weight, BMI, Kellgren & Lawrence score), treatment details (interventions studied, number of comparator arms and comparators studied) and intervention parameters (e.g. duration of individual session, total number of sessions, duration of treatment period), standard care group data (if present, components of standard care), adverse effects, quality of life outcomes and an assessment of study quality.

Outcome data (pain and WOMAC Index scores) were extracted onto an Excel spreadsheet. Where different types of pain were reported such as pain on standing, pain of walking etc., the following hierarchy of pain outcomes was used, determined on the basis of perceived importance to a patient, starting with the most important: walking pain, activity or movement, ascending or descending stairs, weight bearing activity, starting pain, and rest or night pain.

Data were extracted for baseline for each arm and end of treatment follow-up. If end of treatment follow-up was not reported then data were extracted for difference from baseline for each arm. If neither end of treatment follow-up nor difference from baseline were reported, then data extraction focussed on treatment effect (comparison) results. If nothing else was reported then p values were extracted. Outcome data were extracted for different time points: baseline; end of treatment; and any follow-up time point.

4.2.1 Quality assessment strategy

The quality of the individual studies was assessed by one reviewer and checked for agreement by a second reviewer. Any disagreements were resolved through consensus and checked by a third reviewer where necessary. The quality of RCTs was assessed using the checklist advised in CRD's guidance 2008, adapted as necessary to incorporate topic-specific quality issues.

Study quality was assessed using 14 questions. These covered randomisation, allocation concealment, type of placebo, blinding, comparable baseline characteristics, use of power calculation, reporting of eligibility criteria, reporting of losses to follow up, reporting of intention-to-treat data and losses to follow-up (see Appendix 10.2).

Based upon the number of criteria satisfied then studies were graded as excellent, good, satisfactory and poor.

Satisfactory study quality was defined as meeting the following criteria:

- The number of patients randomised to treatment was stated
- Group baseline characteristics comparable
- Eligibility criteria were adequately reported
- Losses to follow up were clearly reported
- Intention-to-treat data were reported (analysed)
- An appropriate type of placebo was used, if relevant

'Poor quality' studies failed to satisfy one or more of the criteria required for satisfactory study quality. Study quality was used in sensitivity analyses for both the meta-analyses and network meta-analysis, with studies that met the criteria for satisfactory studies (or better) being considered separately: these are referred to as 'higher-quality'.

4.3 Data analysis

4.3.1 **Outcome data**

WOMAC pain was the preferred measure of pain for the analyses. When a trial did not measure WOMAC pain then another pain scale was included in the analysis. Further prioritisation of pain scales were made on a clinical basis or on a prevalence basis: the AIMS pain scale was selected over the PPI pain scale because there were a greater number of trials with the AIMS pain scale. There are several different types of pain measured using VAS. The following hierarchy of pain outcomes was determined on the basis of perceived importance to a patient, starting with the most important: walking pain, activity or movement, ascending or descending stairs, weight bearing activity, starting pain, and rest or night pain.

For analysis purposes standardised mean differences (SMDs) were used. A number of reasons underpin the use of SMDs within this review. A variety of scales and pain activities were used in the trials in the review; focussing the analysis upon a single measure (for example, the WOMAC Likert scale) would have seriously limited the amount of evidence that could be combined in an analysis. In addition, transforming the Likert 5 and VAS 0-100 to the same scale may be an extreme assumption given that their standard deviations may not be transformable on the same scale. In a similar vein the standard deviations may vary for different pain measures and in some instances the scales were insufficiently defined.

Where possible, Hedges-g SMDs were estimated[15] and pooled and different doses/regimens of the same type of treatment within a study were pooled (see Appendix 10.3 for the formula). The standard deviation used to calculate Hedges-g was the pooled standard deviation across all of the arms in the trial. Where trials included more than one arm with treatments coded the same, the data across such arms were also pooled.

Reduction in pain recorded as an SMD can be interpreted in the WOMAC VAS 0-100 scale using Table 1 (although it is acknowledged that back-transformation can be of limited value in heterogeneous populations).[16 17] Of the 11 trials reporting a WOMAC VAS 0-100 scale, 4 of these were cumulative. Of the 7 that were not cumulative, one did not report baseline pain scores. Of the six trials that did report baseline pain scores, the mean baseline pain score was 45.19 (SD=8.2).

Table 1: SMD equivalent reduction in pain measured on the WOMAC VAS 0-100 scale

	Reduction in pain score			
SMD	-0.5	-1	-1.5	-2
WOMAC VAS 0-100 scale	-8.25	-16.5	-24.75	-33

Reduction in pain recorded as an SMD can be interpreted in the WOMAC Likert 5 cumulative scale using Table 2. All of the 20 trials that reported this scale also reported the baseline pain score. The mean baseline pain score was 8.96 (SD=1.82).

Table 2: SMD equivalent reduction in pain measured on the WOMAC Likert 5 cumulative scale

	Reduction in pain score			
SMD	-0.5	-1	-1.5	-2
WOMAC Likert 5 cumulative scale	-1.9	-3.8	-5.7	-7.6

Final values and changes from baseline should not be combined in an analysis when SMDs are used as the differences in standard deviations will not reflect differences in the scales.[18] Final values were used in the analysis as there were significantly more trials with final value data than change from baseline data. This maximised the evidence available for analysis: there were only slightly fewer trials with final values than all trials with change from baseline and final and baseline values together. Furthermore, using final values does not require an assumption about the within-patient correlation between the baseline and final values.

Where there were no trial arm data but there were treatment effect data of final values then the treatment effect data were included in the analysis, adjusted or unadjusted. Median data were excluded from the analysis in order to try to maximise consistency in the data. Eight trials were excluded on the basis that they reported medians or that it was unclear whether a mean or a median was used.

Where the number of patients included in the analysis was not reported but the number of patients randomised was, then the number of patients included in the analysis was estimated by multiplying the number of patients randomised by the average proportion of patients included in an analysis across the trials. There were 7 trials out of the 91 that required the number of patients in the analysis to be derived.

Where the standard deviation was not reported then, where possible, the standard error or 95% confidence intervals were used to derive the standard deviation (see Appendix 10.3 for the methods). Where no standard deviation or data that could be used to derive standard deviations were reported in a trial, then a standard deviation was estimated using data from the other included trials that could potentially be used in the defined analyses. The trials with the same or similar scale as that used in the trial with a missing standard deviation were identified and the standard deviations across these trials were pooled. There were 12 trials that needed the standard deviation to be imputed from the 91 trials that were included.

Reporting of adverse events data was sparse and did not warrant quantitative data synthesis; where relevant a narrative synthesis of these data was reported.

4.3.2 **Meta analysis (direct comparisons)**

The meta-analysis part of the review was not intended as a comprehensive stand-alone synthesis, but as a means of informing and complementing the network meta-analysis. In particular, it investigated the within-intervention clinical and statistical heterogeneity. Meta-analyses were conducted using outcomes recorded at the end of treatment only. Comparisons of SMDs generated using pair wise meta-analysis and network meta-analysis are presented in table 41.

All meta-analyses were conducted in RevMan 5.0 (Cochrane Collaboration). A random effects model was used, unless there were four or fewer studies included in the analysis, in which case a fixed-effect model was used, as the estimate of the heterogeneity parameter is likely to be unreliable with small numbers of trials.[19]

Clinical heterogeneity was assessed by investigating the clinical differences between studies regarding participants, interventions and outcomes. The potential sources of the clinical heterogeneity, such as baseline patient population, different durations of intervention, differences in additional treatments, and study quality were identified. Generally trials were pooled, except where

clinical heterogeneity was particularly great. The degree of between-study statistical heterogeneity was investigated using the I^2 statistic.

4.3.2.1 *Network meta-analysis (direct and indirect comparisons)*

An NMA is an extension of meta-analysis, but where a meta-analysis includes only *direct* evidence an NMA can draw on both *direct* and *indirect* evidence. The results from studies that compare interventions A and C are considered to be *direct* evidence for the treatment effect d_{AC}. If a study X compares treatments A and B and a study Y compares treatments B and C, and a treatment effect d_{AC} is calculated from these two studies, then this result is referred to as *indirect* evidence.

A standard meta-analysis combines the results from two or more studies that have comparable populations, interventions, comparators and outcomes. Study quality and other study characteristics are also assumed to be similar. Similarly, to make indirect comparisons, it is assumed that the study characteristics are comparable. This is known as *exchangeability* which can be investigated through the consistency of the direct and indirect evidence. [20-22] It assumes that, had treatment C been included in the study comparing A and B, then the treatment effect d_{AC} would be the same as that found from the study of A and C.[23] Assuming consistency, the treatment effect d_{AC} is the sum of the treatment effects d_{AB} and d_{BC}:

$$d_{AC} = d_{AB} + d_{BC}$$

An NMA can combine both the direct evidence and the indirect evidence for d_{AC}.[23]. As in a meta-analysis, it is the summary treatment effect from each study that is utilised in the NMA, hence the benefit of randomisation in each study is retained.

4.3.2.2 *Interventions*

The 22 main interventions are listed in Table 3. To allow for the potential for interaction effects with adjunct treatments, five adjunct interventions were defined. In total, there were 110 possible treatment combinations.

For analysis, the interventions were grouped and defined in two ways to give three Therapy-plus-adjunct sets of interventions which were analysed separately: 'Therapy-plus-adjunct interventions', 'Therapy-only interventions' and 'Grouped interventions'. The Therapy-plus-adjunct interventions set was defined to evaluate plausible differences in treatment effect between the competing interventions as defined at the level of main intervention plus adjunctive therapy, e.g. acupuncture plus standard care versus acupuncture plus home exercise.

Table 3: The 22 main interventions* and 5 adjunct variations* for each main intervention

Main				Adjunct	
Acupuncture	(Acu)	1		No medication (NoMed)	1
Sham acupuncture	(ShAcu)	2		Treatment as usual/Unclear (UT)	2
Balneotherapy	(Bal)	3		Treatment as usual/unclear + Home exercise/education (UT+ EX)	3
Braces	(Bra)	4		Treatment as usual/unclear + specified analgesics (UT+AN)	4
Exercise - Aerobic (weight bearing)	(ExAe)	5		No medication + home exercise/ education (NoMed+EX)	5
Exercise - Muscle strengthening (non-weight bearing)	(ExMu)	6			
Heat treatment	(Hea)	7			
Ice/cooling treatment	(Ice)	8			
Insoles	(Ins)	9			
Interferential therapy	(Int)	10			
Laser/light therapy	(Las)	11			
Manual therapy	(Man)	12			
NMES	(NMES)	13			
Pulsed electrical stimulation	(PES)	14			
Pulsed electromagnetic fields	(PEMF)	15			
Static magnets	(Mag)	16			
Tai Chi	(Tai)	17			
TENS	(TENS)	18			
Weight loss	(Wei)	19			
Standard care	(SC)	20			
Placebo	(P)	21			
No intervention	(NoTr)	22			

- Abbreviations in parenthesis are those used in Appendix figures

The Therapy-only intervention set grouped any adjunct treatments for an intervention together (i.e. all adjuncts were treated as being the same). So there were a possible 22 interventions in the analysis.

The Grouped interventions set of comparators was defined according to the alternatives available to a general practitioner in the NHS after the initial prescription of medication and advice. This was motivated by the requirements of the related economic evaluation part of the project. Many of the main interventions in Table 3 were considered to be techniques commonly utilised by a physiotherapist, and therefore the GP's choice was whether or not to refer to a physiotherapist. For this reason, these main treatments were collectively defined as physiotherapy. See Table 24 for the list of treatment techniques. It also allowed us to evaluate whether or not the different levels of defining the main interventions significantly affects their relative treatment effects.

4.3.2.3 *Time points*

The NMA analyses were planned on outcomes recorded at 3 different time points:

- The outcome recorded at the end of treatment (end of treatment)
- The outcome recorded at a time point closest to 3 months from the start of treatment, excluding outcomes recorded at less than 4 weeks from the start of treatment (3 months from start of treatment)
- The outcome recorded at a time point closest to 3 months from the end of treatment between 8 and 16 weeks from the end of treatment (3 months from the end of treatment)

The end of treatment outcome measures an immediate treatment effect due to the intervention. The 3 months from the start of treatment outcome measures to some degree the durability of a treatment effect during a treatment and also after the end of treatment. The 3 months from the end of treatment outcome measures the durability of a treatment effect from the end of treatment.

4.3.2.4 *Study quality*

In order to evaluate the impact of study quality on the NMA results, two analyses were done for each set of comparators for each time point. In the initial analysis, any studies were included regardless of quality ('any-quality'). In the sensitivity analysis, only 'higher-quality' studies (good or satisfactory quality studies) were included.

4.3.2.5 *Networks*

For both sets of comparators, a connected network including acupuncture was obtained. An NMA was conducted on all of the trials informing the connected networks. Given the three time points of analysis, the two sets of comparators and two study quality criteria, and two different outcomes (pain and overall WOMAC score), 24 networks were theoretically possible.

4.3.2.6 *The model*

The WinBUGS software was used for this analysis. This is a Bayesian analysis software where prior probabilities are specified for certain parameters and likelihood distributions are defined for the data. The outcome for this analysis is the treatment effect difference, so a normal likelihood distribution was specified for the treatment effect data.

A random effects model was selected as there was likely to be some clinical and methodological heterogeneity within the treatment definitions.

Multiple-arm trials

The data of trial arms with treatments coded the same were pooled.

The treatment effects d_{AB} and d_{CB} in a 3-arm trial are correlated. This was accounted for in the model using multiple-arm trial code for trial arm data produced by Bristol University[24], which was adapted for difference data. The code is presented in Appendix 10.5.

As SMDs are being evaluated, treatment differences are entered into the analysis rather than trial arm data. If there is a 3-arm trial then two comparisons will be entered in the dataset with a shared comparator; the same data for the comparator will be entered twice. A method of dealing with double-counting of comparator data in such situations is to split the comparator sample size and assume a 3-arm trial becomes two independent trials. That is a secondary approach and it ignores the correlation between the treatment effects. Accounting for the correlation in an analysis reduces the variance of the treatment effect estimate. Halving the sample size of the comparator increases the variance of the treatment effect estimate. Accounting for double counting by halving the comparator sample size and retaining the correlation in the analysis should produce treatment effects and variances somewhere in between ignoring double counting and ignoring correlation.

The results from each of these modelling approaches on one of the networks were compared. The treatment effects and variances did not differ greatly and the treatment effect and variance estimates lay mostly in between the other estimates. This approach was selected for the analyses. The results of the tests are presented in the Appendix NMA Table 4.

Consistency analysis

Differences in trial populations and protocols between the trials informing indirect evidence of a comparison and trials informing direct evidence of the comparison can result in different estimates for that comparison. Where the posterior distributions of direct and indirect estimates do not overlap the estimates are said to be inconsistent. To evaluate the consistency between the direct and indirect estimates of any one comparison, the direct evidence was separated from the indirect evidence within the analysis, which is referred to as 'node splitting' in,[20] although it is evidence splitting for a particular

treatment comparison. Dias et al[20] was used to inform the code presented in Appendix 10.5 for the consistency analyses.

Explanantion of network diagrams

Each box represents a treatment. Each solid arrow indicates that there is a data point for that comparison entered into the analysis. A 3-armed trial with arms A, B and C, provides 2 data points in a data set reflecting 2 comparisons A vs B and C vs B. These comparisons are represented by solid arrows and these determine the existence of evidence triangles on which consistency of direct and indirect evidence can be tested. The dotted arrows show comparisons with evidence from 3-armed trials (e.g. A vs C) for which there is no data point in the data set. There were ten 3-arm trials and one 4-armed trial in the any-quality, end of treatment, Therapy-only analysis. The thickness of an arrow represents the number of trials in the analysis that have that comparison that has been directly entered in the analysis. The numbers indicate the level of inconsistency between the direct and indirect evidence for that comparison. A value of 1 represents complete inconsistency, and a value of zero indicates perfect consistency. The consistency results are reported on the network diagrams in the NMA Appendix Figures 1-6, 12, 13.

Model fit and convergence

The model fit was evaluated using the residual deviance. The model fit is the degree to which the model explains the data; the degree to which the variance of the predicted model values are the same as the variances of the individual trial estimates. If a model is a good fit, then the residual deviance should be close to the number of data points in the data set. There is no upper limit to the residual deviance, but a large percentage difference from the number of data points is a poor fit.

The WinBUGS software uses Markov Chain Monte Carlo simulation and Gibbs sampling to estimate model parameters from prior probability and likelihood distributions. Starting values are specified for parameters modelled as a distribution. These should converge to a stable distribution after many sampling iterations. Convergence of the model estimates was assessed by observing the history of the traces of the starting values for selected priors, the Brooks Gelman-Rubin statistic and posterior distributions.[25]

The first 10,000 iterations were discarded and then a further 50,000 iterations were performed. It was tested whether discarding the first 30,000 iterations made any difference to the results, and there was none.

The prior for the between study standard deviation was set to be a uniform distribution with range 0-2. This clearly covered the range of treatment effects within a particular comparison. Other ranges were also tested: 0.8, 5, 10, and greater if unstable estimates obtained across this range.

Model outcomes

The treatment effects of each treatment compared with standard care are presented, as are the results for the Therapy-only interventions network compared with acupuncture. The full results of all the pair wise comparisons for each analysis are published in an online appendix. For brevity only the results where acupuncture is significantly more effective than the comparator at a 95% level of credibility are referred to in the text. Uncertainty was presented using the upper and lower limits of the 95% credible intervals of these estimates. These credible limits describe the boundaries within which it is believed there is a 95% chance that the true value lies. In addition, since only the results of the comparisons compared to standard care were presented and these were listed in order of mean effectiveness for ease of reading, the median rank and 95% credible interval for the rank were presented for each intervention.

For the random effects models, the between study standard deviation τ of the random effects distribution were reported. Consistency analysis was performed for a selection of comparisons appropriately distributed across a network. This was done for each of the End of treatment analyses except for the Grouped Intervention set including only higher-quality studies, and it was done for the 3 months from the start of treatment and Grouped Intervention set analyses, with Any-quality studies

and only higher-quality studies. The residual deviance was also reported along with the number of data points and percentage deviance difference to indicate the model fit.

Publication bias

Where there were a sufficient number of trials for any one main intervention, publication bias was explored by means of funnel plots.

5. RESULTS OF ASSESSMENT OF CLINICAL EFFECTIVENESS

5.1 Quantity and quality of research available

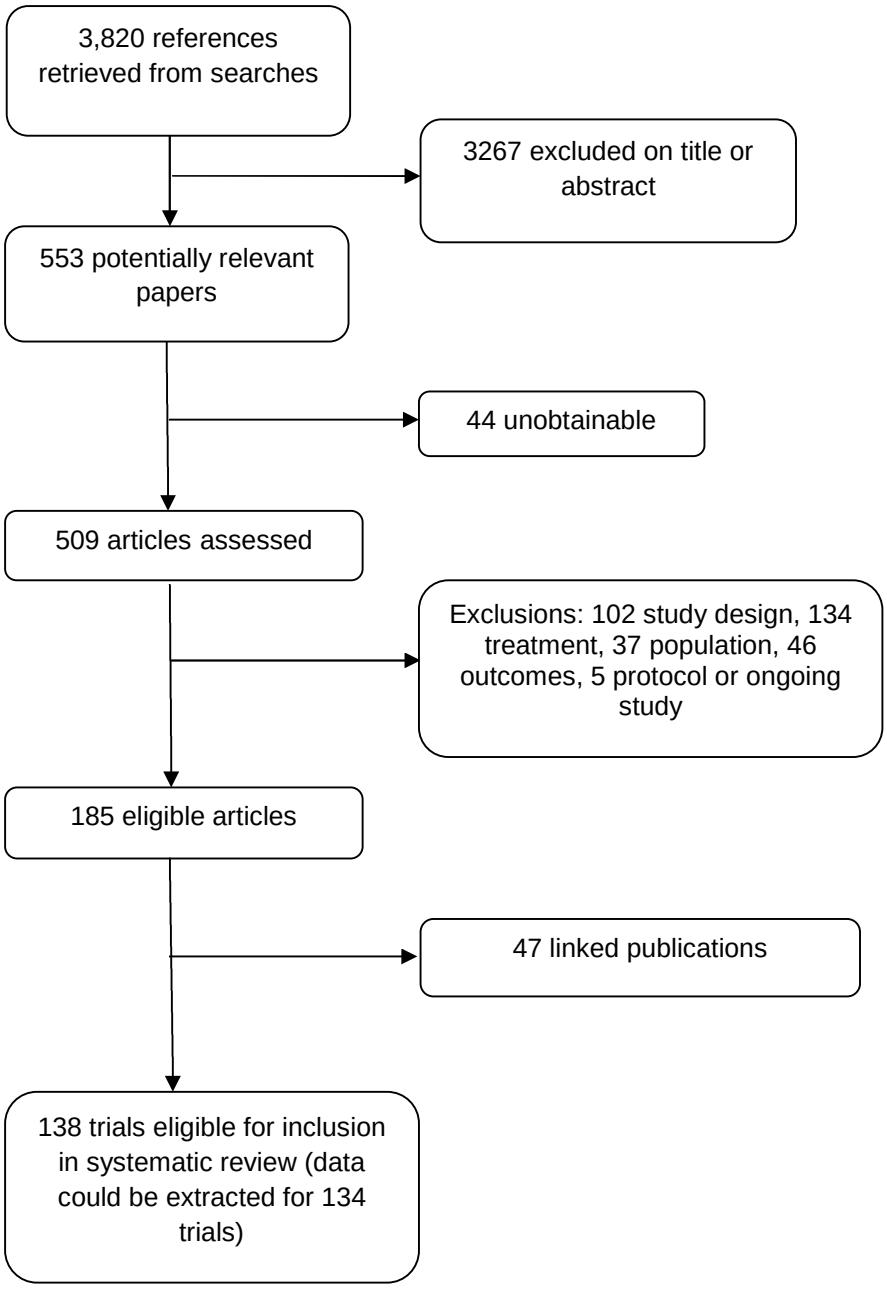


Figure 1: Study flow chart.

The searches retrieved 3,820 references. Of these, following screening of titles and abstracts, 553 full papers were considered potentially relevant to the review of clinical effectiveness of treatments for OA of the knee; 44 of these were unobtainable, leaving 509 papers that were assessed against the inclusion criteria.

A total of 102 studies were excluded on the basis of study design, 99 due to ineligible treatment, 37 due to an ineligible study population, and 46 due to outcomes. A further 35 studies were excluded because the intervention studied was a combination of the interventions of interest. Five studies were either still at the protocol stage or were ongoing. Excluded studies are listed in Appendix 10.8.

A total of 185 papers met the inclusion criteria and after linking all papers, a total of 134 original trials formed the basis of the review; one study[26] was treated as two separate trials, as it comprised two Therapy-plus-adjunct populations: both a general OA of the knee population and patients with varus malalignment (randomisation was stratified). Four articles could not be translated (one Danish, one Czech, and two Turkish), but six trials were included where the primary source of data was published in a language other than English: four German,[27-30] one Chinese[31] and one Spanish.[32] The number of trials for each intervention are discussed under the following sections. A flow chart of studies eligible for the systematic review is presented in figure 1.

Summary results of the quality assessment by intervention are also presented in the following sections; 32 (24%) of the studies in the review were classified as being of either good or satisfactory study quality whilst the remainder were poor. Detailed results can be found at: http://www.york.ac.uk/inst/crd/Documents/StudyQualityAssessmentResults.docx

Due to the limited number of studies relating each intervention a funnel plot was only worthwhile for the comparison of muscle-strengthening exercise versus standard care (see section 6.2.2.2).

5.2 Standard Meta-analysis

5.2.1 **Acupuncture**

5.2.1.1 *Study characteristics*

Twenty-two trials studied acupuncture (see Table 4), with a total of 2167 participant pain scores analysed at the end of treatment (range 14 to 342). All were full papers published between 1988 and 2010 (20 in English, one in German, and one in Chinese); ten were published between 2007 and 2010. The majority of studies were conducted in either China (four RCTs), the UK (four), Germany (three), or the United States (three).

Twenty studies recruited a general population, one[33] studied only participants with both knees affected by osteoarthritis, and one[34] studied only participants awaiting knee surgery. The mean ages of participants ranged from 58 to 85 years, and the proportion of females ranged from 50 to 96%. Mean BMI was reported in only seven studies (range 29 to 33kg/m^2), and mean weight in only four studies (range 60 to 90kg). The methods used for diagnosis were clinical and radiological in seventeen studies, and clinical alone in four studies (the details were unclear in one study). Where reported, most participants had Kellgren & Lawrence scores of at least two (although ten studies did not report details on classification of severity).

Twelve trials compared acupuncture with sham acupuncture, and eleven with standard care. Three trials [35-37] also studied TENS; muscle-strengthening exercise[34] and cooling treatments[37] were the remaining interventions studied. Sessions generally lasted for 20 or 30 minutes. The majority of studies gave between eight and 12 sessions (range: one to 23 sessions), with four to eight weeks being the commonest treatment periods (range: two to 26 weeks). The number of points needled ranged from 2 to 24, with the majority using between 5 and 12 points. Insertion depths were generally around 10-15mm, although eight studies provided no relevant information. Twelve studies reported using electrical stimulation during acupuncture.

Usual (or trial-specific) concomitant treatments, as required, were allowed in eight studies[38-45] with no details provided in six studies. Four studies required at least one of the study arms to take specified

doses of analgesics[31 46-48] and two studies[34 36] provided additional education. Three studies allowed no analgesics to be taken.[29 47 49]

Pain was measured using a variety of scales, with WOMAC pain Likert 5 being the most frequent (7 studies). Nine studies reported overall WOMAC scores, and three reported only the individual WOMAC sub-scores. Thirteen studies assessed adverse effects.

5.2.1.2 *Study Quality*

The number of participants randomised was clearly stated in all 22 studies of acupuncture, but only 12 clearly reported using appropriate randomisation methods, and only ten reported suitable methods for concealing treatment allocation. Eligibility criteria were adequately described in 20 trials, and group baseline characteristics appeared comparable in 17. Two studies were reported as being double-blind[29 49] but 14 reported blinding outcome assessors. Just nine studies reported using a power calculation for sample sizes, but 15 clearly reported using data for the intention-to-treat population. Seventeen studies reported whether there were any losses to follow up, with 13 achieving full follow up for at least 90% of participants.

When the overall study quality ratings were derived (see Appendix 10.2) three studies were rated as being of good quality, six were satisfactory, and 13 were rated as being poor. Full details of study quality are reported in Appendix 10.2.

5.2.1.3 *Results of effectiveness*

Pain

Of the 22 trials included in the review 18 provided final value mean data which could be included in the pair-wise meta-analyses and in the NMA (section 6.3). The four trials that could not be included in the analyses[29 39 47 50] were generally similar to the other 18 trials, though one was the only trial to include patients with grade 1 Kellgren and Lawrence scores (though many trials did not report on this characteristic). One of the four trials omitted from analysis was a large (n=330) trial of satisfactory quality.[50] All four trials compared acupuncture with sham acupuncture; one also reported a comparison with standard care.[47] Three were precluded from our analyses as they presented only change from baseline scores and one did not present end of treatment means.[29 39 47 50] All four reported results favouring acupuncture over sham acupuncture.

The 18 studies that provided final value mean data comprised 21 comparisons: eight of acupuncture versus sham acupuncture; nine of acupuncture versus standard care; three of acupuncture versus TENS, and one of acupuncture versus cooling treatment. When all studies (regardless of study quality) were pooled, treatment with acupuncture was associated with statistically significant reductions in end of treatment pain compared to both sham acupuncture and standard care. Similar results were found when only studies of higher (satisfactory or good) quality were analysed (figures 2 to 5).

Most analyses were associated with statistically significant heterogeneity, although the causes were not immediately apparent. All trials (as far as can be seen from the details reported) included populations with similar populations and severity of knee OA. One trial[34] included patients awaiting knee surgery – possibly more severe than a general population for whom acupuncture would be considered. Sensitivity analysis conducted without this trial reduced the statistical heterogeneity of the pooled satisfactory trials to an I^2 value of 5% and increased the pooled treatment effect, SMD: -1.14 (95% CI: -1.33, -0.95). The trial by Lu (2010)[33] investigated the effect of only a single session of acupuncture and therefore this intervention is not comparable with those assessed in the other trials. Sensitivity analysis conducted without this trial only increased statistical heterogeneity (I^2 = 65%) and did not materially alter the pooled effect (-0.54 (95% CI: -0.87, -0.21).

No significant differences between acupuncture and TENS were found for end of treatment pain (three poor quality studies, figure 6). An SMD could not be calculated for the one study of acupuncture versus ice treatment since the acupuncture arm had an end of treatment pain score of zero.

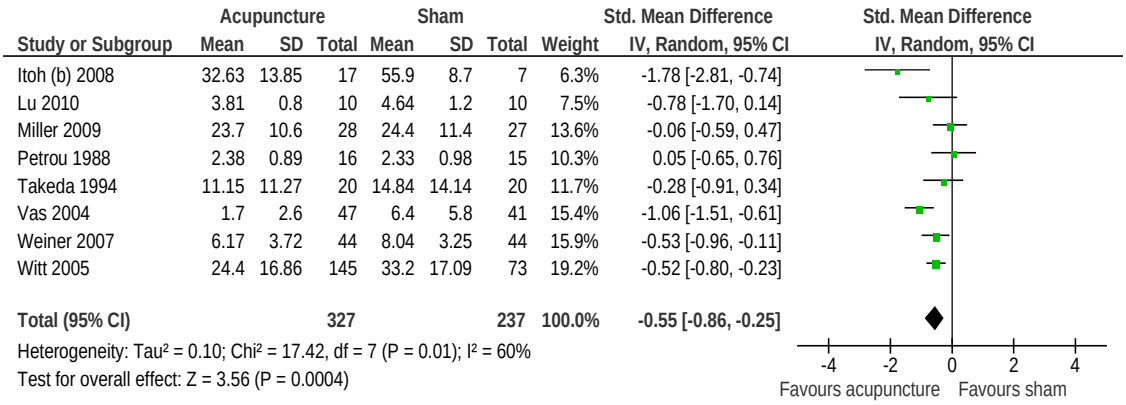


Figure 2: Pain (at end of treatment): acupuncture versus sham acupuncture (all studies).

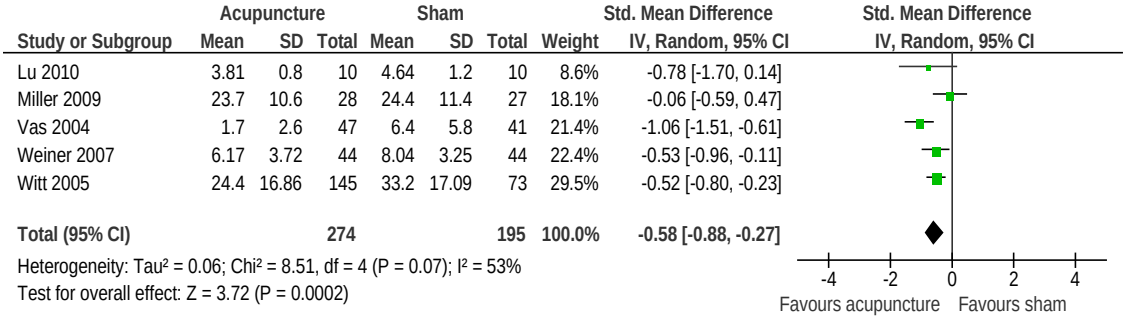


Figure 3: Pain (at end of treatment): acupuncture versus sham acupuncture (good or satisfactory quality studies).

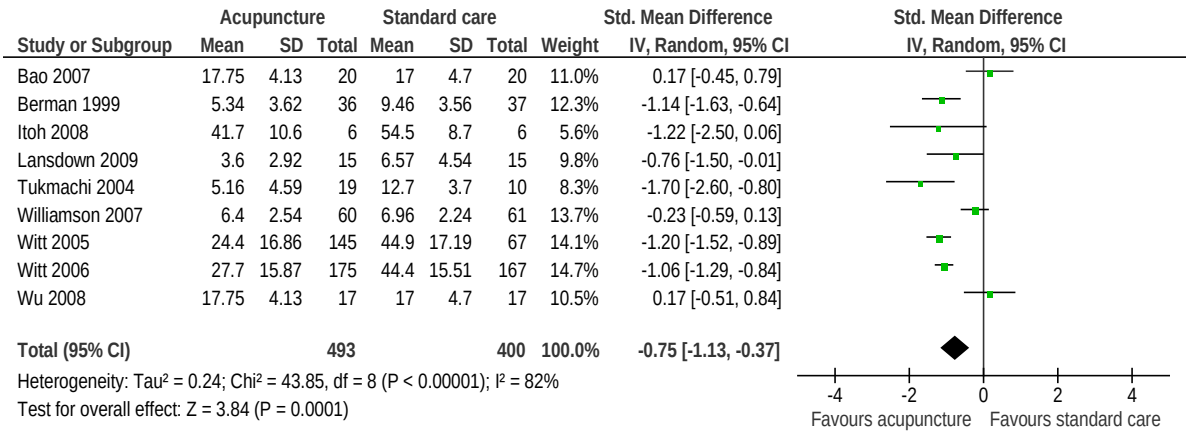


Figure 4: Pain (at end of treatment): acupuncture versus standard care (all studies).

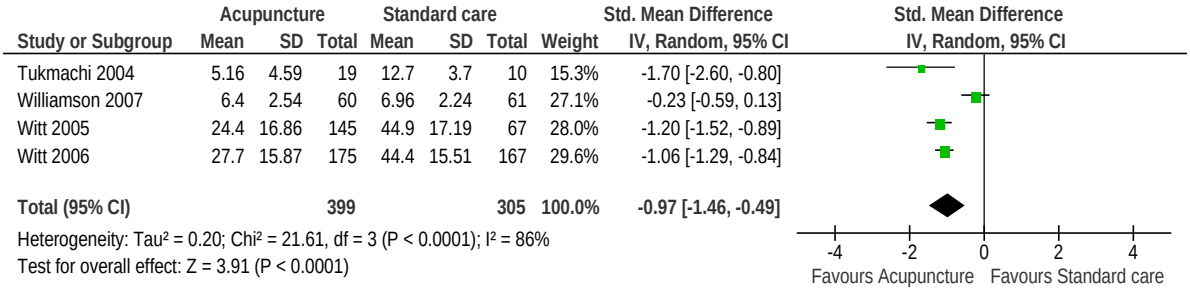


Figure 5: Pain (at end of treatment): acupuncture versus standard care (good or satisfactory quality studies).

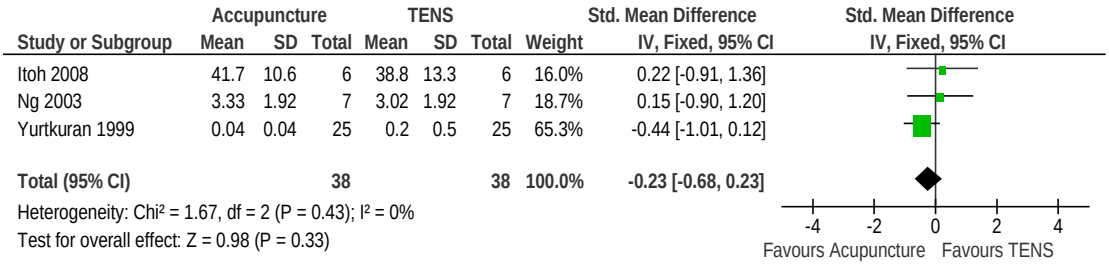


Figure 6: Pain (at end of treatment): acupuncture versus TENS (all studies).

Disability (WOMAC index)

Eight studies provided final value mean data for analysis of the WOMAC index: six provided data for the acupuncture versus standard care comparison, and three provided data for the acupuncture versus sham-acupuncture comparison.

The comparison with standard care (Figure 7) was largely weighted by the two satisfactory quality studies by Witt et al[44 45] and indicates that acupuncture is effective in improving disability; there was no statistical heterogeneity (although the study by Williamson was excluded, as it appeared to cause

heterogeneity in the pain analysis). One study was not suitable for meta-analysis as it only reported changes from baseline.[47]

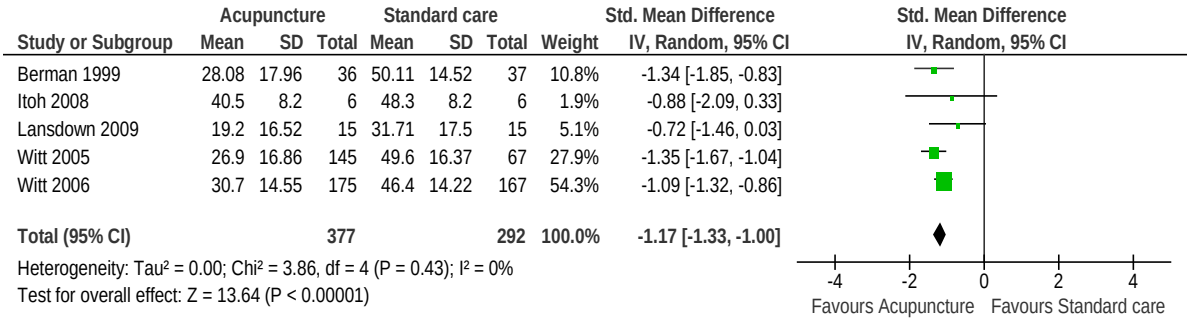


Figure 7: Overall (disability) WOMAC scores: Acupuncture versus Standard care (all studies).

Three studies provided data suitable for the meta-analysis of acupuncture versus sham acupuncture (figure 8). The pooled result suggested significant benefit favouring acupuncture, although this analysis was dominated by one study (of satisfactory quality). One study only reported changes from baseline, so was unsuitable for meta-analysis.[47]

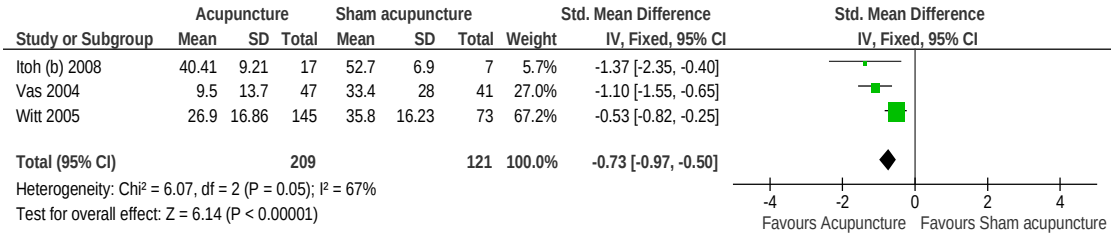


Figure 8: Overall (disability) WOMAC scores: Acupuncture versus Sham acupuncture.

Adverse effects

Of the thirteen studies assessing adverse effects, five stated that no adverse effects were reported by patients, with the remainder either reporting limited specific details, or that occasional minor bruising or bleeding was associated with acupuncture.

Summary of effectiveness of acupuncture

There is evidence from studies of satisfactory quality to suggest that acupuncture is more effective than standard care, and sham acupuncture, in reducing knee OA pain and disability.

Table 4: Acupuncture trials: study details

Author	Treatments (adjunct code, see Table 3)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration ofindividual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Number of points needled	Depth of insertion	Electrical stimulation used?	Pain outcomes assessed	Overall WOMAC score reported?
Bao 2007 [48]	Acupuncture(2) Standard care(4)	40 Y	China	63	62	NR	NR	Clinical	NR/ Unclear	Poor	20	12	4	7	Unclear	Yes	Lysholm scores	No
Berman 1999 [42]	Acupuncture(2) Standard care(2)	73 Y	USA	60	65	NR	mean 32	Clinical and radiological	2 or higher	Poor	20	16	8	5	0.4 to 0.6 inches	Yes	WOMAC pain Likert 5	Yes
Berman 2004 [50]	Acupuncture Sham acupuncture	330 N**	USA	63	66	NR	NR	Clinical and radiological	2 or higher	Satisfactory	20	23	26	9	0.3 to 1.0 inches	Yes	WOMAC pain Likert 5 range 0-20	Other overall score Patient's Global assessment
Itoh 2008 [35]	Acupuncture(1) TENS(1) Standard care(2)	18 Y	Japan	66	range 62-83	NR	NR	Clinical and radiological	2 or higher	Poor	15 to 25 - extra 10 when de qi achieved	5	5	6	10mm	No	Other Pain VAS 10 cm VAS scale (0-100)	Yes
Itoh 2008 [38]	Acupuncture(2) Sham acupuncture(2)	24 Y	Japan	77	73	NR	NR	Clinical and radiological	2 or higher	Poor	30	5	5	6 for standard acupuncture mean of 3.3 for trigger acupuncture	10mm for standard acupuncture 10-30mm for trigger acupuncture	No	Other Pain VAS 100mm	Yes Probably using Likert scale

Author	Treatments (adjunct code, see Table 3)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration ofindividual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Number of points needled	Depth of insertion	Electrical stimulation used?	Pain outcomes assessed	Overall WOMAC score reported?
Jubb 2008 [39]	Acupuncture Sham acupuncture	62 N**	UK	81	65	NR	mean 32	Clinical and radiological	2 or 3	Poor	Manual acu puncture:10 mins Electrical acupuncture : anterior part 10 mins then posterior part10 mins	10	5	9	Varied between 1 and 1.5cm	Yes	WOMAC pain subscale VAS 0-100 Other Pain VAS Overall knee pain (0-100)	Individual WOMAC subs scores reported
Lansdown 2009 [41]	Acupuncture(2) Standard care(2)	30 Y	UK	60	64	NR	NR	Clinical	NR/ Unclear	Poor	10 to 30	up to 10 (as necessary)	Around 10	varied from 4 to 24	Varied from 3 to 30 mm	Unclear/ not stated	WOMAC pain Likert 5	Yes
Lu 2010 [33]	Acupuncture(2) Sham acupuncture(2)	20 Y	China	NR	64	mean 66	NR	Clinical and radiological	2 or 3	Satisfactory	30	1	Unclear/ not stated	5	1-1.5cm	Yes 0.5 mA and 1 ms square pulse and 2 Hz frequenc y.	Other Pain VAS 1-10 VAS (note - SD for post-treatment pain read off graph)	No
Miller 2009 [40]	Acupuncture(2) Sham acupuncture(2)	55 Y	Israel	69	71	NR	NR	Unclear/not stated	NR/ Unclear	Satisfactory	20	16	8	9	Unclear/not stated	Unclear/ not stated	Other pain Likert 5 10 point Likert	No
Molsberger 1994 [29]	Acupuncture Sham acupuncture	97 N*	Germany	63	60	NR	NR	Clinical and radiological	NR (Wirth classification)	Poor	20	10	5	9	0.5-1.5cm	Unclear/ not stated	Other Pain VAS VAS 0-10	No
Ng 2003 [36]	Acupuncture(3) TENS(3) Standard care(3)	14(imputed) Y (only vs TENS)	China	96	85	NR	NR	Clinical	NR/ Unclear	Poor	20	8	2	2	10-15mm	Yes	The Numerical Rating Scale (NRS) of pain,	No
Petrou	Acupuncture(1)	31	Hungary	74	62		NR	Clinical	NR/	Poor	20	8	2	12	Unclear/not	No	Four graded	No

Author	Treatments (adjunct code, see Table 3)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration ofindividual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Number of points needled	Depth of insertion	Electrical stimulation used?	Pain outcomes assessed	Overall WOMAC score reported?
1988 [49]	Sham acupuncture(1)	Y				mean 80			Unclear						stated		standard pain scale (4 points=severe, 3 points=moderate, 2 points=mild, 1 point=none).	
Sangdee 2002 [47]	Acupuncture Sham acupuncture Standard care	186 N**	Thailand	78	63	mean 60	NR	Clinical and radiological	1 or higher	Poor	20	12	4	4	0.5 inch	Yes	WOMAC pain Likert 5 Other Pain VAS Patient's global pain as 100 mm VAS	Yes
Takeda 1994 [51]	Acupuncture(2) Sham acupuncture(2)	40 Y	Canada	50	62	mean 90	Mean 33	Clinical and radiological	NR/ Unclear	Poor	30	9	3	5	30 mm	No	WOMAC pain subscale VAS 0-10 Pain Rating Index (PRI) of McGill Pain Questionnaire.	Individual WOMAC subs scores reported: pain, stiffness and function indices
Tukmachi 2004 [43]	Acupuncture(1 &2) Standard care(2)	29 Y	UK	83	61	NR	NR	Clinical and radiological	2 or 3	Good	20-30	10	5	9	1-1.5cm	Yes	WOMAC pain Likert 5 Likert 0-25 Other Pain VAS VAS 0-10cm	Other overall score Global assessment (VAS 0-10)
Vas 2004 [46]	Acupuncture(4) Sham acupuncture(4)	88 Y	Spain	84	67	NR	mean 33	Clinical and radiological	NR - Used Ahlbäck grade 1 or higher	Good	20	12	12	8	Unclear/not stated	Yes	WOMAC pain Likert 5 Other VAS 0-100	Yes

Author	Treatments (adjunct code, see Table 3)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration ofindividual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Number of points needled	Depth of insertion	Electrical stimulation used?	Pain outcomes assessed	Overall WOMAC score reported?
Weiner 2007 [52]	Acupuncture(2) Sham acupuncture(2)	88 Y	USA	55	71	NR	mean 32	Clinical and radiological	2 or higher	Good	30	6	6	6	until just touch the bone.	Yes (4 points)	WOMAC pain Likert 5	Individual WOMAC subs scores reported: WOMAC function
Williamson 2007 [34]	Acupuncture(2) Muscle strengthening exercise (2) Standard care(3)	181 Y	UK	54	71	NR	mean 32	Clinical and radiological (patients awaiting arthroplasty)	NR/ Unclear	Satisfactory	20	6	6	7 -10	Unclear/not stated	No	Other Pain VAS VAS pain 0-10cm	Yes
Witt 2005 [45]	Acupuncture(2) Sham acupuncture(2) Standard care(2)	285 Y	Germany	66	64	NR	mean 29	Clinical and radiological	2 or higher	Satisfactory	30	12	8	At least 8 for unilateral pain and at least 16 for bilateral pain	Unclear/not stated	No	WOMAC pain subscale VAS 0-100	Yes
Witt 2006 [44]	Acupuncture(2) Standard care(2)	342 Y	Germany	60	61	NR	NR	Clinical and radiological	NR/ Unclear	Satisfactory	Individually prescribed	On average 11	12	Individually prescribed	Unclear/not stated	No	WOMAC pain subscale VAS 0-10	Yes
Wu 2008 [31]	Acupuncture(2) Standard care(4)	34(imputed) Y	China	63	62	NR	NR	Clinical and radiological	NR/ Unclear	Poor	20	12	4	At least 8	Unclear/not stated	Yes	Lysholm score	No
Yurtkuran 1999 [37]	Acupuncture(2) Ice/cooling treatment(2) TENS(2) Placebo TENS(2)	100 Y	Turkey	91	58	NR	NR	Clinical and radiological	NR/ Unclear	Poor Please edit	Unclear	10	2	4	0.5 to 1.0 inches	Yes	Likert 5	No

* No Means, ** Only change from baseline scores reported

5.2.2 **Muscle-strengthening exercise**

5.2.2.1 *Study characteristics*

Thirty trials studied the effectiveness of muscle-strengthening exercise (see Table 5), with a total of 2771 participant pain scores analysed at the end of treatment (range 18 to 366). Twenty-nine were full published papers and one was a conference abstract[53]. All were English language studies, except for one German paper[28], and all were published between 1995 and 2010, half of them since 2005. The majority of studies were undertaken in the UK (six) and USA (five), Denmark (three) and Taiwan (three).

Most studies recruited a general population, but two recruited only patients awaiting knee surgery,[34 54] four recruited only patients with both knees affected by osteoarthritis,[55-58] and one recruited only patients with varus malalignment.[26] The mean age of participants ranged from 53 to 77 years; in 22 studies mean age was between 60 and 69 years. The proportion of females ranged from 31% to 100%; in 17 studies the proportion of females was at least 70%. Mean BMI ranged from 24 to 33 kg/m^2 (14 studies); in nine studies the mean was between 30 and 33 kg/m^2. Mean weight ranged from 55 to 89 kg, though most studies were between 75 and 85kg (eight studies).

The methods used for diagnosis were clinical and radiological in 21 studies, clinical alone in five, radiological alone in three, and the methods were unclear or not stated in one study. Where specified, participants had Kellgren & Lawrence scores of 2 or 3 in two studies, 2 or higher in four studies, 3 or higher in one study and 3 or lower in five. Four studies used other methods of classification including: Ahlbäck criteria; Lequesne score; Altman grading system; and American College of Rheumatology (ACR). Thirteen studies did not report details on classification of severity.

All studies were of land-based exercise (and one study compared land-based with water-based exercise[59]) and in 11 studies home exercise was also incorporated into the intervention (see Table 5). Sessions were undertaken as groups in 10 studies, for individuals in seven studies, and it was unclear or not reported in 13 studies. Where stated, sessions ran from 10 to 80 minutes, though 40-60 minutes was the most frequently reported duration (12 studies). The range of exercise treatment durations for all trials was between 4 and 72 weeks, though for the majority (12 studies) treatment was for 8 weeks, and for all but three studies[60-62] exercise duration was between 4 and 12 weeks. The number of sessions given to participants in the trials ranged from 3 to 216, though half the studies used 16-24 sessions.

The number of review-relevant treatment groups was 2 in 18 studies, 3 in 11 studies and 4 in one study. These studies comprised 33 relevant comparisons with muscle strengthening exercise: 20 were standard care, three were placebo, two were no intervention, two were aerobic exercise, and single comparisons were made with heat treatment, TENS, acupuncture, PES, manual therapy, and NMES. Five studies also compared different types of strengthening exercise with each other, and a relevant comparator.[55 57 59 60 63]

Usual concomitant treatments (as required) were allowed in 13 studies, but in 12 studies such details were unclear or not stated. Education or advice was given as background care in six studies,[34 62 64-67], a specific programme of home exercise (which was not part of the active intervention) was used in two studies,[65 68] and a self-management component was included as part of the exercise sessions in one study.[69] Two studies stated that no medication was permitted.[55 57] One study required (some) participants to take a daily dose of aceclofenac.[70]

Pain was most frequently measured using a VAS 0-10 scale (10 studies), a VAS 0-100 scale (5 studies), a WOMAC VAS 0-100 (4 studies) and a WOMAC 5-point Likert scale (8 studies); five additional measures of pain outcome were also used. An overall WOMAC score was reported in five studies, with WOMAC sub-scores reported in six studies; these included pain in five studies, function in five studies, stiffness in two studies and function in one.

Quality of life data were reported in seven studies (SF36 in four studies[61 67 68 71] one study used EQ-5D, MACTAR (McMaster Toronto Arthritis) and HADS[69], and single studies used AIMS[72] and Knee injury and Osteoarthritis Outcome Score (KOOS).[59] Adverse effects were assessed in eight studies.[26 34 58 59 62 63 68 71]

5.2.2.2 *Study Quality*

Although the number of participants randomised was stated for all 30 trials of muscle-strengthening exercise, only half clearly reported using appropriate methods for randomisation, and only a third used appropriate methods for concealing treatment allocation. One study used cluster randomisation.[69] Eligibility criteria were adequately described in 77% of studies and group baseline characteristics appeared comparable in 61% of studies; the study with two different populations[26] reported comparable baseline characteristics for the neutral knee alignment population, but not for the mal-aligned population. Only one study was described as being double blinded, highlighting the difficulties of blinding treatment-givers and patients in exercise trials, and only 40% of studies reported using blinded outcome assessors. Use of a power calculation for sample sizes was reported in 43% of studies and half the studies reported data for the intention-to-treat population. Seventy per cent of studies clearly reported whether there were any losses to follow up, but only 47% achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived 71% of studies were rated as being of poor quality, and 29% were satisfactory. The study with two different populations[26] was rated as poor for the mal-aligned population and satisfactory for the neutral population. Full details of study quality are reported in Appendix 10.2.

A funnel plot revealed no indication of publication bias (see figure 9)

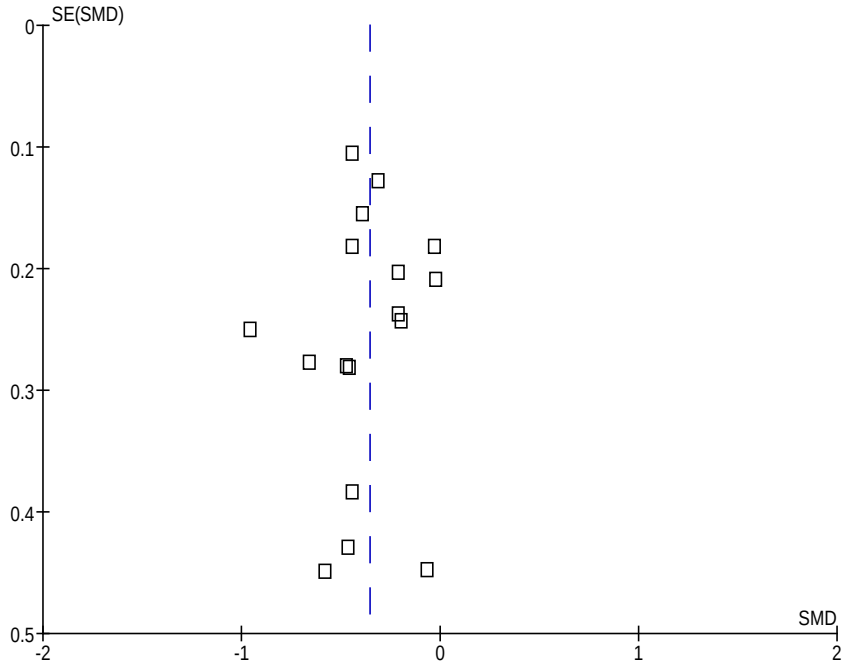


Figure 9: Funnel plot for muscle-strengthening exercise versus standard care comparison.

5.2.2.3 *Results of effectiveness*

Pain

Six trials did not provide mean data following treatment so could not be included in the meta-analyses nor the NMA.[53 58 64 65 70 73] All were rated to be of poor quality. Of these, five reported only medians, and one presented only change from baseline scores. One further study presented only mean differences between intervention and control groups, so could be included in the NMA, but not in the meta-analyses.[63] These seven trials were generally comparable with those that could be analysed, except that one[63] was of whole body vibration exercise and one included only patients whose Kellgren & Lawrence scores were 3 or higher.[58] Most were small trials, though one had a sample size of 217,

and two were of satisfactory quality. These studies either did not report analyses comparing treatment group pain scores, or found no significant differences between groups.

Twenty-three studies reported final mean value data for 23 comparisons: 19 for MSE versus standard care; two MSE versus placebo and two for MSE versus no intervention. The analyses for all studies, and only higher-quality studies (Figures 10 & 11), both showed significant benefit favouring muscle-strengthening exercise over standard care for reducing end of treatment pain. The results of subgroup analyses of trials that did, or did not, incorporate home exercise into the intervention also showed significant benefit favouring muscle-strengthening exercise over standard care when were analysed separately: -0.32 (95% CI: -0.47, -0.17) I^2=0% and -0.33 (95%CI: -0. 57, -0. 22) I^2= 23% respectively. There was almost no statistical heterogeneity associated with the analysis of all studies, or the analysis of studies with and without home exercise, but a high level of statistical heterogeneity was associated with the meta-analysis of higher-quality studies, which may partly be a consequence of the narrower confidence intervals associated with the individual effect estimates. This analysis also included two trials of patients awaiting surgery, whose OA severity might be higher than the general population; statistical heterogeneity was not eliminated by removal of these two trials from the analysis and the resultant pooled treatment effect may be an overestimate (-0.63 95% CI: -0.97, -0.28) .

The two trials comparing MSE with no intervention both found a significant beneficial treatment effect favouring MSE. However, the pooled effect (SMD -0.91, 95% CI -1.33 to -0.50) was associated with a very high degree of heterogeneity (I^2= 86%) such that it cannot be viewed as being reliable. There was a mean weight difference of 16kg between the study populations (one Taiwan, the other Turkey), which may explain the heterogeneity. One small poor quality trial compared muscle-strengthening exercise with placebo exercise[61] SMD -0.63 (95% CI: -1.29, 0.02). Another small poor quality trial compared muscle-strengthening exercise with TENS[74] and the SMD was -0.42 (95% CI: -0.29, 1.13). One study only provided data after six months' follow up (there were no end of treatment results).[69]

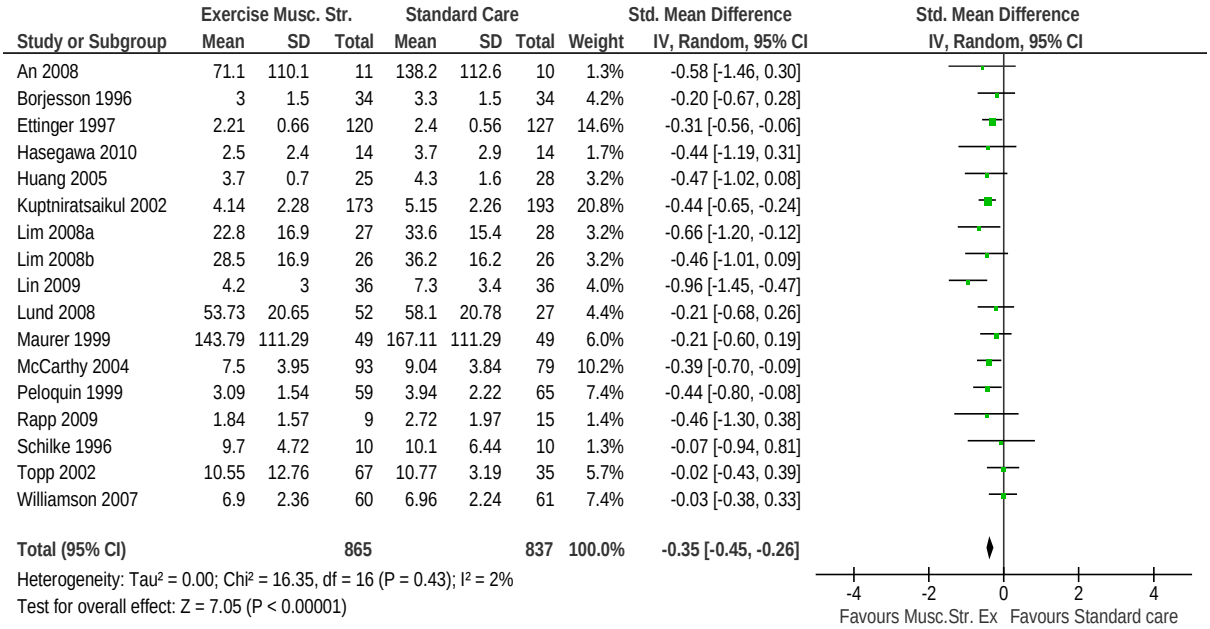


Figure 10: Pain (at end of treatment): muscle-strengthening exercise versus standard care (all studies).

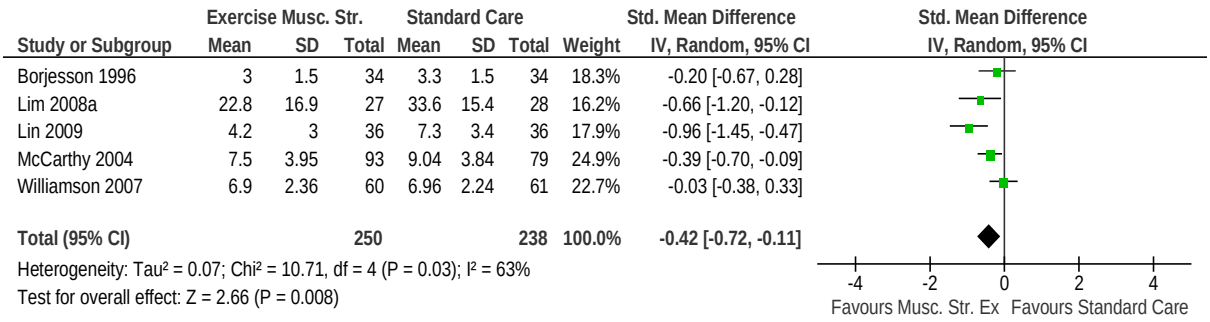


Figure 11: Pain (at end of treatment): muscle-strengthening exercise versus standard care (good or satisfactory quality studies).

Disability (WOMAC index)

Two studies provided overall WOMAC scores at the end of treatment suitable for analysis. For one there was no significant difference between MSE and standard care (SMD -0.13, 95% CI -0.48 to 0.23)[34] and for the other MSE significantly reduced the WOMAC index compared to shortwave diathermy heat treatment (SMD -0.71, 95% CI -1.28 to -0.13).[75]

Adverse effects

Of the nine trials assessing adverse effects, four stated that none were reported, with the remainder reporting limited specific details, or that knee pain, or falls were occasionally associated with muscle-strengthening exercise.

Summary of effectiveness of muscle-strengthening exercise

There is evidence from studies of satisfactory quality to suggest that MSE is more effective than standard care, in reducing knee OA pain.

Table 5: Muscle-strengthening exercise trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Abrahams 2002 [70]	MSE Standard care	56 N*	UK	NR	range 22-83	NR	NR	Clinical and radiological	NR/Unclear	Poor	Unclear/ NR	Unclear/NR	12	Unclear/NR	Other Pain VAS VAS slide indicator version (0-100)but no data reported	No
An 2008 [71]	MSE(2) Standard care(2)	21 Y	China	100	65	NR	26	Clinical	Other method used to classify OA severity ACR Criteria	Poor	30	40	8	No	WOMAC pain subscale VAS 0-100	No
Baker 2001 [61]	MSE(2) Placebo (2)	38 Y	USA	78	69	NR	32	Clinical and radiological	NR/Unclear median 3	Poor	Unclear/ NR	48 (patients exercised 3 times per week for 16 weeks at home, with 12 supervised visits)	16	Yes	WOMAC pain subscale VAS 0-100	Pain and physical function subs scores reported
Bezalel 2010 [75]	MSE(2) Heat treatment(2)	50 Y	Israel	74	74	NR	NR	Unclear/NR	NR/Unclear	Poor	45	4	4	Yes	WOMAC pain Likert 5 Data for post treatment and follow-up read off graph - not very precise	Yes
Borjesson 1996 [54]	MSE(2) Standard care(2)	68 Y	Sweden	50	64	83	NR	Clinical and radiological	Other method used to classify OA severity Ahlbäck grade I-III	Satisfactory	40	15	5	Yes	Borg scale (11-garde category scale 0-no pain to 10-worse pain).	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Callaghan 1995 [65]	MSE Standard care Placebo	27 N*	UK	31	median 53	NR	NR	Radiological	NR/Unclear	Poor	20	8	4	No	Other Pain VAS 0-10 scale	No
Cheing 2002 [74]	MSE(2) TENS(2) Placebo TENS(2)	47 Y	China	89	63	67	28	Clinical and radiological	2 or higher	Poor	20	20	4	No	Other Pain VAS VAS 0-100, with baseline score standardised to 100	No
Durmus 2007 [76]	MSE(2) PES(2)	50 Y	Turkey	100	55	NR	33	Clinical and radiological	3 or lower	Satisfactory	20.	20 sessions	4	No	WOMAC pain subscale VAS 0-10 Other Pain VAS VAS 0-10	All 3 Individual WOMAC subs scores reported
Ettinger 1997 [62]	MSE(2) Aerobic exercise(2) Standard care(3)	364 Y	USA	70	69	NR	53% >30kg/m2	Clinical and radiological	NR/Unclear	Poor	60	216	72	Yes	Other pain Likert 5 Likert 1 (no pain) to 6 (excruciating pain)	No
Gur 2002 [55]	MSE(1) No intervention(1)	23 Y	Turkey	NR	56	79	NR	Radiological	2 or 3	Poor	Unclear/ NR	24	8	No	Other Pain VAS Numeric rating scale 0-10 (10 unbearable pain). the score is the sum of scores for pain at night, after inactivity, on sitting, on rising from a chair, climbing stairs, descending stairs	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Hasegawa 2010 [77]	MSE(2) Standard care(2)	28 Y	Japan	64	77	55	24	Clinical diagnosis was of Knee joint Pain (KJP)	NR/ Unclear	Poor	80 (60 of exercise plus 20 of warm up/down)	12 (also a minimum of 24 home exercise sessions)	12	Yes	Other Pain VAS 0-10 NRS (pain on movement)	No
Hay 2006 [64]	MSE Standard care	217 N**	UK	65	68	NR	NR	Clinical Patients with knee pain – knee OA not specified	NR/ Unclear	Poor	20	3 to 6	10	Yes	WOMAC pain Likert 5 Other Pain VAS (specify) Pain severity over previous 7 days on scale 0-10	No
Huang 2005 [56]	MSE(2) Standard care(2)	98 Y	Taiwan	81	62	NR	NR	Clinical and radiological	NR - Patients with Altman grade II were included	Poor	Unclear/ NR	24	8	Yes but only after 8-week treatment i.e. during post treatment follow-up.	Other Pain VAS Other pain VAS 1-10	No
Hurley 2007 [69]	MSE Standard care	53 Y	UK	70	67	81	30	Clinical	NR/Unclear	Satisfactory	35 to 45	12	6	No, but did include self-management and coping skills	WOMAC pain Likert 5	Yes
Jan 2008 [57]	MSE(1) No intervention(1)	98 Y	Taiwan	81	63	63	NR	Clinical and radiological	3 or lower	Satisfactory	30 (HR) and 50 (LR)	24	8	No	WOMAC pain Likert 5	Only physical subscale reported
Keogan 2007 [53]	MSE Aerobic exercise Standard care	80 N**	Republic of Ireland	64	66	NR	30.6	Clinical and radiological	NR/Unclear	Poor	Unclear/ NR	Unclear/ NR	6	Unclear/NR	Other Pain VAS NR, VAS 0-10 more likely has been used.	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Kuptnira-saikul 2002 [72]	MSE(2) Standard care(2)	366 Y	Thailand	78	68	NR	11% were obese	Radiological	2 or 3	Poor	60	16	8	Unclear/NR	Other AIMS pain subscale (0-10)	No
Lim 2008 [26]	MSE(2) Standard care(2)	107 Y	Australia	Group with more neutral knee alignment: 62 Group with more varus knee malalignment: 48	Group with more neutral knee alignment: mean 62 years Group with more varus knee malalignment: mean 67 years	Group with more neutral knee alignment: mean 78 Group with more varus knee malalignment: mean 81	Group with more neutral knee alignment: mean 29 Group with more varus knee malalignment: mean 29	Clinical and radiological	2 or higher	Satisfactory For neutral group Poor for malaligned group	Unclear/ NR	60 (7 with physiotherapist at weeks 1, 2, 3, 4, 5, 7 and 10)	12	Yes Most exercise was home based, but participants also visited the physiotherapist seven times during the 12 week period.	WOMAC pain Likert 5 (transformed to a 0-100 scale)	No
Lin 2009 [78]	MSE(2) Standard care(2)	72 Y	Taiwan	69	63	62	NR	Clinical and radiological	3 or lower	Satisfactory	Unclear/ NR	24	8	No Told to cease other exercises.	WOMAC pain Likert 5	Pain and function sub scores reported
Lund 2008 [59]	MSE(2) Standard care(2)	79 Y	Denmark	78	68	75	NR	Clinical and radiological	Other method used to classify OA severity Lequesne (1-26)score. Mean score	Poor	50	16	8	Unclear/NR	Other Pain VAS VAS 0-100	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
									of participants was 11.							
Maurer 1999 [67]	MSE(2) Standard care(3)	98 Y	USA	42	65	85	NR	Clinical and radiological	3 or lower	Poor	Unclear/ NR	24	8	Unclear/NR	WOMAC pain subscale VAS 0-100	Pain and function sub scores reported
McCarthy 2004 [68]	MSE(3) Standard care(3)	172 Y	UK	58	65	NR	30	Clinical and radiological		Satisfactory	45	16	8	Yes	WOMAC pain Likert 5 Other Pain VAS VAS pain score 0-100	No (only at baseline)
Peloquin 1999 [66]	MSE(3) Standard care(3)	124 Y	Canada	70	66	NR	30	Clinical and radiological	3 or lower	Poor	60	36	12	No	Other Arthritis Impact Measurement Scale (AIMS2) subscale for pain.	No
Rapp 2009 [28]	MSE(2) Manual therapy(2) Standard care(2)	39 Y	Germany	64	60	83	NR	Clinical and radiological	2 or higher	Poor	45	16	8	Unclear/NR	Other Pain VAS VAS 0-10	No
Røgind 1998 [58]	MSE Standard care	23 N*	Denmark	91	71	71	27	Clinical and radiological	3 or higher	Poor	Unclear/ NR	24 or 26 (2 per week for 3 months)	12	Yes	Other Pain VAS Pain on an 11 point. Separate scores for pain at night, at rest and on weight bearing.	No
Rosemffet 2004 [73]	MSE NMES	18 N*	Argentin a	77	median 60	NR	30.9	Clinical and radiological	2 or higher	Poor	75	16	8	Unclear/NR	Other Pain VAS 20-80 mm VAS scale	Yes
Schilke	MSE(2)	Y	USA	85	66	NR	NR	Clinical	NR/Unclear	Poor	Unclear/	24	8	No	Other Pain VAS	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment Datsuitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Intervention also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
1996 [79]	Standard care(2)	20									NR				OASI 10cm VAS	
Topp 2002 [60]	MSE(2) Standard care(2)	102 Y	USA	73	63	89	NR	Clinical and radiological	NR/Unclear	Poor	50	48	16	Yes	WOMAC pain Likert 5	No
Trans 2009 [63]	MSE(2) Standard care(2)	52 Y	Denmark	100	60	81	30	Clinical and radiological	NR/Unclear	Satisfactory	Up to 10.5	16	8	No Both interventions	WOMAC pain subscale VAS 0-100	All 3 individual WOMAC subs scores reported
Williamson 2007 [34]	MSE(2) Acupuncture(2) Standard care(3)	181 Y	UK	54	71	NR	32	Clinical and radiological Patients on NHS arthroplasty waiting list	NR/Unclear	Satisfactory	60	6	6	No	Other Pain VAS VAS pain 0-10cm	Yes

* No Means, ** Only change from baseline scores reported MSE=Muscle-strengthening exercise

5.2.3 **Aerobic exercise**

5.2.3.1 *Study characteristics*

Nine trials studied aerobic exercise interventions (see Table 6), with over 880 participant pain scores analysed at the end of treatment (range 25 to 364, sample size details not provided in two trials). Seven were full published papers and two were conference abstracts[53 80]. All studies were reported in English, between 1992 and 2009. Four studies were conducted in the United States.
Six studies recruited a general population, two studied only participants with both knees affected by osteoarthritis[80 81], and one studied only overweight or obese participants[82]. The mean/median ages of participants ranged from 54 to 75 years, and the proportion of females ranged from 50 to 100%. Mean BMI was reported in only four studies (range 30 to 34 kg/m^2), and mean weight in only two studies.

The methods used for diagnosis were clinical and radiological in seven studies, and clinical alone in two studies. Only three studies reported using Kellgren & Lawrence scores with the remaining studies not reporting details on classification of severity.

All nine trials compared land-based aerobic exercise with standard care; two multi-armed studies[53 62] also used muscle-strengthening exercise interventions, and one[82] multi-armed study also used a dieting weight loss intervention. Sessions generally lasted for around an hour (range 20 to 90 minutes). The total number of sessions varied greatly from 12 to 234, over periods ranging from six to 78 weeks. Where stated, most studies ran sessions to groups rather than individuals, and four studies[62 82-84] used additional home exercise as part of the intervention. Usual (or trial-specific) concomitant treatments, as required, were allowed in four studies[62 81 83 85], and no details were provided in two studies.[32 76] Four studies[62 82 84 86] provided additional education. One study[80] provided a home exercise plan for the standard care group, and in one study[53] background care details were not provided.

Pain was measured using a variety of scales; only three studies measured WOMAC pain (two using a VAS 0-10 scale, and one using a Likert scale). Only one study reported overall WOMAC scores, one reported individual WOMAC sub-scores, one reported Arthritis Self-efficacy Scale scores and the remaining studies did not report on overall assessment of disability. Adverse effects were only assessed in two studies.[62 82]

5.2.3.2 *Study Quality*

The number of participants randomised was clearly stated in all but one[80] study, but only four[62 81 82 86] clearly reported using appropriate randomisation methods, and only two studies[62 82] reported suitable methods for concealing treatment allocation. Eligibility criteria were adequately described in six trials, but group baseline characteristics appeared comparable in only four, and just two studies[62 82] reported blinding outcome assessors. Three studies[62 82 83] reported using a power calculation for sample sizes, and three [53 62 82] clearly reported using data for the intention-to-treat population. However, all studies but two[53 80] reported whether there were any losses to follow up. Four studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived eight studies were rated as being of poor quality, and one[82] was rated as being of satisfactory quality. Full details of study quality are reported in Appendix 10.2.

5.2.3.3 *Results of effectiveness*

Pain

Four trials did not provide final value data for inclusion in analyses.[53 80 83 84]. Generally they were similar to the remaining trials except that they included the one trial that provided a home exercise plan for the standard care group[80] and one specified grade 3 or higher K+L severity which may be more severe than the other trials. All four were comparisons with standard care, though one also incorporated a comparison with MSE.[53]

Five studies reported final value data for the comparison of aerobic exercise versus standard care; one for aerobic exercise versus MSE[62], and one for aerobic exercise versus weight loss[82].

Pooling all available trials (five of the nine studies) found that treatment with aerobic exercise interventions was associated with a statistically significant reduction in end of treatment pain, compared to standard care (Figure 12) but the analysis was subject to an extremely high level of statistical heterogeneity (88%). Removal of one very small poor quality study reporting a very large treatment effect[81], reduced the I^2 value from 88% to 61% (SMD -0.32, 95% CI -0.62 to -0.01). The other main source of heterogeneity was the one satisfactory quality study, which found no difference in end of treatment pain (Messier).[82] In this trial the population was restricted to overweight or obese participants and also it was the only one where the aerobic exercise intervention was reported to have been delivered to individuals.

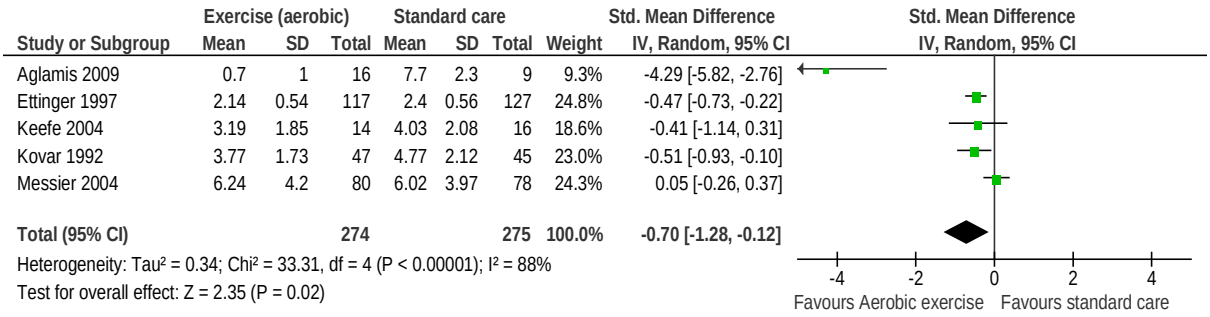


Figure 12: Pain at end of treatment: Aerobic exercise versus standard care.

Disability (WOMAC index)

Only one study reported overall WOMAC, but used change from baseline scores.[81]

Adverse effects

The two studies which assessed adverse effects reported tripping or falling in a very small minority of participants.

5.2.3.4 *Summary of effectiveness of aerobic exercise*

There was no evidence of satisfactory quality to suggest aerobic exercise was more effective than standard care in reducing knee OA pain.

Table 6: Aerobic exercise trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	Mean BMI kg/m^2	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Individual or group sessions?	Also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Aglamis 2009 [81]	AerEx(2) Standard care(2)	25 Y	Turkey	100	56	NR	33	Clinical and radiological	2 or higher	Poor	20	36	12	Unclear/ NR	No	WOMAC pain subscale VAS 0-10	Yes
Bilgici 2004 [80]	AerEx Standard care	NR N*	Turkey	NR	54	NR	NR	Clinical	NR/ Unclear	Poor	60	16	8	Unclear/ NR	Unclear/ NR	WOMAC pain subscale VAS 0-10 Unclear what form of WOMAC was used: no data reported	No
Dias 2003 [84]	AerEx Standard care	NR N*	Brazil	88	median 74, 76	NR	NR	Clinical and radiological	NR/ Unclear	Poor	40	12	6 (followed by 6 weeks home exercise)	Sessions given to groups	Yes for 6 weeks, following the 6 weeks intervention	SF-36 Bodily Pain Insufficient pain outcome data reported for data extraction	No
Ettinger 1997 [62]	AerEx(2) MSE(2) Standard care(3)	364 Y	USA	70	69	NR	53% >30kg/m2	Clinical and radiological	NR/ Unclear	Poor	60	216	72	Sessions given to groups	Yes	Other pain Likert 5 (specify) Likert 1 (no pain) to 6 (excruciating pain)	No
Keefe 2004 [85]	AerEx(2) Standard care(2)	30(imputed) Y	USA	50	59	NR	NR	Clinical	NR/ Unclear	Poor	60	36	12	Sessions given to groups	No	Other (specify) Arthritis Impact Measurement Scales (AIMS)4-item Pain	No, but Arthritis Self-efficacy Scale used
Keogan 2007 [53]	AerEx MSE Standard care	80 N**	Republic of Ireland	64	66	NR	31	Clinical and radiological	NR/ Unclear	Poor	Unclear/ NR	Unclear/ NR	6	Unclear/ NR	Unclear/ NR	Other Pain VAS (specify) NR, VAS 0-10 more likely has been used.	No
Kovar 1992 [86]	AerEx(2) Standard care(3)	92 Y	USA	83	69	mean 77	NR	Clinical and radiological	NR/ Unclear	Poor	90	24	8	Sessions given to groups	No	Other (specify) Arthritis Impact Measurement Scale (AIMS) VAS 10	No

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	Mean BMI kg/m^2	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Individual or group sessions?	Also incorporated home exercise?	Pain outcomes used	Overall WOMAC score reported?
Messier 2004 [82]	AerEx(2) Weight loss(2) Standard care(3)	240 Y	USA	72	69	Mean 94	34	Clinical and radiological	3 or lower	Satisfactory	60	234	78	Sessions given to individuals	Yes	WOMAC pain Likert 5	Pain & physical function subs scores reported
Thorstensson 2005 [83]	AerEx Standard care	56 N**	Sweden	51	56	NR	30	Clinical and radiological	3 or higher	Poor	60	12	6	Sessions given to groups	Yes	Other Pain VAS (specify) KOOS	No

* No means ** Only change from baseline scores reported NR Not reported, Aer Ex= Aerobic exercise, MSE =Muscle-strengthening exercise

5.2.4 **Tai Chi**

5.2.4.1 *Study characteristics*

Five trials studied Tai Chi (see Table 7) with a total of 336 participant pain scores analysed at the end of treatment (range 29 to 182); all were full published papers in English between 2007 and 2010. Two of the five studies were conducted in the United States and the rest in the Far East (China, Hong Kong and Korea). All participants were derived from a general knee OA population and their mean age ranged from 63 to 70 years, and the proportion of females ranged from 75 to 100%. Mean BMI was reported in four studies (range 26 to 30), and mean weight in three studies (range 61 to 73kg). The methods used for diagnosis were clinical in three studies and clinical and radiological in two studies. Use of Kellgren & Lawrence scores was reported in two studies, both of which reported a score of at least two.

All studies were two-armed trials comparing Tai Chi with standard care. Sessions lasted for between 40 and 60 minutes (with the exception of one study [87] where 15 minutes of exercise were undertaken within a 2 hour self-management session), for between six and 72 sessions, over periods ranging from six to 24 weeks. Concomitant treatments comprised education and exercise, or were unclear.

Where specified, pain was measured using a VAS 0-100 scale in one study [88] and a WOMAC 5 point Likert scale in one study [89]. WOMAC pain subscales were used in two studies [90 91]. Quality of life data were reported in three studies (two used SF-36 domains [88 90], and one used HAQ.[87] Adverse effects were assessed in three studies.[88 89 91]

5.2.4.2 *Study Quality*

The number of participants randomised was clearly stated in all studies and all studies used appropriate methods for randomisation, though only two studies reported using appropriate methods for concealing treatment allocation.[88 89] Eligibility criteria were adequately described in all five trials, and group baseline characteristics appeared comparable in three studies but were not comparable[88] or unclear[87] in single studies.
None of the studies were double-blind and none had either patient or treatment-giver blinding. In three studies outcome assessors were blinded[89-91], but it was unclear in one study[88] and outcome assessors were not blinded in another.[87]

Only two studies reported use of a power calculation for sample sizes[87 91], four studies reported data for the intention-to-treat population[87 88 90 91] and all studies reported losses to follow up. Only two studies achieved full follow up for at least 90% of participants.[88 90]

When the overall study quality ratings were derived, three studies were rated as being of poor quality[87-89] and two of satisfactory quality.[90 91] Full details of study quality are reported in Appendix 10.2.

5.2.4.3 *Results of effectiveness*

Pain

Four of the five trials provided data suitable for analysis;[87 89-91] the study which failed to provide these data[88] was of poor quality. One study was the largest trial of Tai Chi, but the intervention was somewhat different from that studied in the other trials as it involved 15 minutes of Tai Chi as part of a two hour self-management session, with only six sessions over six weeks given.[87] Meta-analysis including this study showed a significant reduction in pain with Tai Chi (figure 13). During the analyses it was found that the poor quality study by Ni et al (2010) [89] had been retracted by its publishing journal; doubts exist about whether the trial actually took place. Meta-analysis of the two satisfactory quality studies indicated no significant improvement in pain with Tai Chi, when compared with standard care (Figure 14).

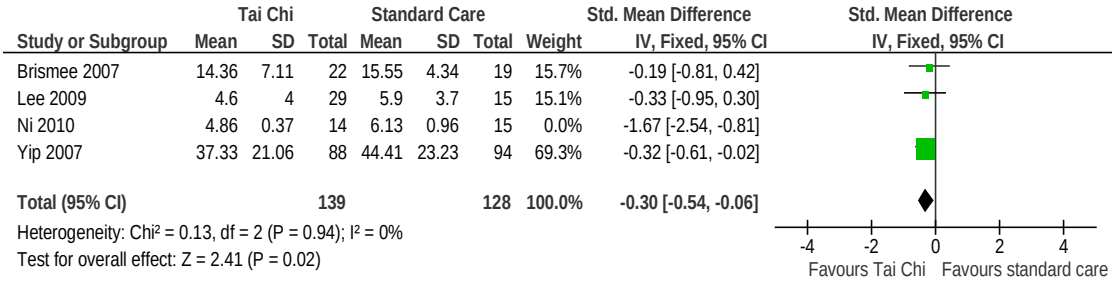


Figure 13: Pain (at end of treatment): Tai Chi versus standard care (all studies).

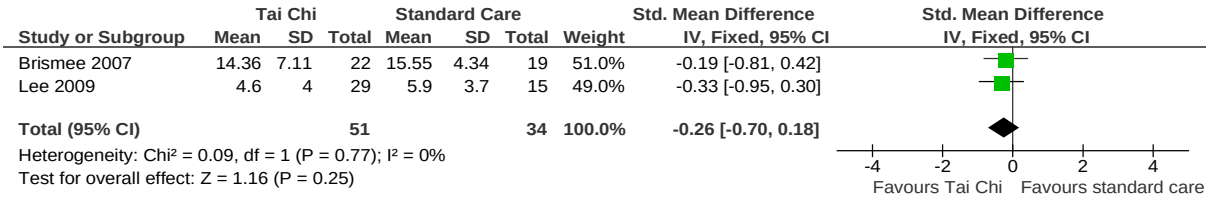


Figure 14: Pain (at end of treatment): Tai Chi versus standard care (satisfactory quality studies).

Disability (WOMAC index)

Two studies provided overall WOMAC scores at the end of treatment suitable for analysis.[90 91] When results from these studies were pooled there was no significant difference between Tai Chi and standard care (figure 15).

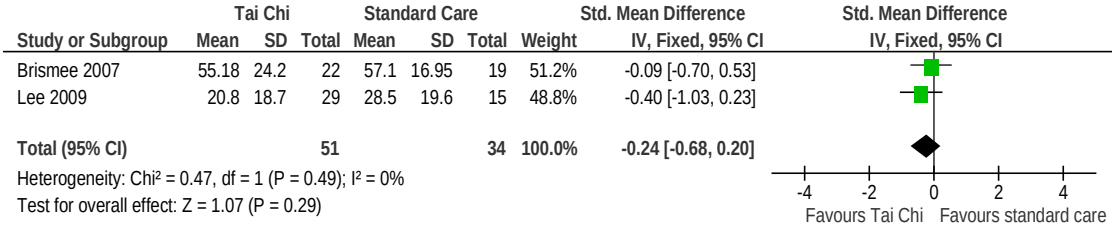


Figure 15: WOMAC Index (at end of treatment): Tai Chi versus standard care (both satisfactory quality studies).

Adverse effects

Two studies assessed adverse effects. In one study there were sporadic complaints of minor muscle soreness, foot and knee pain in the Tai Chi group. In the third study an increase in knee pain was reported by a single patient in the Tai Chi group but this was resolved.

5.2.4.4 *Summary of effectiveness of Tai Chi*

There was no evidence of satisfactory quality to suggest Tai Chi was more effective than standard care in reducing knee OA pain.

Table 7: Tai Chi trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session	Total number of sessions	Duration of treatment period (weeks)	Pain outcomes used	Overall WOMAC score reported?
Brismee 2007 [91]	Tai Chi(2) Standard care(3)	41 Y	USA	83	70	73	28	Clinical	NR	Satisfactory	40 minutes	36 (18 classes and 18 home sessions)	12	WOMAC pain subscale VAS 0-100 (reported as 7 to 35) VAS 0-10	Yes (scale 26-130)
Wang 2009 [88]	Tai Chi Standard care	40 N*	USA	75	65	NR	30	Clinical and Radiological	2 or higher	Poor	60 minutes	*24*	12	WOMAC pain subscale VAS 0-100 VAS 0-10	All 3 individual WOMAC subs scores reported
Lee 2009 [90]	Tai Chi(2) Standard care(2)	44 Y	Korea	93	69	61	26	Clinical and Radiological	2 or higher	Satisfactory	60 minutes	16	8	WOMAC used but reported as 26-130	Yes (scale 26-130) All 3 individual WOMAC subs scores reported
Ni 2010 [89] ***RETRACTED STUDY***	*Tai Chi(2) Standard care(3)*	*29 Y*	*China*	*100*	*63*	*66*	*27*	*Clinical*	*NR*	*Poor*	*40 minutes*	*72*	*24*	*WOMAC pain Likert 5*	*Yes*
Yip 2007 [87]	Tai Chi Standard care	182 Y	Hong Kong	84	65	NR	NR	Clinical	NR	Poor	15 minutes (within a 2 hour self-management session)	6	6	VAS 0-100	No

* No Means

5.2.5 **Weight loss (dieting)**

5.2.5.1 *Study characteristics*

Four trials studied weight loss interventions (see Table 8), with a total of 781 participant pain scores analysed at the end of treatment (range 74 to 389). All were full papers published in English between 2004 and 2009. Two studies were conducted in the United States.

All studies recruited overweight or obese participants. The mean ages of participants ranged from 61 to 70 years, and the proportion of females ranged from 26 to 89%. Mean/median BMIs ranged from 33 to 36 kg/m^2, and mean/median weight from 93 to 98kg. The methods used for diagnosis were clinical and radiological in two studies, and clinical alone in two studies. Three studies reported using Kellgren & Lawrence scores, which varied considerably by study.

All four trials compared weight loss using dieting with standard care; one three-armed trial [82] also studied an aerobic exercise intervention. The duration of dieting periods ranged from eight to 104 weeks. The differences in weight loss between dieting and standard care groups ranged between 3.5 and 8 kg.

Usual concomitant treatments, as required, were allowed in one study[92]; one study[93] had two standard care arms, with one group receiving home exercise and the other educational leaflets; education was also given to both groups in one study[94] and just to the standard care group in another.[82]

Pain was measured using a WOMAC Likert scale in three studies and a WOMAC VAS (0-100 scale) in one study. Two studies reported overall WOMAC scores, with the remaining two studies not reporting an overall assessment of disability. One study[82] reported adverse effect details**.**

5.2.5.2 *Study Quality*

The number of participants randomised was clearly stated in all but four studies, but only two[82 93] clearly reported using appropriate randomisation and treatment allocation concealment methods. Eligibility criteria were adequately described in all four trials, and all trials also had comparable group baseline characteristics. One study[82] reported blinding outcome assessors.

Only one study[94] failed to report using a power calculation for sample sizes, but only two studies[82 93] clearly reported using data for the intention-to-treat population. However, all studies reported whether there were any losses to follow up, but none achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived two studies were rated as being of poor quality, and two were rated as being of satisfactory quality.[82 93] Full details of study quality are reported in Appendix 10.2.

5.2.5.3 *Results of effectiveness*

Pain

One poor quality trial did not provide final value mean data and could not be included in the standard or NMA.[92] This trial had used the shortest treatment duration (only eight weeks) but it was the only one to report the percentage of patients who had lost 10% of body weight: its findings suggested significant benefit from weight loss, compared to standard care.

The pooled result from the three trials that provided final value mean data[82 93 94] found no effect of weight loss on pain, when compared to standard care (Figure 16). However this analysis was subject to considerable statistical heterogeneity (I^2 =79%) and probable clinical heterogeneity given that the weight loss intervention varied or was unclear across the three trials. The Miller (2006) trial, which reported the greatest treatment effect, whilst being the only poor quality trial, did report the largest weight loss (8 kg - see table 8).

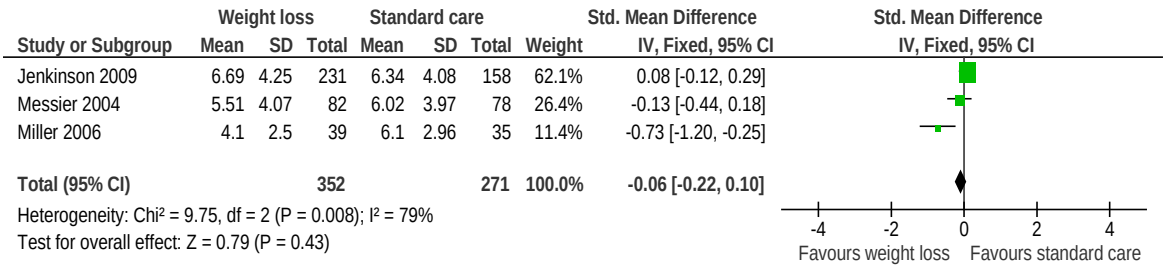


Figure 16: Pain (at end of treatment): Weight loss (dieting) versus standard care.

The comparison of weight loss with aerobic exercise from a single trial[82] found no significant difference in pain relief (SMD 0.18, 95% CI -0.13 to 0.48).

Disability (WOMAC index)

One poor quality study provided overall WOMAC scores at the end of treatment suitable for meta-analysis.[94] The result indicated a significant benefit favouring weight loss over standard care (SMD -0.78, 95% CI -1.25 to -0.30).

Adverse effects

The only study to assess adverse events reported no events occurring for the weight loss group.[82]

5.2.5.4 *Summary of effectiveness of weight loss*

There was no evidence of satisfactory quality to suggest weight loss was more effective than standard care in reducing knee OA pain.

Table 8: Weight loss trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Mean Age (years)	Weight (in kg: mean, unless stated)	BMI (kg/m^2, Mean, unless stated)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of dieting period (weeks)	Difference in weight loss between treatment groups (kg)	Weight loss cut of used to analyses effect on pain or WOMAC, e.g. 10% of body weight	Pain outcome	WOMAC score
Christensen 2005 [92]	Weight loss Standard care	78 N**	Denmark	89	63	97	36	Clinical and radiological	2 or 3	Poor	8	6.6 (95%CI 5.3 to 7.9, p<0.0001) Percentage of patients achieving 10% body weight loss was 50% in Weight Loss group and 0% in Standard Care group	Cut off used 10%	WOMAC pain subscale VAS 0-100	Overall WOMAC score reported
Jenkinson 2009 [93]	Weight loss (2 & 3) Standard care(3)	389 Y	UK	66	61	Median 93	Median 33	Clinical	4 or lower 41% had a K&L score of 0	Satisfactory	104	NR/unclear	Cut off used Dietary interventions aimed for a weight loss of 0.5-1.0kg per week. Weight loss not reported.	WOMAC pain Likert 5	No
Messier 2004 [82]	Weight loss(2) Aerobic exercise(2) Standard care(3)	240 Y	USA	72	69	94	34	Clinical and radiological	3 or lower	Satisfactory	78	Mean weight loss by group: Weight loss group 4.61kg, exercise group 3.46 kg, standard care group 1.1kg.	Cut off used. Average weight loss goal was 5%	WOMAC pain Likert 5	Individual WOMAC subs scores reported Pain, physical function
Miller 2006 [94]	Weight loss(3) Standard care(3)	74 Y	USA	26	70 (all >/= 60)	98	35	Clinical	NR/ Unclear	Poor	26	8 kg	Cut off used 10% weight loss was goal for intervention group.	WOMAC pain Likert 5	Overall WOMAC score reported WOMAC sub scores Pain, stiffness and function

** Only change from baseline scores reported

5.2.6 **Balneotherapy**

5.2.6.1 *Study characteristics*

Fourteen trials studied balneotherapy interventions (see Table 9) with a total of 1,008 participant pain scores analysed at the end of treatment (range 20 to 309); all were full published papers reported in English, between 1995 and 2010. Five of the studies were conducted in Israel.

All studies recruited a general population with the exception of two that studied patients with both knees affected.[95 96] The mean ages of participants ranged from 54 to 70 years, and the proportion of females ranged from 47 to 100%. Mean BMI (26 to 32 kg/m^2) was reported in three studies and in another 61% of the study had a BMI > 24. Mean weight (71 to 77 kg) was reported in three studies.
The methods used for diagnosis were clinical and radiological in 10 studies and clinical in four studies. Seven studies reported using Kellgren & Lawrence scores (patients with a broad range of scores were recruited) and one the Lequesne index of severity.[97]

One study was a four armed trial which compared three groups for balneotherapy (Dead Sea; Sulphur pools; and a combination of both) with placebo.[97] One study was a three-armed trial which compared balneotherapy with either heat treatment or standard care.[98] The remaining studies were two armed trials comparing balneotherapy with either placebo or standard care. One study included regular exercise for both groups[95] and another a home exercise programme.[99] Sessions lasted between 20 and 65 minutes. The total number of sessions varied from 6 to 30, over periods ranging from 10 days to 6 weeks.

Pain was measured using a variety of scales; two studies measured WOMAC pain (using a Likert scale) and five used the WOMAC pain subscales. Two studies reported overall WOMAC scores, four reported individual WOMAC sub-scores and two Lequesne Index scores. Six studies reported quality of life outcomes and five reported adverse effects.

5.2.6.2 *Study Quality*

The number of participants randomised was clearly stated in all but two studies,[100 101] but six clearly reported using appropriate randomisation methods,[96 99-103] and only four studies[95 99 102 104] reported suitable methods for concealing treatment allocation.

Eligibility criteria were adequately described in 11 trials and group baseline characteristics appeared comparable in eight and nine studies reported blinding outcome assessors. Only two studies reported using a power calculation for sample sizes[99 105] and five studies clearly reported using data for the intention-to-treat population.[96 98 102-104] However, all studies bar three[104 106 107] reported whether there were any losses to follow up. Eight studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived 12 studies were rated as being of poor quality and two of satisfactory quality[96 102]. Full details of study quality are reported in Appendix 10.2.

5.2.6.3 *Results of effectiveness*

Pain

The six trials that did not provide final value data for analysis[98 100 102 103 105 108] were generally similar to those that did, except most were comparisons with standard care and included the only direct comparison of balneotherapy with heat treatment.[98]. They reported positive effects for balneotherapy compared with standard care or placebo.

Seven poor quality studies provided data suitable for meta-analysis of balneotherapy versus placebo (Figure 17). The pooled result indicates that balneotherapy was associated with a non-significant reduction in end of treatment pain compared with placebo, but there was significant heterogeneity between the studies (I^2=69%). The one study that significantly favoured balneotherapy over placebo[95] included regular exercise for both treatments arms.

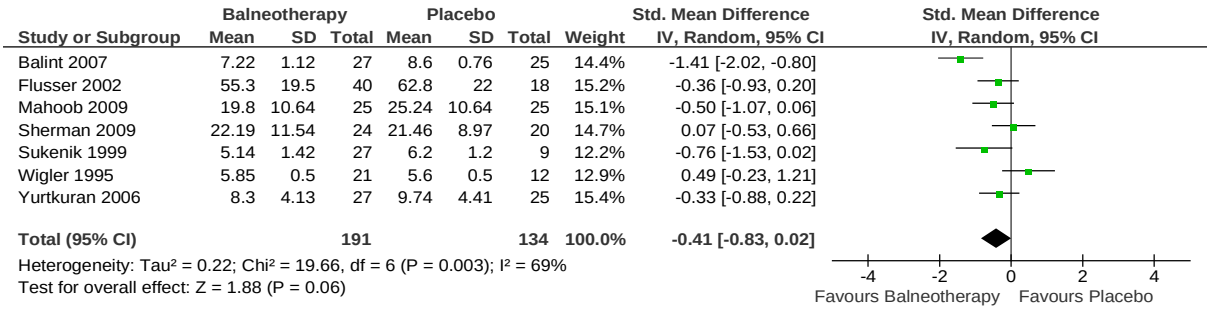


Figure 17: Pain (at end of treatment): balneotherapy versus placebo (all studies).

All studies ineligible for meta-analysis failed to present end of treatment means (see table), and reported positive effects for balneotherapy compared with standard care or placebo.

One study comparing balneotherapy with standard care, provided data suitable for meta-analysis, with balneotherapy appearing to offer more benefit (SMD -1.01, 95% CI -1.48 to -0.54).[96]

Disability (WOMAC index)

Two studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis. [95 99] The results were heterogeneous; in one study placebo appeared more effective (SMD 2.25, 95% CI 1.54 to 2.95)[95] and in the other balneotherapy was more effective (SMD -0.58, 95% CI -1.14 to -0.02).

Adverse effects

Of the eight studies that assessed adverse effects, three stated that no adverse effects were reported, with the remainder reporting limited specific details or increased pain, itching or increased dieresis.

5.2.6.4 *Summary of effectiveness of balneotherapy*

The result of the meta-analysis, which included only poor quality trials showed a non-significant benefit of balneotherapy. Both satisfactory quality trials found a significant benefit of balneotherapy: one compared with placebo and one with standard care.

Table 9: Balneotherapy trials: Study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Temperature	Type of treatment	Pain outcomes used	Overall WOMAC score reported?
Balint 2007 [95]	Bal(3) Placebo (3)	52 Y	Hungary	63	Range 50 to 75	NR	NR	Clinical	NR	Poor	30	20	4	36 degrees C	Bathing mineral water	WOMAC Pain subscale	Yes
Cantarini 2007 [98]	Bal Heat treatment, standard care	74 N*	Italy	63	64	*71*	NR	Clinical and radiological	3 or lower	Poor	35	15	3	Water 38 degrees C, mud packs 45 degrees C	Bathing mineral water and mud packs	Pain VAS 0-100.	No
Fioravanti 2010 [96]	Bal(2) Standard care(2)	80 Y	Italy	75	70	NR	*26*	Clinical and radiological	3 or lower	Satisfactory	35	12	2	Mud packs (45 degrees) and mineral bath (38 degrees).	Bathing mineral water and mud packs	WOMAC pain subscale VAS 0-100	No
Flusser 2002 [107]	Bal(2) Placebo (2)	58 Y	Israel	85	65	*76*	NR	Clinical and radiological	2 or 3	Poor	20	15	3	30 to 35 degrees C	Mud packs	Pain VAS 0-10	Lequesne Index
Forestier 2010 [105]	Bal Standard care	309 N*	France	47	64	NR	*30*	Clinical and radiological	1 or higher	Poor	65	18	3	Mineral hydrojet (37 degrees), massages by physiotherapist (38 degrees), mineral matured mud (45 degrees), and collective mineral water pool (32 degrees).	Bathing mineral water and mud packs	WOMAC pain subscale VAS 0-100 Other Pain VAS 0-100	WOMAC Function
Karagulle 2007 [102]	Bal Standard care	20 N*	Turkey	85	60	NR	NR	Clinical and radiological	2 or higher	Satisfactory	30	20	10 days	38+-1 degrees Celsius	Bathing mineral water	Pain VAS 0-10	Lequesne Algofunctional Index
Kovacs 2002 [100]	Bal Placebo	68 N*	Hungary	58 to 78*	NR	NR	NR	Clinical and radiological	Not reported/ Unclear	Poor	30	15	15 days	36 degrees Celsius	Bathing mineral water	Pain VAS	No
Mahoob 2009 [104]	Bal(2) Placebo (2)	50 Y	Iran	100	Range 44 to 79	NR	NR	Clinical	NR	Poor	20	30	30 days	NR	50 g of mud gel	WOMAC pain subscale VAS	Individual WOMAC subs scores for pain, function and stiffness

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Temperature	Type of treatment	Pain outcomes used	Overall WOMAC score reported?
Nguyen 1997 [103]	Bal Standard care	64 N*	France	81	NR	NR	*61% had BMI > 24*	Clinical	NR	Poor	Un-clear	Un-clear	3	NR	Bathing mineral water	Pain VAS 0-100	No
Sherman 2009 [101]	Bal(2) Placebo (2)	44 Y	Israel	80	67	NR	NR	Clinical and radiological	1 or higher	Poor	20	12	6	35-36 degrees	Bathing mineral water	WOMAC pain subscale VAS 0-10 Pain VAS 0-100	Individual WOMAC subs scores for pain, function and stiffness
Sukenik 1999 [97]	Bal(2) Placebo(2)	36 Y	Israel	89	63	NR	NR	Clinical and radiological	Lequesne index of severity	Poor	20	14	2	Dead sea Sulphur pools 37 degrees; Sweet water control 24-25 deg	Bathing mineral water	VAS 0-10	No
Tishler 2004 [108]	Bal Standard care	68 N*	Israel	78	64	NR	NR	Clinical	NR	Poor	30	6	6	37 degrees	Bathing mineral water	WOMAC pain Likert 5 Other Pain VAS (specify) VAS 0-100	WOMAC subs scores reported for pain, function and stiffness
Wigler 1995 [106]	Bal(2) Placebo(2)	33 Y	Israel	88	Mean 65	NR	NR	Clinical and radiological	NR	Poor	40	7	2	Water 38 degrees, mud pack 45 degrees	Bathing mineral water and mud packs	Pain VAS 0-10	No
Yurtkuran 2006 [99]	Bal(5) Placebo (5)	52 Y	Turkey	97	54	*77*	*32*	Clinical and radiological	2 or 3	Poor	20	10	2	37 degrees	Bathing mineral water	WOMAC pain Likert 5 Pain VAS 0-100	Yes

* No Means Bal: Balneotherapy

5.2.7 **Insoles**

5.2.7.1 *Study characteristics*

Six trials studied insoles (see Table 10) with a total of 669 participant pain scores analysed at the end of treatment (range 30 to 172). All were full papers published in English between 2001 and 2009. Two studies were conducted in the United States.

Four studies recruited a general knee OA population and two[109 110] recruited only patients with knee malalignment. The mean ages of participants ranged from 58 to 68 years, and the proportion of females ranged from 54 to 100%. Mean BMIs ranged from 25 to 33 kg/m^2 (mean weight was only reported in one study).

The methods used for diagnosis were clinical and radiological in four studies, radiological alone in one study, and clinical or radiological in one study. All studies reported using Kellgren & Lawrence scores, with four recruiting patients with scores of 2 or more, one 1 or more,[111] and one 4 or lower.[112]

All trials compared insoles with placebo. Where stated, the time participants spent wearing insoles ranged from three hours per day, to all day; and most studies stated insoles should be worn every day of the study duration (which ranged from 6 weeks to two years). Two trials used ankle support as well as insoles.[110 112]

Usual concomitant treatments or trial analgesics, as required, were allowed in five studies (details were unclear in one study) and one study[113] also allowed use of analgesic or corticosteroid injections.

Pain was measured using a WOMAC VAS (0-100 scale) in four studies, with other types of VAS pain scales used in the remaining two trials. Two studies reported overall WOMAC data with the remaining two studies not reporting an overall assessment of disability (although one reported Individual WOMAC subs scores). Two studies assessed adverse effects.[110 111]

5.2.7.2 *Study Quality*

One study[111] had a crossover design. The number of participants randomised was clearly stated in all studies, with four clearly reporting use of appropriate randomisation procedures, two[109 112] of which also used appropriate allocation concealment methods. All studies had adequately described eligibility criteria, and comparable group baseline characteristics. One study[111] was reported as being double-blind, and two[110 112] reported blinding outcome assessors.

Although three studies reported using a power calculation for sample sizes, five studies reported using data for the intention-to-treat population. All studies reported whether there were any losses to follow up, and five studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived one study was rated as good quality, four were rated as satisfactory, and one rated as poor. Full details of study quality are reported in Appendix 10.2.

5.2.7.3 *Results of effectiveness*

Pain

Three trials did not provide final value mean data for analysis.[111 112 114] They comprised three of the four biggest trials of insoles but otherwise were generally similar to the trials that could be included in the analysis.[109 110 113]

The result from pooling the three studies (Figure 18) (all of satisfactory or good quality) comparing insoles with placebo was subject to significant heterogeneity, which appears to be due to one very small study that also included ankle supports as part of the intervention[110]: this study showed a significant effect of insoles in reducing pain (SMD -0.84 (95% CI -1.59, -0.09). Removal of this study

resulted in the I^2 value falling to 0% and the pooled estimate of effect was an SMD of 0.24 (95% CI -0.02 to 0.51), i.e. no effect of insoles.

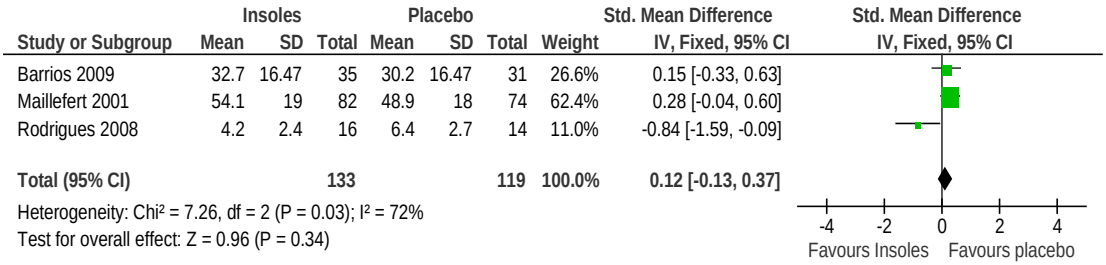


Figure 18: Pain (at end of treatment): insoles versus placebo.

Three studies could not be included in the meta-analysis: one did not report means, and two reported only differences from baseline, with none reporting that insoles had significant benefit when compared directly to placebo, or standard care.

Disability (WOMAC index)

One good quality study provided overall WOMAC scores at the end of treatment suitable for meta-analysis.[110] No significant difference was found between insoles (with ankle supports) and placebo (SMD -0.45, 95% CI -1.18 to 0.28).

Adverse effects

Of the two studies assessing adverse effects, one reported that none occurred, and one reported occasional blistering of the toes.

5.2.7.4 *Summary of effectiveness of insoles*

Evidence of satisfactory quality suggested that the use of insoles did not significantly reduce knee OA pain. However, one good quality trial of insoles with ankle support indicated a significant benefit.

Table 10: Insoles trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Mean Age (in years)	Mean Weight (in kg)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Time spent wearing (per day)	Number of days worn per week	Duration of treatment period	Was an ankle support also worn?	Pain outcome	Overall WOMAC score reported?
Baker 2007 [111]	Insoles Placebo	172 N*	USA	59	68	NR	33	Clinical or radiological	1 or higher	Poor	approximately 420 minutes (7 hours)	Unclear/not stated	6 weeks phase 1: 6 weeks, phase 2: 6 weeks and wash out period: 4 weeks.	No	WOMAC pain subscale VAS 0-100	No
Barrios 2009 [109]	Insoles(4) Placebo(4)	66 Y	USA	56	62	NR	33	Clinical and radiological	2 or higher	Satisfactory	Full day (wear time was gradually increased over 3-4 days).	7	1 year	No	WOMAC pain subscale VAS 0-100	All 3 individual WOMAC subs scores reported
Maillefert 2001 [113]	Insoles(4) Placebo(4)	156 Y	France	74	65	NR	29	Clinical and radiological	2 or higher	Satisfactory	Unclear/not stated wear permanently	Unclear/not stated wear permanently	Unclear/not stated up to 24 months	No	WOMAC pain subscale VAS 0-100	No
Nigg 2006 [114]	Insoles Placebo	123 N**	Canada	54	58	85	30	Clinical and radiological	2 or higher	Satisfactory	As much as possible	7	12 weeks	No	WOMAC pain subscale VAS 0-100	Yes
Rodrigues 2008 [110]	Insoles(2) Placebo(2)	30 Y	Brazil	100	62	NR	30	Clinical and radiological	2 or higher	Good	180 to 360 minutes	7	8 weeks	Yes	Other Pain VAS 0-10 night, rest and move	Yes
Toda 2008 [112]	Insoles Placebo	122 N**	Japan	88	64	NR	25	Radiological	4 or lower	Satisfactory	300 to 600 minutes	7	12 weeks	Sock-type ankle support	Other Pain VAS 0-100%	No

* No Means, ** Only change from baseline scores reported

5.2.8 **Static magnets**

5.2.8.1 *Study characteristics*

Three trials assessed static magnets (see Table 11) with a total of 131 participant pain scores analysed at the end of treatment; all were full published papers in English between 2002 and 2008. Two of the studies were conducted in the United States and the other in Taiwan.

All participants were derived from a general knee OA population and their mean age ranged from 63 to 65 years, and the proportion of females ranged from 60 to 79%. The mean BMI was 27 in one study and the median 30 in a second study; mean weight was reported in one study (64 kg). The methods used for diagnosis were clinical in one study and clinical and radiological in two studies. Kellgren & Lawrence scores were reported in one study (a score of at least one was required) and another study used the Ahlbäck classification (grade I).

All studies compared static magnets with placebo. The strength of magnetic field (Gauss or Tesla) was 35mT (as measured with a Lakeshore 430 gauss meter), [115] 1.08 T, [116] or 40-850 Gauss.[117] In one study[115] magnets were worn during waking hours and in another study they were worn for at least 6 hours a day.[117] In the third study[116] the duration was not specified but was noted by the patients. The duration of treatment periods ranged from 2 weeks to 12 weeks. Usual (or trial-specific) concomitant treatments, as required, were allowed in two studies. One study required at least one of the study arms to take specified doses of analgesics.[117]

Pain was measured using the HAQ 0-100 pain scale in one study, [115] WOMAC VAS 0-10 pain subscale in one study[116] and both the WOMAC VAS 0-100 pain subscale and another Pain VAS 0-100 in the final study.[117] Quality of life data were reported in one study which used HAQ.[115] Adverse effects were assessed in two studies.[115 117]

5.2.8.2 *Study Quality*

Two studies[115 117] clearly stated the number of participants randomised and used appropriate methods for randomisation, but this was not clear or not stated in the other study[118]. All studies reported using appropriate methods for concealing treatment allocation and use of an appropriate placebo. Eligibility criteria were adequately described in all trials. Group baseline characteristics appeared comparable in one study[115] but were unclear[116] or not comparable [117] in the other studies. One study reported use of a power calculation for sample sizes[115], one study reported data for the intention-to-treat population[117] and all studies reported losses to follow up. Two studies achieved full follow up for at least 90% of participants.[116 117]

When the overall study quality ratings were derived all three studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.8.3 *Results of effectiveness*

Pain

A single poor quality study provided final values mean data for analysis.[116] This trial was poorly reported and the duration of treatment was only 2 weeks, compared with 6 and 12 weeks in the other two trials. Its findings indicated a beneficial effect of static magnets over placebo (SMD -0.82, 95% CI -1.46 to -0.19).The other trials results were less favourable: in one study pain scales improved significantly in both groups and in the other efficacy did not significantly differ between the groups.

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

Two studies assessed adverse effects, with skin irritation and muscle soreness reported in one study and some patients reported mild discomfort from wearing the cotton/elastic knee sleeve in another study.

5.2.8.4 *Summary of effectiveness of static magnets*

There was no evidence of satisfactory quality to suggest that static magnets significantly reduced knee OA pain.

Table 11: Static magnet trials: Study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Strength of magnetic field (Gauss or Tesla)	When to wear magnets	Duration of treatment period (weeks)	Pain outcomes used	Overall WOMAC score reported?
Chen 2008 [115]	Static magnets Placebo	42 N*	Taiwan	79	65	64	27	Clinical and Radiological	Ahlbäck classification (grade I)	Poor	35mT, as measured with a Lakeshore 430 gauss meter	Only during waking hours	12	HAQ pain scale 0-100	No
Hinman 2002 [116]	Static magnets Placebo	43 Y(only one study)	USA	60	63	NR	NR	Clinical	NR	Poor	1.08 T	Not specified***	2	WOMAC pain subscale VAS 0-10	No
Wolsko 2004 [117]	Static magnets Placebo	46 N**	USA	69	63+	MR	Median 30+	Clinical and Radiological	1 or higher	Poor	40-850 Gauss	At least 6 hours a day	6	WOMAC pain subscale VAS 0-100 and VAS 0-100 with 5 scales (max 500)	No

* No Means, ** Only change from baseline scores reported.

*** Magnet group: 1.0 to 23.5 hours per day; mean number of hours a magnet was worn 116.33 (range 25-235)
Placebo group: 1 to 24 hours per day and the mean number of hours a magnet was worn 85.12 (range 6.5-213)

5.2.9 **Braces**

5.2.9.1 *Study characteristics*

Two trials assessed braces with a total of 227 participant pain scores (see Table 12); both were full published papers in English published in 1999 and 2006; studies were conducted in the Netherlands and Canada.

All participants had varus or valgus/malalignment; this patient group differed from all other trials for other interventions. Mean age ranged from 50 to 59 years, and the proportion of females ranged from 28 to 83%. Mean BMI was not reported in either study, but in one study[119] only patients with a BMI<35 were included and in the other study the mean weight was 59kg.[120] The methods used for diagnosis were clinical and radiological in both studies. The Kellgren & Lawrence score was 2 or higher in one study[119] whilst the other study used an Ahlbäck classification score of >0.[120] Both studies were two-armed trials comparing braces with standard care.

In one study[119] braces were worn whilst awake or during troublesome activities for 7 days a week. The duration of the treatment period ranged from 24 weeks to 12 months.

Pain was measured using a VAS 0-10 in one study[120] and through a WOMAC pain subscale VAS 0-100 and Pain VAS 0-100 for 6-minute walking test in the other.[119] Quality of life data were reported in both studies and comprised EQ5D[120] and the McMaster-Toronto Arthritis Patient Preference Disability Questionnaire.[119] Adverse effects were assessed in one study.[120]

5.2.9.2 *Study Quality*

Both studies clearly stated the number of participants randomised, used appropriate methods for randomisation and concealing treatment allocation. Blinding was not applicable for these studies. Eligibility criteria were adequately described in both trials. Group baseline characteristics were not comparable in one study[120] or were unclear in the other study.[119].Both studies reported use of a power calculation for sample sizes, one study reported data for the intention-to-treat population[120] and both studies reported losses to follow up. One study achieved full follow up for at least 90% of participants.[119]
When the overall study quality ratings were derived both studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.9.3 *Results of effectiveness*

Pain

One of the two studies provided data suitable for analysis, but this study only provided data for end of treatment differences between treatment groups, so could only be included in the NMA analyses.[120] One of the studies reported a significant difference in pain compared with standard care and the other a borderline significant improvement in pain severity.

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

One study assessed adverse effects, with skin irritation and ‘bad fit’ reported for those using braces.

5.2.9.4 *Summary of effectiveness of braces*

There was no evidence of satisfactory quality to suggest that braces significantly reduced knee OA pain.

Table 12: Braces trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Time spent wearing (per day)	Number of days worn per week	Duration of treatment period	Concomitant treatment	Pain outcomes used	Overall WOMAC score reported?
Brouwer 2006 [120]	Braces(2) Standard care(3)	117 Y	Netherlands	50	59	NR	29	Clinical and radiological	Ahlbäck score >0	Poor	NR	NR	12 months		Other Pain VAS 0-10	No
Kirkley (1999) [119]	Braces Standard care	110 N**	Canada	28	59	NR	BMI<35 were in-cluded	Clinical and radiological	2 or higher	Poor	Braces worn during troublesome activities	7 days	Assumed to be 24 weeks		WOMAC pain subscale VAS 0-100 VAS 0-100 for 6-minute walking test	Yes

** Only change from baseline scores reported

5.2.10 **TENS**

5.2.10.1 *Study characteristics*

Seventeen trials studied TENS interventions (see Table 13) with a total of at least 730 participant pain scores analysed at the end of treatment (range 12 to 116, sample sizes were sometimes not clearly stated); all but one were full published papers reported in English, between 1981 and 2009; there was one conference abstract.[121] Five of the studies were conducted in the USA with three each in Australia and China.

All studies recruited a general population with the exception of one that used patients awaiting knee surgery.[122] The mean/median ages of participants ranged from 56 to 85 years, and the proportion of females ranged from 48 to 97%. Mean BMI (26 to 31 kg/m^2) was reported in five studies as was mean weight (57 to 88 kg).
The methods used for diagnosis were clinical and radiological in 12 studies, clinical in two studies, radiological in one study and not reported in two. Five studies reported using Kellgren & Lawrence scores, with four recruiting patients with scores of at least two.

Three studies were four armed trials: one compared TENS at either 2 Hz, 100 Hz, an alternating frequency of 2 Hz and 100 Hz with placebo,[123] another compared three different durations of TENS with placebo (20, 40 and 60 minutes),[124] and another compared TENS with acupuncture, ice/cooling treatment and placebo.[37] Five studies were three-armed trials and nine studies two-armed trials.
Sessions lasted between 15 and 60 minutes. The total number of sessions varied from 1 to 63, over periods ranging from one session to 9 weeks.
Usual (or trial-specific) concomitant treatments, as required, were allowed in eight studies and no details were provided in eight studies. No medication was allowed in two studies. [125 126]

Pain was measured using a variety of scales but only one measured WOMAC pain.[127] Two studies reported overall WOMAC score[35 128] and two reported individual WOMAC sub-scores.[127 129] Two studies reported quality of life outcomes[127 128] and five assessed adverse effects.[37 127-130]

5.2.10.2 *Study Quality*

The number of participants randomised was clearly stated in all but one study,[123] but seven clearly reported using appropriate randomisation methods, and five studies reported suitable methods for concealing treatment allocation. Eligibility criteria were adequately described in 13 trials and group baseline characteristics appeared comparable in seven. Five studies were reported as being double blind and 10 reported blinding outcome assessors. Only four studies used a power calculation for sample sizes[127-129 131] and two studies clearly reported using data for the intention-to-treat population.[37 126] However, all studies bar three[36 121 132] reported whether there were any losses to follow up. Nine studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived all studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.10.3 *Results of effectiveness*

Pain

Of the 17 TENS trials only 10 reported final value mean data suitable for analysis. The seven trials that could not be included in the standard or NMA analyses [36 121-123 127 130 133] included a comparison with inferential therapy that included Kellgren & Lawrence grade 1 OA[127]. They also included the three trials that administered a high number of sessions to patients (56[127] and 63).[131 133] One trial reported only differences between treatment groups, so was suitable for the NMA analyses, but not the meta-analyses.[131]

The nine trials that did provide final value mean data included six comparisons with placebo[37 74 124 125 128 129], three comparisons with acupuncture,[36 37 134] two comparisons with ice treatment[37 126], one comparison with heat treatment[132], and one comparison with standard care.[35]

When all six (poor quality) trials that compared TENS with placebo were pooled TENS was associated a significant reduction in end of treatment pain compared with placebo (figure 19), but this was subject to significant heterogeneity (I^2=77%). When the two trials in which patients had received only a single session of TENS were removed[125 129] heterogeneity increased to 86% and the pooled SMD increased to -0.93 (95% CI: -1.88, 0.01). However, when one of these studies, whose estimate of effect was much more favourable than the other trials, and in which the baseline groups were not comparable.[124] was removed as a sensitivity analysis, the statistical heterogeneity fell to I^2=0%, with the treatment effect being reduced to -0.55 but remaining statistically significant (CI -0.82, -0.29) (Figure 20).

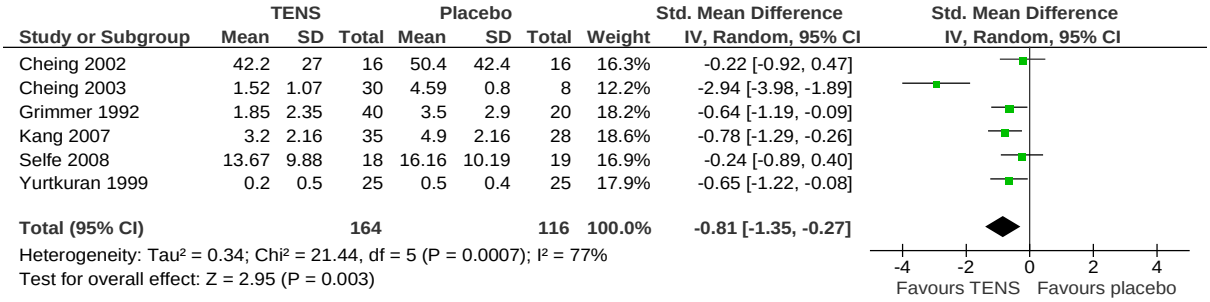


Figure 19: Pain (at end of treatment): TENS versus placebo (all studies).

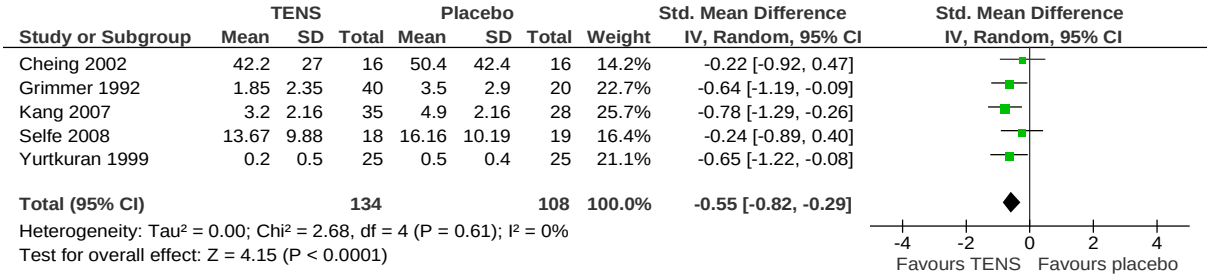


Figure 20: Pain (at end of treatment): TENS versus placebo (sensitivity analysis.

When the two (poor quality) trials that compared TENS with ice treatment were pooled TENS was associated with a small, non-significant reduction in end of treatment pain compared with ice treatment (Figure 21). There was no statistical heterogeneity despite one of the trials[126] having tested only a single session of treatment.

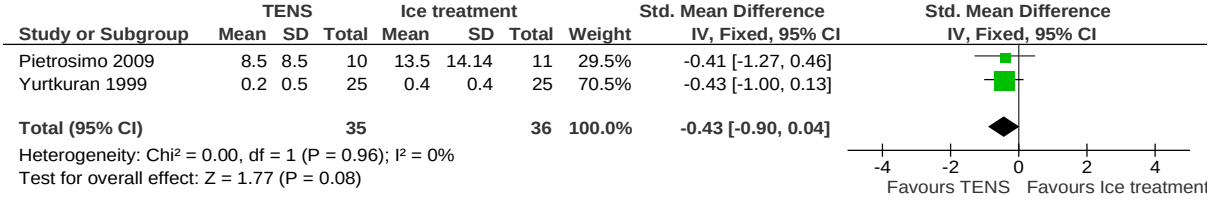


Figure 21: Pain (at end of treatment): TENS versus ice treatment (all studies).

No significant differences between acupuncture and TENS were found for end of treatment pain (three poor quality studies, see Figure 6). A single small study compared TENS with standard care and found that TENS did not appear to offer any improved benefit (SMD -1.29, 95% CI -2.58 to 0.01).[35] Of those studies ineligible for meta-analysis, two presented only change from baseline scores and six did not present end of treatment means (see Table 13). All studies reported results favouring TENS.

Disability (WOMAC index)

One very small study compared TENS with standard care, but did not find a significant difference (SMD -1.08, 95% CI -2.32 to 0.17)[35] and another compared TENS with placebo and also did not find a significant difference (SMD
-0.20, 95% CI -0.84 to 0.45).[128]

Adverse effects

Of the five studies that assessed adverse effects, three stated that no adverse effects were reported, whilst in one study a patient developed a mild skin reaction to the electrode jelly [130] and in another a patient reported muscle soreness deemed probably to be related to the device. [127]

5.2.10.4 *Summary of effectiveness of TENS*

There was no evidence of satisfactory quality to suggest that use of TENS significantly reduced knee OA pain. Amongst poor quality studies TENS was associated a significant reduction in end of treatment pain compared with placebo.

Table 13: TENS trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Frequency of current (Hz)	Intensity of current (mA)	Pulse width/duration (micro secs)	Electrodes placed at acupuncture points?	Pain outcomes used	Overall WOMAC score reported?
Alcidi 2007 [132]	TENS(2) Heat treatment(2)	40 Y	Italy	85	66	NR	NR	Clinical and radiological	NR	Poor	20	5	1	50	'Well tolerated tingling'	0.5	No	Other Pain VAS 1-100	No
Burch 2008 [127]	TENS Interferential therapy	116 N**	USA	72	62	86	31	Clinical and radiological	1 or higher	Poor	35	56	8	0.2	60	300	Unclear/ not stated	WOMAC pain Likert 5	All 3 individual WOMAC subs scores reported
Cheing 2002 [74]	TENS(2) MSE(2) Placebo(2)	47 Y	China	89	63	67	28	Clinical and radiological	2 or higher	Poor	60	20	4	80	To achieve a tingling sensation	140	Yes	Other Pain VAS 0-100	No
Cheing 2003 [124]	TENS(2) Placebo(2)	38 Y	China	89	66	66	NR	Clinical and radiological	2 or higher	Poor	20 to 60	10	2	100	NR	200	Yes	Other Pain VAS 0-10	No
Grimmer 1992 [125]	TENS(1) Placebo(1)	60 Y	Australia	62	67	NR	NR	Clinical and radiological	NR	Poor	30	1	1 session	80	NR	NR	Yes	Other Pain VAS	No
Itoh 2008 [35]	TENS(1) Acupuncture(1) standard care(2)	12 Y (only one study)	Japan	66	62-83	NR	NR	Clinical and radiological	2 or higher	Poor	15	5	5	4	NR	NR	No	Other Pain VAS 0-100	Yes
Kang 2007 [129]	TENS(2) Placebo(2)	63 Y	USA	71	57	NR	NR	Clinical and radiological	NR	Poor	30	1	1 session	NR	16% at start, increased to 23% at 15 minutes	NR	No	Other Pain VAS 0-10	All 3 individual WOMAC subs scores reported
Law 2004 [123]	TENS Placebo	34 N*	Hong Kong	97	83	57	26	Clinical and radiological	2 or higher	Poor	40	10	2	2, 100, or alternating 2/100	25-35mA	2Hz: 576 100Hz:200	Yes	Other Pain VAS 0-100	No

Lewis 1988 [121]	TENS Placebo	unknown N*	Australia	NR	NR	NR	NR	Unclear/not stated	NR	Poor	NR	NR	9	Unclear/ not stated	Unclear/ not stated	Unclear/ not stated	Unclear/ not stated	McGill Pain Questionnaire, Pain Index	No
Lewis 1994 [131]	TENS(1) Standard care(4)	56 Y	Australia	58	66	NR	NR	Clinical and radiological	NR	Poor	30 to 60	63	3	70	Adjusted until comfortable	100 micro secs	Spleen 9 and 10, stomach 34 and 35	VAS 0-100	No
Lewis 1984 [133]	TENS Placebo	30 N*	UK	73	Median 61	NR	NR	Unclear/not stated	NR	Poor	30-60	63	7	70	NR	NR	Yes	VAS	No
Ng 2003 [36]	TENS Electro-acupuncture Standard care	14 N*	China	96	85	NR	NR	Clinical	NR	Poor	20	8	2	2	NR	200	Yes	Numerical Rating Scale	No
Pietrosimone 2009 [126]	TENS(1) Ice/cooling treatment(1) No intervention)1)	33 Y	USA	48	56	88	30	Radiological	NR	Poor	45	1	1 day	150	NR	150 micro secs	No	VAS	No
Selfe 2008 [128]	TENS(2) Placebo(2)	37 Y	USA	68	67	NR	31	Clinical and radiological	NR	Poor	20 to 30	17	8	NR	NR	NR	Some-times	WOMAC pain VAS 0-10 Numeric Rating Scale	Yes
Smith 1983 [130]	TENS Placebo	30 N*	UK	67	68	NR	NR	Clinical	NR	Poor	20	8	4	32 to 50	Adjusted until comfortable	80	Yes	Other Significant pain relief***	No
Taylor 1981 [122]	TENS Placebo	20 N**	USA	90	72	NR	NR	Clinical and radiological	NR	Poor	30	NR	2	NR	NR	NR	No	Likert 5	No
Yurtkuran 1999 [37]	TENS(2) Acupuncture(2) ice/cooling(2) Placebo(2)	100 Y	Turkey	91	58	NR	NR	Clinical and radiological	NR	Poor	20	10	2	4	0.4-2.5 volts	1000 ms	Yes	Likert 5	No

* No Means, ** Only change from baseline scores reported. ***Defined as any of the following two criteria being fulfilled (and no worsening in the remaining criterion): 1) a ten point or 50% decrease in weekly pain score, 2) a 50% decrease in analgesic intake compared to baseline and 3) a 5-point improvement in the weekly sleep disturbance score. MSE Muscle-strengthening exercise

5.2.11 **Pulsed electrical stimulation (PES)**

5.2.11.1 *Study characteristics*

Six trials studied PES (see Table 14), with a total of over 210 participant pain scores analysed at the end of treatment (range 18 to 71, sample sizes were sometimes not clearly stated); all were full published papers with five published in English and one in Spanish[32], predominantly between 2005 and 2008. Two studies were conducted in the United States.

All studies recruited a general population, with mean ages of participants (where stated) ranging from 55 to 66 years, and the proportion of females ranging from 46 to 100%. Mean BMI was reported in only three studies (range 28 to 33 kg/m^2), and mean weight in only one study. The methods used for diagnosis were clinical and radiological in five studies, and radiological alone in one study. Three studies were of patients with Kellgren & Lawrence scores of three or four, and one studied patients with a score of ≤3. One study used Gupta criteria (grade II or III), and the remaining study did not report details on classification of severity.

Five trials compared PES with placebo (one trial[135] also had a standard care group, and one trial[135] also compared different doses of PES), and one trial compared PES with muscle-strengthening exercise. Treatment was generally given for around 20 minutes, although two studies[136 137] treated patients with sessions lasting six or more hours. The number of sessions used ranged from six to 39, over periods ranging from two to 13 weeks. Current frequencies ranged from 35Hz to 27MHz. Usual (or trial-specific) concomitant treatments, as required, were allowed in four studies[135-138], and no details were provided the two studies.[32 76]

Pain was measured using a VAS 0-10 scale in three studies, a WOMAC VAS 0-100 scale in two studies and a WOMAC VAS 0-10 point scale in one study; one study also measured pain using AIMS. Only one study reported overall WOMAC scores, one reported individual WOMAC sub-scores, one reported Lequesne scores and the remaining studies did not report on overall assessment of disability. Adverse effects were monitored in only two studies[136 137] and no studies reported quality of life data.

5.2.11.2 *Study Quality*

The number of participants randomised was clearly stated in all studies, but only one[137] clearly reported using appropriate randomisation, and only two studies[135 137] reported suitable methods for concealing treatment allocation.
Eligibility criteria were adequately described in five trials, and group baseline characteristics appeared comparable in four. Three studies clearly reported using an appropriate placebo treatment.[32 137 138] Although three studies[136-138] were described as being double-blind only two[32 137] clearly reported blinding participants, one[137] reported blinding treatment-givers and three[135 137 138] reported blinding outcome assessors. Only one study[138] reported use of a power calculation for sample sizes, and only one study[76] reported data for the intention-to-treat population. However, all studies bar one[32] reported whether there were any losses to follow up. Only half the studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived five studies were rated as being of poor quality, and one[76] was rated as being of satisfactory quality. Full details of study quality are reported in Appendix 10.2.

5.2.11.3 *Results of effectiveness*

Pain

Of the five trials comparing PES with placebo, three small, poor quality studies provided final value mean data for analysis. Generally these trials were not different from the other trials except that the number of sessions in one trial was low (6).[138] The pooled result indicating a small non-significant

effect of PES over placebo at the end of treatment was not subject to statistical heterogeneity (figure 22).

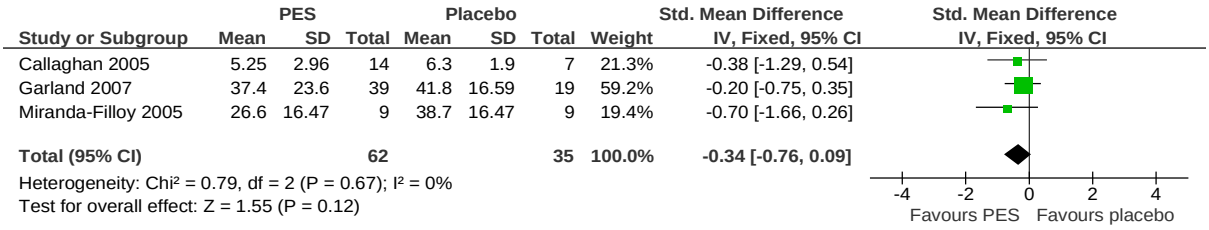


Figure 22: Pain (at end of treatment): PES versus placebo.

In one study the comparator was muscle-strengthening exercise with PES appearing to offer more benefit (SMD -1.19, 95% CI -1.80 to -0.59).[76]

Disability (WOMAC index)

One study provided overall WOMAC scores at the end of treatment suitable for meta-analysis. [137] There was no significant difference between PES and placebo (SMD -0.28, 95% CI -0.83 to 0.27).

Adverse effects

Both studies assessing adverse events reported skin reactions at electrode sites in around a fifth of participants (for all treatment groups).

5.2.11.4 *Summary of effectiveness of PES*

The result of the meta-analysis showed a non-significant benefit of PES compared with placebo. However, the one satisfactory quality trial indicated that PES was more effective than MSE in reducing knee OA pain.

Table 14: Pulsed electrical stimulation (PES) trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session	Total number of sessions	Duration of treatment period (weeks)	Frequency of current	Pain outcomes used	Overall WOMAC score reported?
Callaghan 2005 [138]	PES(2) Placebo (2)	21 Y	UK	48	60	NR	28	Radiological	3 or higher	Poor	20 minutes	6	2	High frequency group: 27MHz Low frequency group: 27MHz	VAS 0-10 and pain subscale of AIMS	No
Durmus 2007 [76]	PES(2) MSE(2)	50 Y***	Turkey	100	55	NR	33	Clinical and radiological	3 or lower	Satisfactory	20 minutes	20	4	50 Hz	WOMAC pain subscale VAS 0-10 VAS 0-10	All 3 individual WOMAC subs scores reported
Fukuda 2008 [135]	PES Standard Care Placebo	NR N**	Brazil	100	61	67	NR	Clinical and radiological	NR - grade II or III using Gupta criteria	Poor	38 minutes (33 KJ dose) 19 minutes (17 KJ dose, and placebo)	39 (3 per week for 3 months (13 weeks)	13	33 KJ vs 17 KJ (27.12 Hz with a pulse frequency of 145Hz)	Other Pain VAS VAS (unclear what scoring).	Other overall score Lequesne
Garland 2007 [137]	PES(2) Placebo (2)	58 Y	USA	66	66	NR	31	Clinical and radiological	3 or higher	Poor	6 hours or more	Unclear ('Each day')	13	100Hz	WOMAC pain subscale VAS 0-100	Overall WOMAC score reported
Miranda-Filloy 2005[32]	PES(2) Placebo (2)	18(imputed) Y	Spain	80	Older than 40 years	NR	NR	Clinical and radiological	3 or higher	Poor	20 minutes	24	8	35 Hz	WOMAC pain subscale VAS 0-100 assumed as scor approx 30 Other Pain VAS VAS 0-100	No
Zizic 1995 [136]	PES Placebo	71 N**	USA	46	NR	NR	NR	Clinical and radiological	NR/ Unclear	Poor	6 to 10 hours a day	Unclear/ NR	4	100Hz	Other Pain VAS VAS 0-10	No

** Only change from baseline scores reported *** Only one study of this comparison MSE =Muscle-strengthening exercise

5.2.12 **Pulsed electromagnetic fields (PEMF)**

5.2.12.1 *Study characteristics*

Six trials studied PEMF (see Table 15); with a total of 521 participant pain scores analysed at the end of treatment (range 40 to 176). Five were full published papers and one was a conference abstract.[139] All were published between 1994 and 2005 in English, except for one paper in German[27]. Two of the six studies were conducted in the United States.

All studies recruited a general population (although the details were unclear for one study.[140] The mean age of participants ranged from 60 to 69 years, and the proportion of females ranged from 28 to 80%. Mean BMI was only reported in two studies (range 27 to 29 kg/m^2), and mean weight in only one study. The methods used for diagnosis were clinical and radiological in three studies, radiological alone in one study, and were unclear or not stated in two studies. Use of Kellgren & Lawrence scores was reported in one study, with one study using Lequesne scores, and four studies not reporting details on classification of severity.

All six studies were two-armed parallel trials comparing PEMF with placebo. Where stated, treatment was given for between six and 120 minutes, for between eight and 30 sessions (except for one trial that studied 147 sessions[141]), over periods ranging from nine days to six weeks (but most studies gave PEMF for six weeks). Pulse frequencies varied from one to 50Hz. Usual concomitant treatments (as required) were allowed in four studies[27 141-143], with no other medication permitted in one study,[140] and no details provided in the study reported as a conference abstract.[139]

Pain was measured using a VAS 0-10 scale in one study, a VAS 0-100 scale in two studies and a 5 point Likert scale in three studies. One study reported overall WOMAC scores, one reported individual WOMAC sub-scores, one reported Knee Society Scores and the remaining studies did not report on overall assessment of disability. Quality of life data were reported in two studies (one used EQ5D[141], and one used Activities of Daily Living).[143] Adverse effects were assessed in four studies.[27 141-143]

5.2.12.2 *Study Quality*

The number of participants randomised was clearly stated in four studies, and four studies reported using appropriate methods for concealing treatment allocation. However, only two studies reported using appropriate methods for randomisation.[141 143]

Eligibility criteria were adequately described in four trials, and group baseline characteristics appeared comparable in half the studies. Four studies clearly reported using an appropriate placebo treatment.[27 140 141 143] All studies were described as being double-blind, with five clearly stating that both patients and outcome assessors were blinded; half the trials blinded the treatment-givers.

Only two studies reported use of a power calculation for sample sizes[141 142] and only one study reported data for the intention-to-treat population.[141] However, all studies bar one[140] reported whether there were any losses to follow up. Half the studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived all six studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.12.3 *Results of effectiveness*

Pain

The four poor quality studies[27 139 140 142] that provided final values for analysis were similar to the two that did not[141 142] except that one of the latter studied a very high number of sessions (147).[141] The calculated pooled result (figure 23) was associated with considerable statistical heterogeneity, with no indication that PEMF provides benefit in pain reduction when compared to placebo; the one study suggesting benefit from PEMF had a small sample size and standard deviations that had to be

imputed for the meta-analysis. One of the four studies reported only a treatment effect size, making it suitable for NMA, but not for the pair wise meta-analysis.[140]

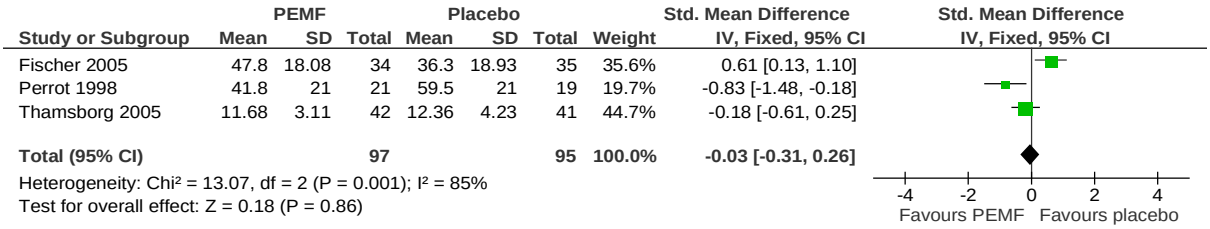


Figure 23: Pain (at end of treatment): PEMF versus placebo.

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

Of the four studies assessing adverse effects, one reported that none occurred,[143] and the others reported very few adverse events.

5.2.12.4 *Summary of effectiveness of PEMF*

There was no evidence of satisfactory quality to suggest that use of PEMF significantly reduced knee OA pain.

Table 15: Pulsed electromagnetic fields (PEMF) trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Frequency of pulse (Hz)	Strength of magnetic field	Pain outcomes used	Overall WOMAC score reported?
Fischer 2005 [27]	PEMF(2) Placebo(2)	69 Y	Slovenia	72	60	NR	29.3	Radiological	NR/Unclear	Poor	16	30	6	Unclear/not stated Low frequency	Varied between 3.4 and 13.6 µTesla	VAS 0-100 and Likert 5	No (knee society score)
Jacobson 2001 [140]	PEMF(1) Placebo(1)	176 Y	USA	NR	NR	NR	NR	Unclear/not stated	NR/Unclear	Poor	6	8	2	0.976 to 7.7 Hz	2.74*10-7 to 3.4*10-8 Gauss	Scale from 1 to 10 (no further details)	No
Perrot 1998 [139]	PEMF(2) Placebo(2)	40 Y	France	80	69	NR	NR	Unclear/not stated	NR	Poor	60	9	1.3	Unclear/not stated	Unclear/not stated	Other Pain VAS VAS 0-100	No
Pipitone 2001 [141]	PEMF Placebo	69 N*	UK	28	median around 63 (range 40-84)	NR	NR	Clinical and radiological	NR/Unclear	Poor	Unclear/ not stated	147	6	7.8Hz in morning and afternoon, and 3Hz in the evening	<0.5 Gauss	WOMAC pain Likert 5	Yes
Thamsborg 2005 [142]	PEMF(2) Placebo(2)	83 Y	Denmark	54	60	NR	27	Clinical and radiological	1 or higher	Poor	120	30	6	50 Hz	Unclear/not stated	WOMAC pain Likert 5	All 3 individual WOMAC sub scores reported
Trock 1994 [143]	PEMF Placebo	84 N**	USA	70	67	80	NR	Clinical and radiological	NR/Unclear	Poor	30	18	around 4 to 6	5 Hz for 10 minutes, 10 Hz for 10 minutes and 12 Hz for 10 minutes.	10-15 Gauss for 10 minutes and 15-25 Gauss for 20 minutes	Other Pain VAS VAS 0-10	No

* No Means, ** Only change from baseline scores reported

5.2.13 **Neuromuscular electrical stimulation (NMES)**

5.2.13.1 *Study characteristics*

Two trials studied NMES (see Table 16) with a total of 52 participant pain scores analysed at the end of treatment. Both were full papers published in English between 2003 and 2004, and both recruited a general population, with mean/median ages of participants ranging from 60 to 71 years. The proportion of females was around 80%, and the mean BMI around 30 kg/m^2, for both studies (mean weights were not reported). Both studies used clinical and radiological methods to diagnose participants. One study was of patients with a Kellgren & Lawrence score of at least one, and one trial studied patients with a score of at least two.

One trial compared NMES with muscle-strengthening exercise and the other used standard care as a comparator. Treatment was given for 15 or 30 minutes in 24 or 36 sessions over periods ranging from eight to 12 weeks. Other treatment parameters were generally not well-reported. Both trials were unclear in stating whether usual concomitant treatments were allowed, but one trial[144] did provide all participants with a 12 week self-management course.

Pain was measured using a VAS 20-80 scale in one study, and the McGill Pain Rating Index in the other. Only one study reported overall WOMAC scores. Adverse effects were monitored in only one study[144] and neither study reported quality of life data.

5.2.13.2 *Study Quality*

The number of participants randomised was clearly stated in both studies, but only one[73] clearly reported using appropriate randomisation methods; neither reported using suitable methods for concealing treatment allocation. Although both studies adequately described eligibility criteria, neither had comparable baseline characteristics between the groups. Neither study reported data for the intention-to-treat population, use of blinding of outcome assessors, or use of a power calculation for sample sizes. One study[144] reported clearly on whether there were any losses to follow up, but neither study achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived both studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.13.3 *Results of effectiveness*

Only the comparison with standard care (in which both groups also received a programme of self-management) provided final value mean data, reporting no significant difference between the groups (SMD 0.46, 95% CI -0.23 to 1.14).[144] No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis. One study assessed adverse effects, with none being reported for participants receiving NMES.

5.2.13.4 *Summary of effectiveness of NMES*

There was no evidence of satisfactory quality to suggest that use of NMES significantly reduced knee OA pain.

Table16: NMES trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean or median)	Weight (in kg: mean, median, or range)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Frequency of current (Hz)	Intensity of current	Pulse width/duration (micro secs)	Electrodes placed at acupuncture points?	Pain outcomes used	Overall WOMAC score reported?
Rosemffet 2004 [73]	NMES Muscle strengthening Exercise	18 N*	Argentina	77	median 60	NR	30.9	Clinical and radiological	2 or higher	Poor	30	24	8	25	According to patient tolerance	Unclear/ NR	Unclear/ NR	20-80 VAS scale	Yes
Talbot 2003 [144]	NMES(3) Standard care(3)	34 Y***	USA	79	mean 71	NR	30	Clinical and radiological	1 or higher	Poor	15	36	12	Unclear/ NR	Unclear/NR	300	Unclear/ NR	McGill Pain Rating Index	No

* No Means, *** Only one study of this comparison

5.2.14 **Interferential therapy**

5.2.14.1 *Study characteristics*

Four trials studied the effectiveness of interferential therapy (Table 17) with a total of over 180 participant pain scores analysed at the end of treatment (range 26 to 106, sample sizes were sometimes not clearly stated). All were English language studies, and all were full published papers, except for one conference abstract.[145]

All studies recruited a general population with mean ages of participants ranging from 59 to 67 years. The proportion of females (range 67 to 72%), mean BMIs (range 28 to 31 kg/m^2) and mean weights (range 78 to 86kg) were only reported for two trials.
Diagnoses were made using clinical and radiological methods in three studies, and clinical methods alone in one study. Only one study reported details on classification of disease severity (recruiting participants with Kellgren & Lawrence scores of one or more).

Two studies compared interferential therapy with placebo (with one[146] study also comparing four different modes of therapy), although comparisons were also made with no intervention[146] and TENS[127]. Treatment was administered for between 20 and 35 minutes, for between eight and 56 sessions over periods ranging from 12 days to eight weeks. Current frequencies and pulse widths (where stated) varied greatly.

Usual concomitant treatments (as required) were allowed in one study[127], and one study prohibited use of medication but did provide a specific programme of home exercise.[147] No relevant details were provided in two studies.[145 146]
Pain was measured using a pain VAS 0-10 scale in two studies, a WOMAC 5 point Likert scale in one study, and no details were provided for one study. One study reported individual WOMAC subs scores - although no overall score was provided - but the three other studies reported no relevant data. Quality of life data (using a VAS scale) and assessment of adverse effects were reported in only one study.[127]

5.2.14.2 *Study Quality*

The number of participants randomised was clearly stated in three studies, but only one study reported using appropriate methods for randomisation and allocation concealment.[127] Eligibility criteria were adequately described in all trials, but group baseline characteristics appeared comparable in only one study.[147] The three placebo-controlled studies all reported use of an appropriate placebo treatment. Three studies blinded patients, two blinded outcome assessors and one study blinded treatment-givers. Only one study reported use of a power calculation for sample sizes, and whether there were any losses to follow up (although less than 90% of the population were followed up fully).[127] None of the studies clearly reported data for the intention-to-treat population.

When the overall study quality ratings were derived all four studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.14.3 *Results of effectiveness*

Two poor quality studies provided final value data for analysis.[145 147] both were small trials, and both had to have their standard deviations imputed. Also the one comparison with placebo prohibited the use of medication, making it different from other trials[147] The results found interferential therapy to be more effective than placebo or no intervention: SMD -0.93 (95% CI: -1.74, -0.11) and SMD -1.64 (95% CI: -2.48, -0.81) respectively.

No studies provided overall WOMAC scores at the end of treatment. The only study to assess adverse effects found none occurring in the interferential therapy group.

5.2.14.4 *Summary of effectiveness of inferential therapy*

There was no evidence of satisfactory quality to suggest that use of interferential therapy significantly reduced knee OA pain.

Table 17: Interferential therapy trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (years: mean, unless stated)	Weight (in kg: mean, median, or range)	Mean BMI (kg/m^2)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of individual session (minutes)	Total number of sessions	Duration of treatment period (weeks)	Frequency (Hz)	Pulse width/duration	Pain outcomes used	Overall WOMAC score reported?
Adedoyin 2002 [147]	Interferential(5) Placebo(5)	26 (imputed) Y	Nigeria	67	59	mean 78	28	Clinical and radiological	NR/ Unclear	Poor	20	8	4	100Hz for 15 mins, reduced to 80Hz for further 5 mins.	0.033s	VAS 0-10,	No
Burch 2008 [127]	Interferential TENS	106 N**	USA	72	62	86	31	Clinical and radiological	1 or higher	Poor	15 of interferential followed by 20 of patterned stimulation	56	8	Base of 5000Hz, and a pre-modulated beat frequency sweeping between 1-150Hz	between 3 and 102 micro seconds	WOMAC pain Likert 5	All 3 individual WOMAC subs scores reported
Defrin 2005 [146]	Interferential(2) Placebo(2) No intervention(2)	55 (imputed) Y	Israel	NR	67	NR	NR	Clinical	NR/ Unclear	Poor	20	12	4	Constantly ranged between 30Hz and 60Hz.	NR/ Unclear	VAS 0-10	No
Young 1991 [145]	Interferential Placebo	NR N*	Canada	NR	18 to 75	NR	NR	Clinical and radiological	NR/ Unclear	Poor	25	10	1.7	NR/Unclear	NR/ Unclear	Reported WOMAC but unclear how assessed	No

* No Means, ** Only change from baseline scores reported

5.2.15 **Heat treatment**

5.2.15.1 *Study characteristics*

Six trials studied heat treatment interventions (see Table 18) with a total of 349 participant pain scores analysed at the end of treatment (range 30 to 104); all were full published papers reported in English, between 2004 and 2010 except for one study published in 1974. Two studies were conducted in Italy, with others in the United States, Israel, UK and Thailand.

All studies recruited a general population. The mean ages of participants ranged from 61 to 74 years, and the proportion of females ranged from 63 to 100%. Mean BMI (26 kg/m^2) and mean weight (71 kg) were reported in single studies. The methods used for diagnosis were clinical and radiological in three studies, radiological in one study, clinical in one study and unclear in one study. Only two studies reported using Kellgren & Lawrence scores.

Treatment was with shortwave diathermy in four studies, with single trials of radiofrequency electromagnetic radiation[132] and a heat-retaining knee sleeve.[148] Two studies were three-armed trials, one compared heat treatment with balneotherapy and standard care[98] and another with ice/cooling treatment and placebo.[149] The remaining studies were two-armed trials, two of which compared heat treatment to placebo and single studies compared to TENS[132] and muscle strengthening exercise.[75]

Sessions lasted between 15 and 20 minutes. The total number of sessions varied from 5 to 10, over periods ranging from 5 days to 4 weeks. One study used an additional home exercise programme as part of the intervention.[150]

Pain was measured using a variety of scales; three studies measured WOMAC pain (one using a VAS 0-10 scale, and two using a Likert scale). Two studies reported overall WOMAC scores and one reported individual WOMAC sub-scores. One study reported quality of life outcomes using the AIMS1[98] and three studies reported adverse effects.[98 148 150]

5.2.15.2 *Study Quality*

The number of participants randomised was clearly stated in all studies, but three [75 148 150] clearly reported using appropriate randomisation methods, and only two studies[148 150] reported suitable methods for concealing treatment allocation. Eligibility criteria were adequately described in four trials, but group baseline characteristics appeared comparable in only two,[148 149] and four studies reported blinding outcome assessors. Only two studies used a power calculation for sample sizes[75 150] and three studies clearly reported using data for the intention-to-treat population.[75 98 149] However, all studies bar one[132] reported whether there were any losses to follow up. Four studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived all studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.15.3 *Results of effectiveness*

5.2.15.4 *Pain*

Two trials did not provide final value mean data suitable for analysis.[98 150] One was a comparison with balneotherapy and standard care[98] Neither suggested a treatment benefit for heat treatment. They were not dissimilar to the four trials which could be analysed;[75 132 148 149] these included two comparisons with placebo[148 149] (figure 24), a comparison with TENS[132] (SMD 0.06, 95% CI -0.56 to 0.68), a comparison with MSE[75](SMD 0.80, 95% CI 0.23 to 1.38), and a comparison with ice/cooling treatment[149] (SMD 0.69, 95% CI -0.02 to 1.41), none of which demonstrated a benefit of heat treatment.

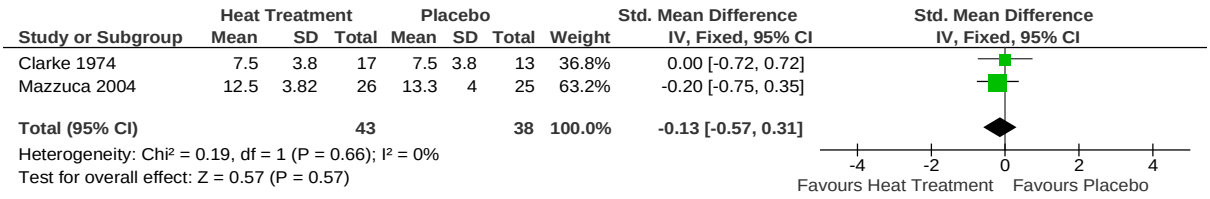


Figure 24: Pain (at end of treatment): heat treatment versus placebo (all studies).

Disability (WOMAC index)

One study provided overall WOMAC scores at the end of treatment suitable for meta-analysis; MSE significantly reduced the WOMAC index compared to shortwave diathermy heat treatment (results reported in section 6.2.2.3).

Adverse effects

Of the three studies that assessed adverse effects, no adverse effects for heat treatment were reported.

5.2.15.5 *Summary of effectiveness of heat treatment*

There was no evidence of satisfactory quality to suggest that use of heat treatment significantly reduced knee OA pain.

Table 18: Heat treatment trials: Study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean) kg/m^2	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Type of treatment	Pain outcomes used	Overall WOMAC score reported?
Alcidi [132]	Heat(2) TENS(2)	40 Y (only one study)	Italy	*85*	*66*	NR	NR	Clinical and radiological	NR	Poor	20	5	0.7	Radiofrequency electromagnetic radiation	VAS *1-100*	No
Bezalel [75]	Heat(2) MSE(2)	50 Y (only one study)	Israel	*74*	*74*	NR	NR	NR	NR	Poor	20	6	4	Shortwave diathermy	WOMAC pain Likert 5	Yes
Cantarini [98]	Heat Balneotherapy, Standard care	74 N*	Italy	*63*	*64*	*71*	NR	Clinical and radiological	3 or lower	Poor	15	10	3	Shortwave diathermy	VAS *0-100*	No
Clarke [149]	Heat(2) Ice/cooling treatment(2) placebo(2)	30 Y	UK	*69*	*61*	NR	NR	Radiological	NR	Poor	NR	9	3	Shortwave diathermy	Likert *0-3; max score 17*	No
Mazzuca [148]	Heat(2) Placebo(2)	51 Y	USA	*77*	*63*	NR	NR	Clinical and radiological	2 or higher	Poor	NA	NA	4	Heat-retaining knee sleeve	WOMAC pain Likert 5	Individual WOMAC sub scores
Rattanachaiyanont [150]	Heat Placebo	104 N**	Thailand	*100*	*63*	NR	*26*	Clinical	NR	Poor	20	9	3	Shortwave diathermy	WOMAC pain VAS 0-10	Yes

* No Means, ** Only change from baseline scores reported MSE = Muscle strengthening exercise

5.2.16 **Ice/cooling treatment**

5.2.16.1 *Study characteristics*

Four trials assessed ice/cooling treatment (see Table 19) with a total of 211 participant pain scores analysed at the end of treatment (range 33 to 100); all were full published papers in English published between 1974 and 2009 except for a conference abstract in German[30]; two studies were conducted in Turkey and others in the UK and USA.

All participants were drawn from a general knee OA population. Mean age ranged from 56 to 61 years, and the proportion of females ranged from 48 to 91%; demographic data was not reported for one study[30]. Mean BMI was reported in only one study (30 kg/m^2) as was the mean weight (88kg)[126]; the three remaining studies did not report this data. The methods used for diagnosis were clinical and radiological in two studies[30 37] and radiological in two studies.[126 149] Kellgren & Lawrence scores were not reported in any study.

One study was a two-armed trial comparing ice/cooling treatment with standard care.[30] There were two three armed trials with ice/cooling treatment compared with: heat treatment and placebo;[149] and TENS and no intervention.[126] A four-armed trial compared ice/cooling treatment with acupuncture, TENS and placebo TENS.[37]

The duration of sessions ranged from 10 to 20 minutes and the numbers of sessions from 1 to 40. The duration of the treatment period ranged from 1 day to 3 weeks.

Pain was measured using Pain Likert scales in two studies[37 149] and a pain VAS in another.[126] Quality of life data were not reported for any of the studies and adverse effects were reported in one study.[37]

5.2.16.2 *Study Quality*

All studies clearly stated the number of participants randomised though it was unclear whether any study used appropriate methods for randomisation and only one study used appropriate methods for concealing treatment allocation.[126] None of the studies conducted any form of blinding, with the exception of one study which blinded outcome assessors.[37]. Eligibility criteria were adequately described in only one trial.[37] Group baseline characteristics were comparable in two studies[126 149] but insufficient details were reported in the other studies.[30 126] None of the studies reported use of a power calculation for sample sizes. Three studies reported data for the intention-to-treat population, losses to follow up and full follow up for at least 90% of participants;[37 126 149] in the other study[30] these were unclear or not reported.

When the overall study quality ratings were derived all studies were rated as being of poor quality. Full details of study quality are reported in Appendix 10.2.

5.2.16.3 *Results of effectiveness*

Pain

One trial did not provide final value mean data for analysis.[30] This trial was similar to the other trials except that it had investigated a more intensive therapy regimen (40 sessions over three weeks). Three trials could be analysed[37 126 149] and comprised two comparisons with TENS,[37 126] one with heat treatment[149] and one with acupuncture.[37] The two trials comparing ice cooling with TENS were not pooled because one of the treatments was given only once. There was no significant difference between the groups for both studies (SMD -0.41, 95% CI -1.27 to 0.46[126], and SMD 0.43, 95% CI -0.13 to 1.00[37]). The comparison of ice/cooling treatment with heat treatment found no statistically significant difference between treatments (SMD -0.69, 95% CI -1.41 to 0.02).[149] An SMD could not be calculated for the one study of ice treatment versus acupuncture since the acupuncture arm had an end of treatment pain score of zero.

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

None of the studies assessed adverse effects for ice/cooling treatment.

5.2.16.4 *Summary of effect of ice/cooling treatment*

There was no evidence of satisfactory quality to suggest that use of ice/cooling treatment significantly reduced knee OA pain.

Table 19: Ice/cooling trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Pain outcomes used	Overall WOMAC score reported?
Arman 1988 [30]	Ice/cooling Standard care	33 N*	Turkey	NR	NR	NR	NR	Clinical & radiological	NR	Poor	10	40	3	NR	No
Clarke 1974 [149]	Ice/cooling(2) Heat treatment(2) placebo(2)	45 Y (only study)	UK	69	61	NR	NR	Radiological	NR	Poor	NR	9	3	Likert 0-3 (max score of 17)	No
Pietrosimone 2009 [126]	Ice/cooling(1) TENS(1) No intervention(1)	33 Y	USA	48	56	88	30	Radiological	NR	Poor	20	1	1 day	VAS	No
Yurtkuran 1999 [37]	Acupuncture(2) Ice/cooling(2) TENS(2) Placebo TENS(2)	100 Y	Turkey	91	58	NR	NR	Clinical & radiological	NR	Poor	20	10	2	Likert 1-5	No

* No Means

5.2.17 **Laser or light therapy**

5.2.17.1 *Study characteristics*

Seven trials studied laser or light therapy interventions (see Table 20) with a total of over 260 participant pain scores analysed at the end of treatment (range 29 to 60, sample sizes were sometimes not clearly stated); five were full published papers and two were conference abstracts. All studies were reported in English, between 1989 and 2009. Two studies were conducted in Turkey, with others in Korea, China, Sweden, Denmark and Israel.

Four studies recruited a general population, two [151 152] studied only participants with both knees affected by osteoarthritis, and in one study the population was unclear.[153] The mean/median ages of participants, where reported, ranged from 58 to 74 years, and the proportion of females ranged from 68 to 90%. Mean BMI was reported in only two studies (range 29 to 30 kg/m^2) and mean weight in no studies. The methods used for diagnosis were clinical and radiological in six studies and unclear in one study. Four studies reported using Kellgren & Lawrence scores (with all patients scoring at least 2), in two studies the method of diagnosis was not reported and in one study they used OA severity radiographic grades.

Three studies were three-armed trials, one compared different light doses with placebo[154], another red light and infrared light with placebo[151] and another different light intensities with placebo.[152] The remaining studies were two-armed trials compared with placebo[153 155 156] or standard care.[157]

Sessions lasted between 5 and 30 minutes. The total number of sessions varied from 6 to 56, over periods ranging from 10 days to 8 weeks. One study[154] used additional home exercise as part of the intervention for all arms of the trial where all groups underwent straight leg raising exercise over 14 weeks.

Usual (or trial-specific) concomitant treatments, as required, were allowed in three studies and no details were provided in two studies. Two studies required at least one of the study arms to take specified doses of analgesics.[152 154]

Pain was measured using a variety of scales; only three studies measured WOMAC pain (one using a VAS 0-10 scale, and two using a Likert scale). Only one study reported overall WOMAC scores and three reported individual WOMAC sub-scores. None of the studies reported quality of life outcomes and one study reported adverse effects.[152 155-157]

5.2.17.2 *Study Quality*

The number of participants randomised was clearly stated in all studies, but only one[155] clearly reported using appropriate randomisation methods, and only two studies[155 156] reported suitable methods for concealing treatment allocation.
Eligibility criteria were adequately described in four trials, but group baseline characteristics appeared comparable in only three, and just two studies[155 156] reported blinding outcome assessors. Only one study used a power calculation for sample sizes[156] and four studies clearly reported using data for the intention-to-treat population.[152 154-156] However, all studies bar two[153 157] reported whether there were any losses to follow up. Five studies achieved full follow up for at least 90% of participants.

When the overall study quality ratings were derived four studies were rated as being of poor quality, and three were rated as being of satisfactory quality.[152 154 155] Full details of study quality are reported in Appendix 10.2.

5.2.17.3 *Results of effectiveness*

Pain

Four of the seven trials did not present final value data for analysis,[153 154 156 157] one of which[154] was rated to be of satisfactory quality. The four included three comparisons with placebo and one with standard care, for two of which the sample size was unclear. Otherwise they were generally similar to the three analysed trials, which were all placebo comparisons (one poor quality and two satisfactory quality studies).[151 152 155] The pooled result found no significant difference between laser or light

therapy interventions and placebo for end of treatment pain (Figure 25) but a high degree of heterogeneity was present. A sensitivity analysis in which the one poor quality study, which was the only study that reported a significant benefit of laser therapy, was removed[151] markedly reduced the heterogeneity (I^2=2%) but the pooled result was still close to zero and not statistically significant (Figure 26).

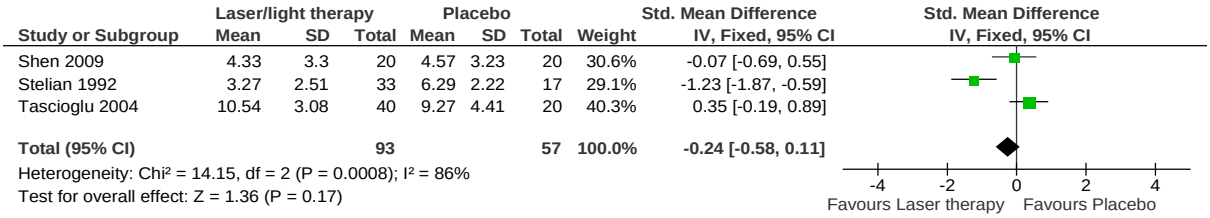


Figure 25: Pain (at end of treatment): laser/light therapy interventions versus placebo (all studies).

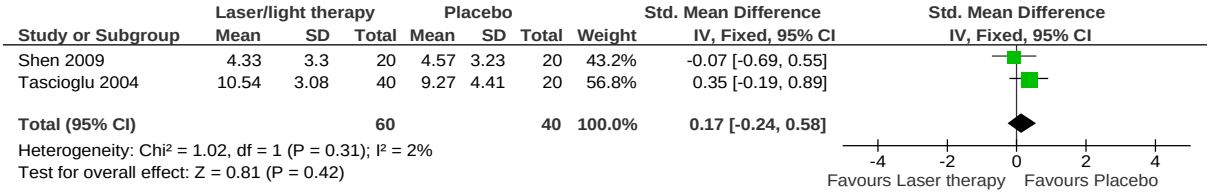


Figure 26: Pain (at end of treatment): laser/light therapy interventions versus placebo (satisfactory studies).

Of the four studies ineligible for meta-analysis three compared laser or light therapy interventions with placebo and one with standard care. Two of the trials did not report means and two presented change from baseline scores. In one study there were no significant differences in pain compared with placebo,[156] whilst in another, which was of satisfactory quality, and which reported only median values, pain was significant reduced compared with placebo.[154]

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

Of the four studies that assessed adverse effects, three reported that there were no adverse effects and one study reported that three patients had an area of reddened skin at the irradiated site that disappeared within few hours without treatment.

5.2.17.4 *Summary of effectiveness of laser or light therapy*

The satisfactory quality studies that were included in the analysis provided no evidence to suggest that use of laser/light therapy significantly reduced knee OA pain. However, one satisfactory quality study, which reported only medians, did suggest a benefit of laser therapy over placebo.

Table 20: Laser/light therapy trials: Study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyses? (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks)	Light dose	Pain outcomes used	Overall WOMAC score reported?
Bulow 1994 [156]	Laser Placebo	29 N*	Denmark	83	Median 74	NR	NR	Clinical and radiological	NR	Poor	15	9	3	Accumulated dose 202.5J	Likert scale: 0-3, maximum score 126	No
Gur 2003 [154]	Laser Placebo	90 N*	Turkey	80	60	NR	30	Clinical and radiological	2 or higher	Satisfactory	5	10	2	Group 1: 30 J accumulated dose Group 2: 20 J accumulated dose	VAS 0-10	Yes
Kim 2006 [157]	Laser Standard care	unclear N**	Korea	NR	NR	NR	NR	Clinical and radiological	2 or higher	Poor	30	56	8	402 laser diodes, 650nm, 50mW	WOMAC pain subscale VAS 0-10	Individual WOMAC subs scores reported
Nivbrant 1989 [153]	Laser Placebo	unclear N**	Sweden	NR	NR	NR	NR	Clinical and radiological	OA severity radiographic grades 2-4	Poor	20	6	6	NR	VAS	No
Shen 2009 [155]	Laser(2) Placebo(2)	40 Y	China	90	58	NR	NR	Clinical and radiological	2 or higher	Satisfactory	20	12	4	283J	WOMAC pain Likert 5	All 3 individual WOMAC subscale scores
Stelian 1992 [151]	Laser(2) Placebo(2)	50 Y	Israel	68	68	NR	NR	NR	NR	Poor	30	20	10 days	Red light 10.3J; infrared light 11.1J	Other Pain VAS 0-10 Pain Likert 0-5 scale	No
Tascioglu 2004 [152]	Laser(4) Placebo(4)	60 Y	Turkey	70	62	NR	29	Clinical and radiological	2 or 3	Satisfactory	10 for 1st group, 5 for 2nd group	10	2	Group 1: 3 joule per tender joint (5 joints in all), dose/treatment 15 joule. Group 2: 1.5 joule per tender joint (5 joints in all), dose/treatment 7.5 joule.	WOMAC pain Likert 5 Other Pain VAS 0-100	All 3 individual WOMAC subscale scores

* No Means ** Only change from baseline scores reported

5.2.18 **Manual therapy**

5.2.18.1 *Study characteristics*

Five trials assessed manual therapy (see Table 21) with a total of 367 participant pain scores analysed at the end of treatment (range 39 to 114); all were full published papers in English published between 2003 and 2009; two studies were conducted in Australia and others in the USA, Germany and South Africa.

All participants were drawn from a general knee OA population. Mean age ranged from 56 to 68 years, and the proportion of females ranged from 63 to 78%. Mean BMI was reported in only one study (mean 29 kg/m^2) [158] and the mean weight was reported in two studies (and ranged from 82 to 83kg).[28 159] The methods used for diagnosis were clinical and radiological in all studies, but the Kellgren & Lawrence score (2 or higher) was reported for only one study.[28]

Two studies were two-armed trials comparing manual therapy with standard care.[158 159] and one study compared manual therapy with placebo.[160] The remaining studies were three armed trials: one compared two types of manual therapy (accessory mobilisation and by hand) with placebo[161] and the other compared manual therapy with muscle-strengthening exercise or standard care.[28] One study had a cross-over design.[161]

Where stated, the duration of each session ranged from 10 to 60 minutes and the number of sessions from 1 to 16. The duration of the treatment period ranged from a single 10 minute session[161] to 8 weeks.

Pain was measured using a pain VAS in all five studies, though studies also used a WOMAC pain Likert,[161] WOMAC pain subscale VAS 0-100[158] and the numerical Pain Rating Scale.[159] The overall WOMAC score was reported in one study.[158] Quality of life data were not reported for any study. Adverse effects were reported for two studies.[158 159]

5.2.18.2 *Study Quality*

All studies clearly stated the number of participants randomised, two studies used appropriate methods for randomisation[158 160] and one used appropriate methods for concealing treatment allocation.[160] One study reported that the study was double blind, with blinding of both the patient and outcome assessor[161] and another study reported blinding of both the patient and outcome assessor [160] Eligibility criteria were adequately described in all studies, whilst group baseline characteristics were comparable in two studies,[159 161] or had significantly different WOMAC pain at baseline.[158]Two studies reported use of a power calculation for sample sizes.[158 161] Three studies reported data for the intention-to-treat population,[158 160 161] and four reported losses to follow up.[158-161] Follow up for at least 90% of participants was reported in three studies.[159-161]

When the overall study quality ratings were derived one study was considered to be of good quality[161] and the rest of poor quality.[28 158-160]. Full details of study quality are reported in Appendix 10.2.

5.2.18.3 *Results of effectiveness*

Pain

Two studies (both poor quality) provided data suitable for a meta-analysis comparison of manual therapy with standard care.[28 159] No significant difference between manual therapy and standard care were found for end of treatment pain (Figure 27).

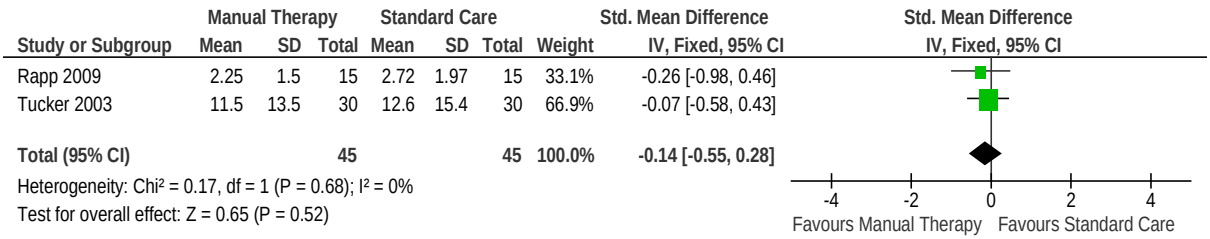


Figure 27: Pain (at end of treatment): manual therapy versus standard care (all studies).

One study compared manual therapy with placebo compared with the placebo group manual therapy yielded a significant decrease in pain (SMD -0.65, 95% CI -1.27 to -0.02).[160]
Of the two studies ineligible for meta-analysis, both presented change from baseline scores and both (including the one good quality trial[161]) reported results favouring manual therapy over placebo[158] or standard care.[161]

Disability (WOMAC index)

No studies provided overall WOMAC scores at the end of treatment suitable for meta-analysis.

Adverse effects

Of the two studies that assessed adverse effects, one stated that a single patient reported increased discomfort and withdrew from the trial, and the other reported no adverse effects.

5.2.18.4 *Summary of effectiveness of manual therapy*

There was no evidence of satisfactory quality to suggest that manual therapy significantly reduced knee OA pain. However, one good quality trial whose data could not be included in the analysis did indicate a benefit of manual therapy over placebo.

Table 21: Manual therapy trials: study details

Author	Treatments (adjunct code, if suitable for analysis)	Number analysed at end of treatment, Data suitable for pain analyes? (Y/N)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean)	BMI (Mean)	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality	Duration of session (minutes)	Number of sessions	Duration of treatment period (weeks, unless stated)	Pain outcomes used	Overall WOMAC score reported?
Moss 2007 [161]	MT Placebo	114 N**	Australia	66	65	NR	NR	Clinical & radiological	NR	Good	10	1	10 mins	WOMAC Likert 5 VAS 0-10	No
Perlman 2006 [158]	MT Standard care	68 N**	USA	78	68	NR	29	Clinical & radiological	NR	Poor	60	12	8	WOMAC VAS 0-100 VAS 0-100	Yes
Pollard 2008 [160]	MT(2) Placebo(2)	43 Y (only study)	Australia	NR	56	NR	NR	Clinical & radiological	NR	Poor	NR	6	2	VAS 0-10	No
Rapp 2009 [28]	MT(2) MSE(2) Standard care(2)	39 Y	Germany	64	60	83	NR	Clinical & radiological	2 or higher	Poor	20	16	8	VAS 0-10	No
Tucker 2003 [159]	MT(2) Standard care(4)	103 Y	South Africa	63	59	82	NR	Clinical & radiological	NR	Poor	NR	8	3	VAS 0-100 Numerical Rating Scale (NRS)	No

** Only change from baseline scores reported MT = Manual Therapy MSE= Muscle-strengthening exercise

6. NETWORK META-ANALYSIS

6.1 Quantity and Quality of data

The total number of trials potentially available for analysis was 91. A flow chart indicating the reasons studies were unsuitable for NMA is presented in Figure 28.

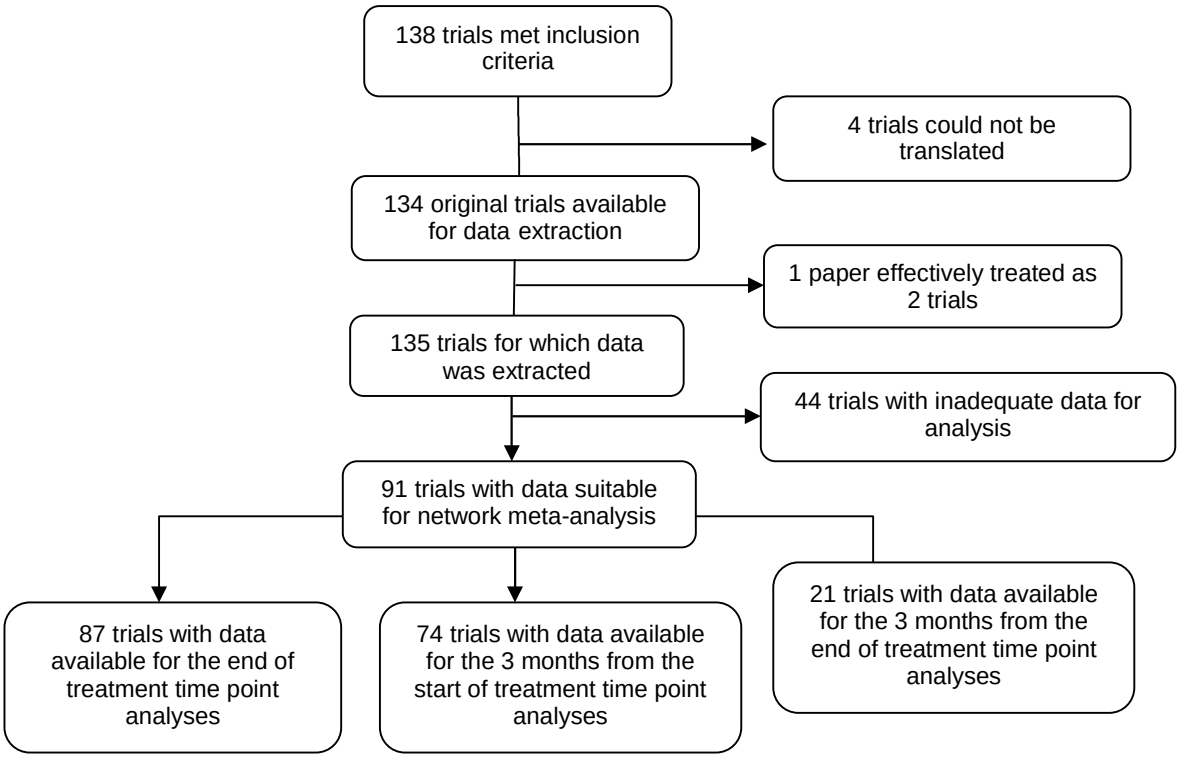


Figure 28: Flow chart of trials available for analysis.

In these 91 trials, a variety of pain scales were reported. The frequency of these scales is reported in Table 22. Some VAS scales utilised scores for only the range specified, as in 0-100, some used cumulative (summed) scoring, and others standardised the cumulative scores (e.g. a cumulative score of 300 out of 500 became 60 on a 0–100 scale). These are not distinguished in Table 22. If a trial reported more than one scale, only the primary scale prioritised for use in an analysis is counted here.

The number of trials with each of these pain scales by the main interventions is reported in Appendix 10.4.3 NMA Appendix Table 1. The variation in the mean standard deviations across the main interventions that utilised the two main scales (VAS 0-10 and WOMAC Likert 5) is reported in Appendix 10.4.3 NMA Appendix Tables 2 and 3.

Table 22: A frequency table of the primary scales reported in the 91 trials. The primary scale is the scale used in an analysis

AIMS VAS 0 - 10	2
AIMS2 pain subscale (0-10 scale)	1
AIMS-4 item	1
Borg Scale 0-10	1
KSS 10 pt Likert	1
Likert	2
Likert 1-6	1
Likert Four graded pain scale (1-none,4-severe)	1
Lysholm pain scale (no pain =25, constant pain =0)	1
Lysholm scores	1
McGill	1
NR VAS 0 - 10	1
NRS VAS 0 - 10	2
Numerical Pain Rating Scale and VAS 0-100	1
Scale 1 (minimal) to 10 (maximal) and pain diary	1
VAS 0 - 10	18
VAS 0 - 100	12
WOMAC	
Likert 5	20
VAS 10	6
VAS 100	11
std mean (VAS 100)	3
Not specified	3

Three time points were defined in the methods section (section 6.2.4.2). However, for the 3 months after end of treatment time point (evaluated in 21 trials) no connected network that included acupuncture existed (only two very small other networks existed, each comprising three interventions). Therefore networks were constructed only for end of treatment and 3 months from start of treatment. In many cases these were very similar since for many trials for End of Treatment and 3 months from start of treatment were the same time point. The 12 connected networks for the primary outcome (pain) obtained for these analyses, and the sensitivity analysis excluding outliers, are presented in figures 1-13. In the figures, each solid arrow indicates that there is a direct comparison entered into the analysis. A 3-armed trial with arms A, B and C, provides 2 data points in a data set reflecting 2 comparisons A vs B and C vs B. These comparisons are represented by solid arrows and these determine the existence of evidence triangles on which consistency of direct and indirect evidence can be tested. The numbers in the figures indicate the level of inconsistency between the direct and indirect evidence for that comparison. A value of 1 represents complete inconsistency, and a value of zero indicates perfect consistency. Tables presenting the comparisons included in each connected network for the overall WOMAC score analyses are presented in Appendix 10.4.2 as Tables 1-12.

Three networks were evaluated for each time point:
Therapy-plus-adjunct interventions network: a total of 110 possible interventions were defined. There were 22 primary interventions and each primary intervention had 5 possible variations of concomitant treatments. These are detailed in Table 3.
Therapy-only interventions network comprising 22 primary interventions. These are also detailed in Table 3.

Grouped interventions network: there were 13 possible interventions in the general network. These 13 categories included all of the 22 main interventions in the specific network. All concomitant variations of each main intervention were considered the same. The categories are listed in Table 24.
For each network at each time point, an analysis was planned including:
Studies of any-quality
Higher-quality studies

Study and population characteristics across all interventions in the review are summarised in Table 25. Appendix 10.4.1 presents these details for each trial included in the NMA and illustrates that they are generalisable to a general OA of the knee population.

Table 24: The 13 interventions in the grouped interventions set

Intervention	
Acupuncture	1
Sham acupuncture	2
Balneotherapy	3
Braces	4
Exercise - Aerobic (weight bearing)	5
Exercise - Muscle strengthening (non-weight bearing)	6
Physiotherapy (heat treatment; ice/cooling treatment, interferential therapy; laser/light therapy; manual therapy; NMES; Pulsed electrical stimulation (PES); Pulsed electromagnetic fields (PEMF); static magnets; TENS)	7
Insoles	8
Tai Chi	9
Weight loss	10
Standard care	11
Placebo	12
No intervention	13

Table 25: Summary characteristics of trials included in the systematic review

Intervention	No. of trials eligible for the review (no. of pts*)	Type of population recruited (no. of studies)	Range of mean ages (years)	Range of BMIs (kg/m^2) (where reported)	Range of % female	Comparators (no. of treatment arms†)
Acupuncture	22 (2167)	General(20), both knees affected(1), awaiting surgery(1)	58-85	29-33	50-96	Sham acupuncture(12), standard care(11), TENS(3), muscle strengthening exercise(1), ice/cooling(1)
Balneotherapy	14 (1008)	General(12), both knees affected(2)	‡54-70	26-32	47-100	Placebo(8), standard care(6), heat treatment(1)
Braces	2 (227)	Knee mal-alignment(2)	59	NR	38-83	Standard care(2)
Aerobic exercise	9 (880)	General(6), both knees affected(2), overweight or obese(1)	‡54-75	33-34	50-100	Standard care(9), muscle strengthening exercise(2), weight loss(1)
MSE	30 (2771)	General (23), both knees affected(4), awaiting surgery(2), knee mal-alignment(1)	‡53-77	24-33	31-100	Standard care(19), placebo(3), no treatment(2), aerobic exercise(2), heat treatment(1), TENS(1), acupuncture(1), PES(1), manual therapy(1), NMES(1)
Heat treatment	6 (349)	General(6)	61-74	NR	63-100	Placebo(3), standard care(1), TENS(1), muscle strengthening exercise(1), balneotherapy(1), ice/cooling(1)
Ice/cooling treatment	4 (211)	General (4)	56-61	30 (one study)	48-91	TENS(2), acupuncture (1), standard care(1), heat treatment(1), placebo(1), no treatment(1)
Insoles	6 (669)	General(4), knee mal-alignment(2)	58-68	29-33	54-100	Placebo(6)
Interferential therapy	4 (180)	General (4)	59-67	28 (one study)	67-72	Placebo(2), TENS(1), no treatment(1)
Laser/light therapy	7 (260)	General (5), both knees affected(2)	58-74	29-30	68-90	Placebo(6), standard care(1)
Manual therapy	5 (367)	General (5)	56-68	NR	63-78	Standard care(3), placebo(2), muscle strengthening exercise(1)
NMES	2 (52)	General(2)	60-71	30-31	77-79	Standard care(1), muscle strengthening exercise(1)
PES	6 (210)	General(6)	55-66	28-33	46-100	Placebo(5), standard care(1), muscle strengthening exercise(1)
PEMF	6 (521)	General(6)	60-69	27-29	28-80	Placebo(6)
Static magnets	3 (131)	General(3)	63-65	NR	60-79	Placebo(3)
Tai Chi	4 (307)	General (4)	65-70	26-28	75-93	Standard care(4)
TENS	17 (730)	General(16), awaiting surgery(1)	56-85	28-31	48-97	Placebo(11), standard care(3), acupuncture (3), Ice/cooling(2), heat treatment(1), interferential (1), no treatment, muscle strengthening exercise(1)
Weight loss (dieting)	4 (781)	Overweight or obese(4)	61-70	33-35	26-89	Standard care(4), aerobic exercise(1)

* Number of patients analysed by the primary studies for end of treatment pain - this was not always clearly stated †Different doses of the same treatment in a trial were pooled, counting as one arm ‡ Trial reported mean age by treatment group, and contained a group with a mean age ≥55. Some studies compared two or more different intervention NR=Not reported

Study quality
Table 26 shows the number of trials for each main intervention that are good or satisfactory, or poor: 45% or more of the acupuncture, sham acupuncture, muscle-strengthening exercise, insoles, laser/light therapy, and Tai Chi trials are of good or satisfactory (higher) quality.

Table 26: Trials by main intervention and quality across the 91 trials that had adequate data available for analysis

	Number of trials						
	Study quality				Study quality		
Treatment	Higher*	Poor	Any	Treatment	Higher*	Poor	Any
Acu	8	10	18	NMES	0	1	1
Bal	1	8	9	PES	1	3	4
Bra	0	1	1	PEMF	0	4	4
ExAe	1	5	6	Mag	0	1	1
ExMu	7	17	24	Tai	2	2	4
Hea	0	4		TENS	0	11	11
Ice	0	3	3	Wei	2	1	3
Ins	3	0	3	SC	13	30	43
Int	0	2	2	P	6	28	34
Las	3	1	4	NoTr	1	3	4
Man	0	3	3	ShAcu	5	3	8

* Good or satisfactory quality

6.2 Presentation of results of NMA

Given the large number of interventions in the network the results can be presented in a variety of ways. In this report most of the results presented for the comparison with standard care. The full results are available at: http://www.york.ac.uk/inst/crd/Documents/FullResultsPain.xlsm
http://www.york.ac.uk/inst/crd/Documents/FullResultsWOMAC.xlsm

6.3 Pain - end of treatment analyses

There were 87 trials with data that could be used in the end of treatment analyses of pain. The number of trials actually included in each analysis depended on how many of them formed part of a connected network with acupuncture.

There was considerable variation in the average treatment duration across the main interventions. See Appendix 10.4.3 Figure 1.

The treatment effects of each treatment compared with standard care are presented. For the Therapy-only interventions network acupuncture is compared with each of the other interventions. For brevity, only the results where acupuncture is significantly more effective than the comparator at a 95% level of credibility are referred to in the text.

6.3.1 The therapy-plus-adjunct interventions set

6.3.1.1 *All studies (any-quality)*

Of the potential 110 Therapy-plus-adjunct interventions, 35 interventions formed part of a connected network with acupuncture and the evidence was informed by 79 trials. The network, presented in Figure 29, shows strong evidence of a high level of inconsistency between the direct and indirect evidence involving the PES/UT, heat treatment/UT, placebo and TENS/UT nodes. This suggests that there is a possibility of bias or lack of exchangeability between the comparisons and hence the

credible intervals of the treatment effect for these treatments may be underestimated. Stable estimates were produced by the model for the SMDs and for the between-study SD of the random effects distribution, τ. The mean between-study SD across the comparisons was 0.32 SMD.

Table 27 presents the treatment effects of each intervention compared to standard care with usual treatment in order of mean effectiveness. The median effectiveness rank and the uncertainty in the rank are also presented. All interventions other than NMES with treatment as usual plus home exercises/education (NMES/UT+EX) had a mean estimate that favoured the intervention over standard care with usual treatment. PES, acupuncture, balneotherapy, sham acupuncture, laser/light treatment, static magnets and Tai Chi all showed a treatment benefit over standard care with the 95% credible intervals not crossing the line of no effect, regardless of the adjunctive treatment. The 95% credible intervals of the treatment effects for muscle-strengthening exercise, aerobic exercise, interferential treatment, and insoles only marginally crossed the line of no effect.

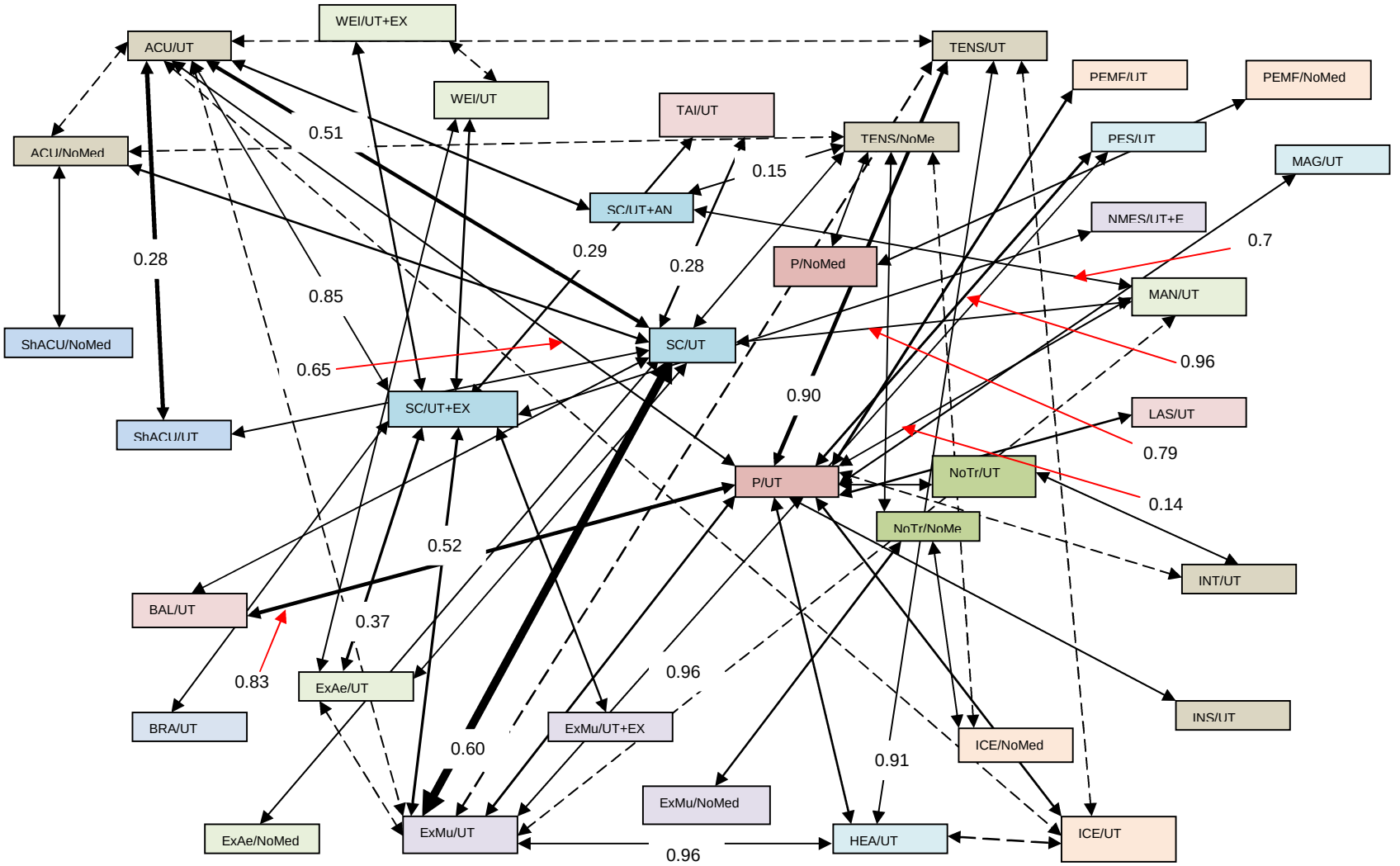


Figure 29: End of treatment analysis/any-quality trials/therapy-plus-adjunct interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

Table 27: Reduction in pain compared to standard care/usual treatment: end of treatment, therapy-plus-adjunct intervention set, all-quality studies.

Comparator: Standard care/usual care (23 trials)

Treatment	Trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	Trials	SMD	(95% CrI)	Median rank (CrI)
ExAe/NoMed	1	-4.29	(-5.88 to -2.71)	1 (1,2)	Bal/UT	6	-0.58	(-1.06 to -0.09)	20 (10,30)
Int/UT	1	-1.92	(-3.80 to -0.04)	3 (1,31)	Wei/UT+EX	2	-0.55	(-1.21 to 0.09)	21 (7,33)
ExMu/NoMed	2	-1.58	(-3.23 to 0.04)	4 (2,30)	P/NoMed	2	-0.54	(-1.76 to 0.67)	21 (5,34)
ShAcu/NoMed	1	-1.31	(-2.73 to 0.10)	6 (2,33)	ExAe/UT	5	-0.54	(-1.01 to -0.06)	21 (10,31)
Acu/NoMed	3	-1.26	(-2.30 to -0.23)	7 (2,28)	Wei/UT	2	-0.53	(-1.15 to 0.09)	21 (8,33)
TENS/NoMed	4	-1.19	(-2.06 to -0.32)	8 (3,25)	ShAcu/UT	6	-0.52	(-0.96 to -0.07)	21 (11,32)
Ins/UT	1	-1.08	(-2.15 to 0.00)	9 (2,31)	Bra/UT	1	-0.45	(-1.3 to 0.39)	23 (6,34)
Mag/UT	1	-1.07	(-2.07 to -0.07)	9 (2,30)	PEMF/NoMed	1	-0.43	(-1.85 to 0.97)	24 (4,35)
SC/UT+AN	4	-1.04	(-1.62 to -0.46)	10 (4,22)	NoTr/NoMed	3	-0.43	(-1.92 to 1.05)	24 (4,35)
Acu/UT	14	-1.03	(-1.34 to -0.72)	10 (4,18)	ExMu/UT	20	-0.43	(-0.65 to -0.21)	24 (16,30)
TENS/UT	6	-0.94	(-1.49 to -0.40)	12 (4,23)	Tai/UT	3	-0.37	(-0.86 to 0.11)	25 (12,34)
PES/UT	4	-0.93	(-1.51 to -0.34)	12 (4,25)	PEMF/UT	3	-0.33	(-0.96 to 0.31)	27 (13,34)
Las/UT	2	-0.88	(-1.65 to -0.13)	13 (4,29)	SC/UT+EX	12	-0.30	(-0.70 to 0.09)	27 (18,33)
Ice/NoMed	1	-0.87	(-2.68 to 0.94)	14 (2,35)	Hea/UT	4	-0.29	(-0.85 to 0.28)	27 (15,34)
Man/UT	3	-0.83	(-1.44 to -0.22)	14 (5,28)	P/UT	25	-0.25	(-0.66 to 0.17)	28 (20,33)
ExMu/UT+EX	2	-0.72	(-1.37 to -0.07)	17 (5,31)	NoTr/UT	1	-0.07	(-1.55 to 1.40)	31 (5,35)
Ice/UT	2	-0.66	(-1.49 to 0.18)	18 (5,33)	NMES/UT+EX	1	0.16	(-0.87 to 1.18)	33 (14,35)

Data points: 94 Residual deviance: 102 % deviance difference: 8.5% τ (between study standard deviation): 0.32 (95%CrI:0.21 to 0.46)

SMD Standardised mean difference; CrI Credible intervals. +EX with home exercises/education Results in order of mean treatment effectiveness

Acupuncture with usual care showed a statistically significant treatment benefit over placebo, muscle-strengthening exercise with usual treatment, heat treatment with usual care, NMES with usual care, PEMF with usual care, Tai Chi with usual care and sham acupuncture with usual care. However, some of these results lacked face validity. For example, acupuncture did not show a statistically significant treatment benefit over muscle-strengthening exercise, PEMF or sham acupuncture without usual care. Full results are at http://www.york.ac.uk/inst/crd/Documents/FullResultsPain.xlsm

The comparisons of the impact of the adjuncts for all the interventions with more than one adjunct in the network are presented in the caterpillar plot in Figure 30. There was no evidence of a difference in the adjuncts for most of the interventions. This suggests a lack of power in distinguishing between these treatment effects. Aerobic exercise with no medication was statistically significantly more effective than aerobic exercise with treatment as usual at a 95% level of credibility, but this lacks face validity. Standard care with usual treatment and analgesia had a statistically significant treatment benefit over standard care with usual treatment, and over standard care with exercise, at a 95% level of credibility.

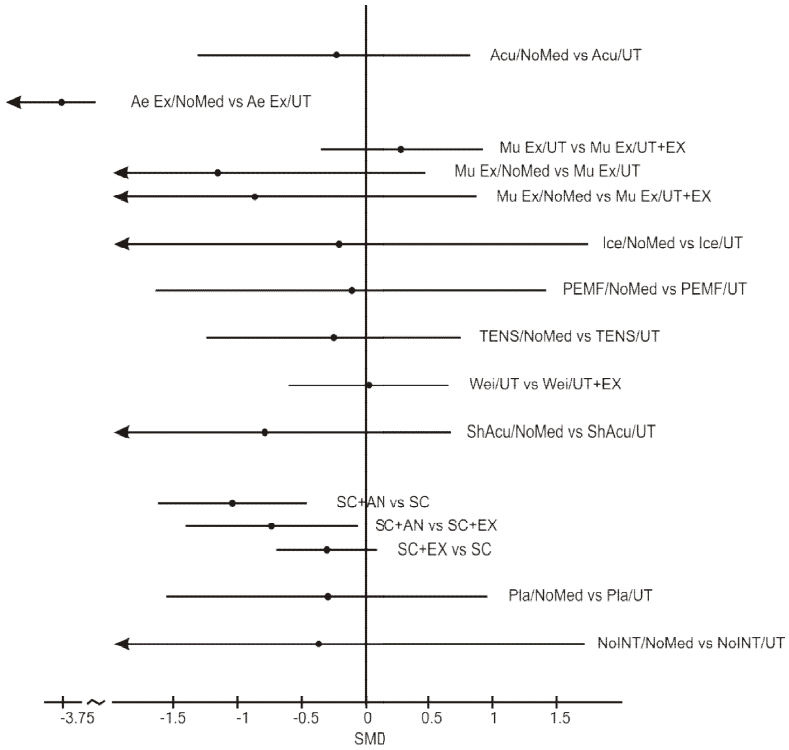


Figure 30: Caterpillar plot evaluating the impact of adjuncts on the outcomes from interventions.

Adjuncts: UT='treatment as usual', T+EX/ED= 'treatment as usual' plus specified home exercise or education, T+AN='treatment as usual' plus specified analgesics, NoMed = no medication
Comparators: Pla=placebo, SC=Standard Care, ShAcu=Sham acupuncture, NoINT=No intervention
Interventions: Acu=acupuncture, Ae Ex= Aerobic exercise, Mu Ex = muscle-strengthening exercise, Ice=Ice/cooling treatment, Wei = Weight loss intervention

6.3.1.2 *Sensitivity analysis: Higher-quality studies only*

When only higher-quality studies were included in a sensitivity analysis to evaluate the impact of excluding poor quality studies, only 13 out of the 35 interventions in the main analysis formed part of a connected network with acupuncture and the evidence was informed by 17 trials. The network is presented in Figure 31. Stable estimates were produced by the model for the SMDs and for the between-study SD of the random effects distribution, - the mean between-study SD across the comparisons was 0.22 SMD, a slightly higher estimate than in the analysis including poor quality

studies, which may reflect fewer studies informing the estimate. There was no strong evidence of inconsistency between the direct and indirect evidence. The consistency results are shown in Figure 31.

The estimates of effect of the included interventions are similar to those in the analysis of any-quality trials (Table 28). The change in the ranks and the rank uncertainty reflects the smaller set of interventions and different effect estimates. PES, acupuncture, balneotherapy, sham acupuncture, and muscle-strengthening exercise all showed a treatment benefit over standard care with the 95% credible intervals not crossing the line of no effect, regardless of the adjunctive treatment. The 95% credible intervals of the treatment effect of Tai Chi only marginally crossed the line of no effect.

Acupuncture with usual care showed a statistically significant treatment benefit over muscle-strengthening exercise with usual treatment and sham acupuncture with usual care. http://www.york.ac.uk/inst/crd/Documents/FullResultsPain.xlsm

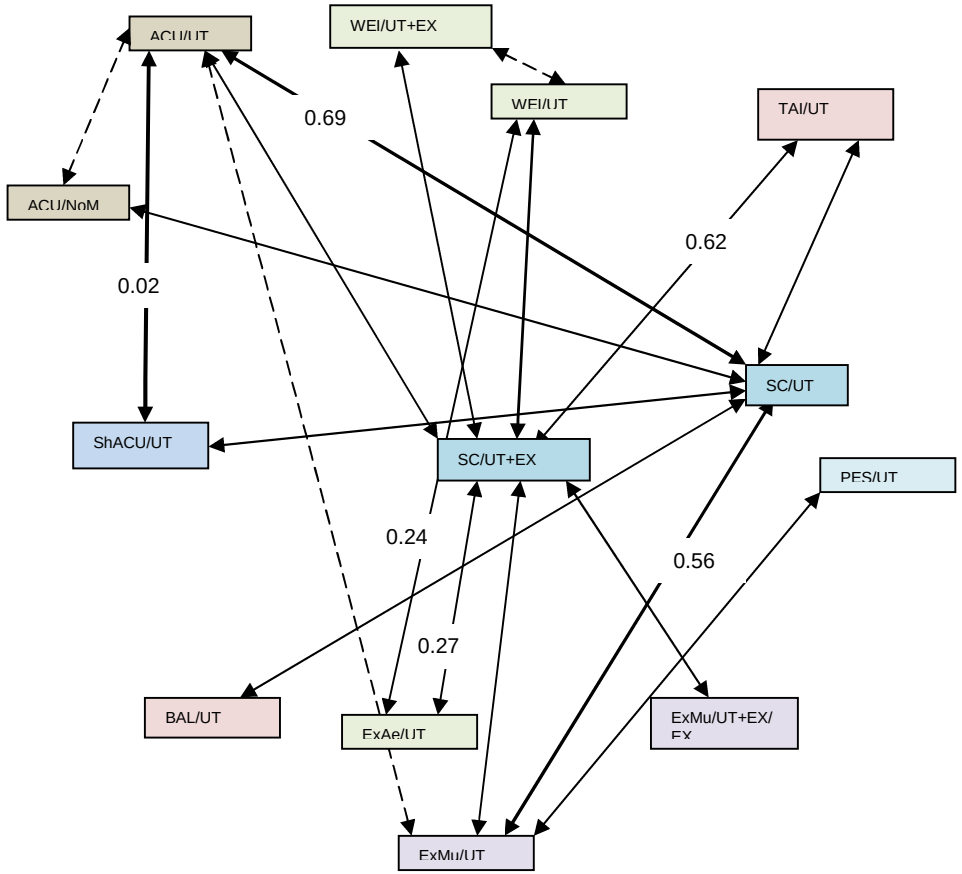


Figure 31: End of treatment analysis/higher-quality trials/therapy-plus-adjunct interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency.

The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

Table 28: Reduction in pain compared to standard care/usual treatment: end of treatment analysis, therapy-plus-adjunct intervention set, higher-quality studies only.

Comparator: Standard care/usual care (7 trials)

Treatment	Trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	Trials	SMD	(95% CrI)	Median rank (CrI)
PES/UT	1	-1.70	(-2.52 to -0.88)	1 (1,5)	Wei/UT	2	-0.59	(-1.24 to 0.04)	8 (3,12)
Acu/NoMed	1	-1.45	(-2.75 to -0.15)	2 (1,12)	SC/UT+EX	5	-0.56	(-1.05 to -0.07)	8 (5,11)
Acu/UT	7	-1.07	(-1.38 to -0.75)	4 (2,7)	Tai/UT	2	-0.55	(-1.15 to 0.03)	9 (3,12)
Bal/UT	1	-1.01	(-1.69 to -0.37)	4 (1,11)	ExMu/UT	5	-0.51	(-0.84 to -0.17)	9 (5,12)
ExMu/UT+EX	1	-0.96	(-1.69 to -0.22)	4 (2,10)	ExAe/UT	1	-0.46	(-1.17 to 0.25)	10 (4,13)
ShAcu/UT	4	-0.66	(-1.07 to -0.24)	7 (4,12)	Wei/UT+EX	1	-0.43	(-1.14 to 0.28)	10 (4,13)

Data points: 23 Residual deviance: 23.39 % deviance difference: 1.6% τ (between study standard deviation): 0.22 (95%CrI:0.01 to 0.49)

SMD Standardised mean difference; CrI Credible intervals. Results in order of mean treatment effectiveness

6.3.2 **Therapy-only interventions set**

6.3.2.1 *All studies (any-quality)*

All 22 interventions formed part of a connected network with acupuncture and the evidence was informed by 87 trials. The network is presented in Figure 32. Stable estimates were produced by the model for the SMDs and for the between-study SD of the random effects distribution, τ. The mean between-study SD across the comparisons was 0.43 SMD. There was evidence of a high level of inconsistency for the treatment effect estimates involving PES. This suggests that there is a possibility of bias or lack of exchangeability between the comparisons and hence the credible intervals of the treatment effect for PES at least may be underestimated.

Table 29 and Figure 35 present the treatment effects compared with standard care in order of mean effectiveness. The median effectiveness rank and the uncertainty in the rank are also presented. All of the interventions apart from PEMF, placebo, NMES and no intervention had a mean estimate that favoured the intervention. Interferential treatment, acupuncture, PES, TENS, aerobic exercise, and muscle-strengthening exercise all showed a treatment benefit over standard care with the 95% credible intervals not crossing the line of no effect. The 95% credible intervals for balneotherapy only marginally crossed the line of no effect. Table 30 and figure 36 present the results for each comparator versus acupuncture. The mean estimate favoured acupuncture for all comparisons apart from with Interferential treatment and static magnets. Acupuncture was significantly better at reducing pain than muscle-strengthening exercise, heat treatment, insoles, PEMF, NMES, placebo, and no intervention at a 95% level of credibility.

A few treatment effects changed significantly when between the Therapy-plus-adjunct interventions and Therapy-only Interventions networks, such as the treatment effect for insoles compared to standard care. In this case the change occurred because extra trials made it into the network as formerly disjointed networks became combined due to the broader definition of treatment groups. For example, in the Therapy-plus-adjunct interventions set, the treatment effect (SMD) for insoles was -1.06. This changed to -0.01 in the Therapy-only Intervention set, as two extra trials were incorporated.

Table 31 reports the results compared to standard care in terms of change in the WOMAC VAS 0-100 scale.

6.3.2.2 *Sensitivity analysis: excluding extreme data*

Four trials were excluded in a sensitivity analysis: two poor quality studies[81 124] on the basis of extreme data, one study of satisfactory quality [34] on the basis that the population had more severe OA of the knee, and one poor quality study [87] on the basis that the intervention was very different from all the others in its class. In general, the results were not sensitive to these changes, although the model fit improved as the percentage deviance difference reduced from 16% to 2.9%, and the between-study standard deviation was reduced. The consistency results are reported in Figure 33. The removal of the outliers affected the inconsistency values in some cases, but there was still strong evidence of inconsistency for the PES, placebo and muscle-strengthening exercise evidence triangle, suggesting that that particular inconsistency was not responsible for the difference in model fit. Table 29 presents the results versus standard care and Table 30 presents the results versus acupuncture. The median effectiveness rank and the uncertainty in the rank are also presented. Acupuncture remained more effective than sham acupuncture and muscle-strengthening exercise, but it was no longer statistically significantly better.

6.3.2.3 *Sensitivity analysis: Higher-quality trials only*

When only higher-quality studies were included in a sensitivity analysis to evaluate the impact of excluding poor quality studies, only 10 interventions formed part of a connected network with acupuncture and the evidence was informed by 19 trials (see Tables 29 and 30). The network is presented in Figure 34. Insoles, laser/light therapy and placebo comparators were not part of this analysis, as they did not form part of a connected network with acupuncture (even though their trials had adequate quality).

Stable estimates were produced by the model for the SMDs and for the between-study SD of the random effects distribution, τ. The mean between-study SD across the comparisons was 0.33 SMD, a slightly higher estimate than in the analysis including poor quality studies which may reflect fewer studies informing the estimate. The consistency results are presented in Figure 34. There was no strong evidence of inconsistency. This may be due either to fewer studies in the analysis and therefore lower power to identify inconsistency, or to the smaller number of evidence triangles in the smaller networks. The lack of evidence of inconsistency in these cases does not necessarily mean that the model accounts for all the differences in trial population and designs between different comparisons.

The estimates of effectiveness of the included interventions are similar to those in the main analysis, except PES and balneotherapy became more effective (and to a lesser extent acupuncture) compared with standard care, and aerobic exercise and weight loss became less effective.

Table 29 and Figure 37 present the treatment effects compared with standard care in order of mean effectiveness. The median effectiveness rank and the uncertainty in the rank is also presented. For the interventions remaining, the mean estimates favoured the intervention compared with standard care except for aerobic exercise and weight loss. PES, acupuncture, balneotherapy and muscle-strengthening exercise all showed a treatment benefit over standard care with the 95% credible intervals not crossing the line of no effect. The 95% credible intervals for sham acupuncture only marginally crossed the line of no effect.

Table 30 and Figure 38 present the results for each comparator versus acupuncture. Apart from the PES and balneotherapy comparisons, the mean estimate favoured acupuncture. Acupuncture was statistically significantly better at a 95% level of credibility than sham acupuncture, muscle-strengthening exercise, weight loss, aerobic exercise, and no intervention.

Table 31 reports the results compared to standard care in terms of change in the WOMAC VAS 0-100 scale.

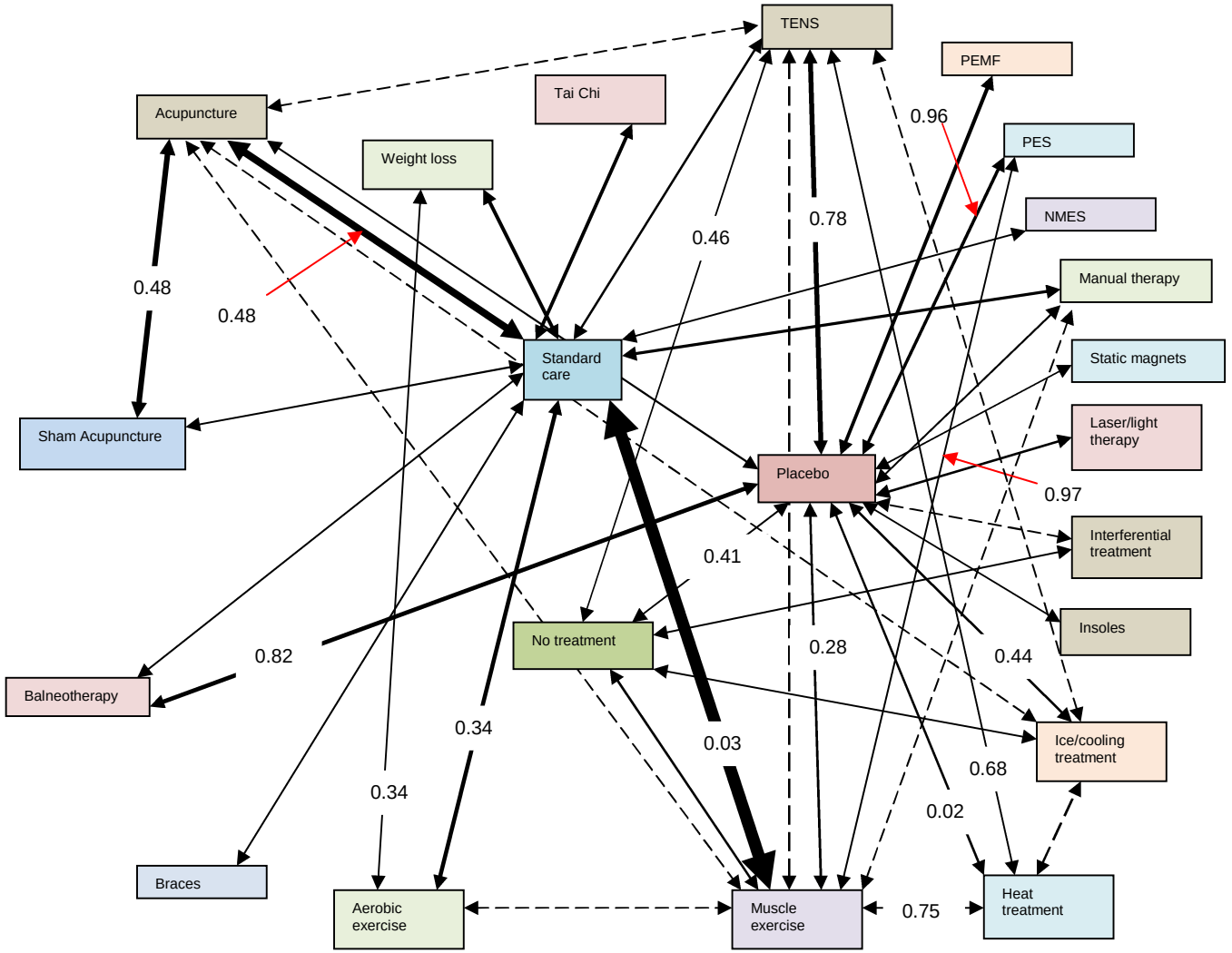


Figure 32: End of treatment analysis, any-quality trials, therapy-only intervention set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency.

The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

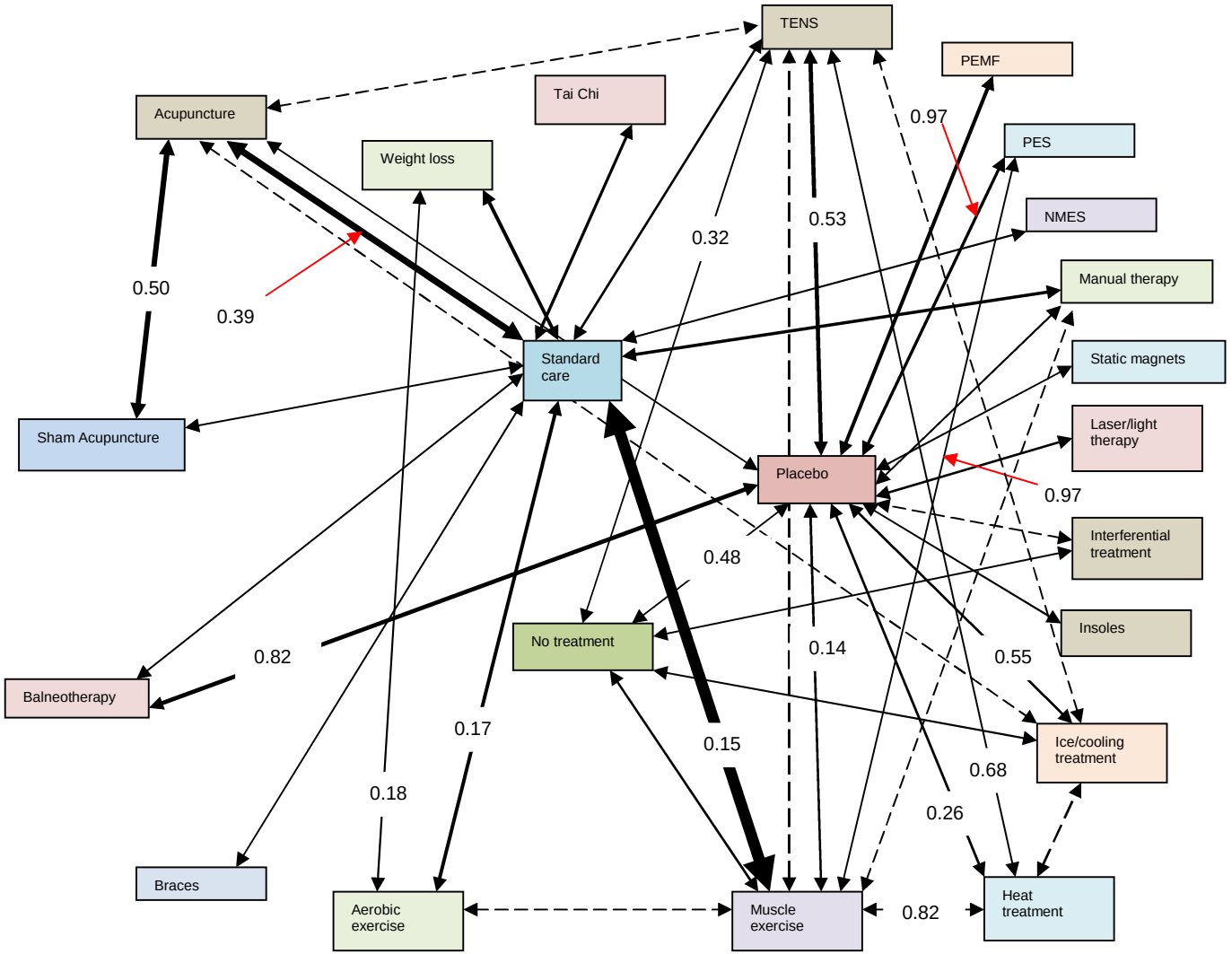


Figure 33: End of treatment analysis, any-quality trials (excluding outliers), therapy-only intervention set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency.

The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

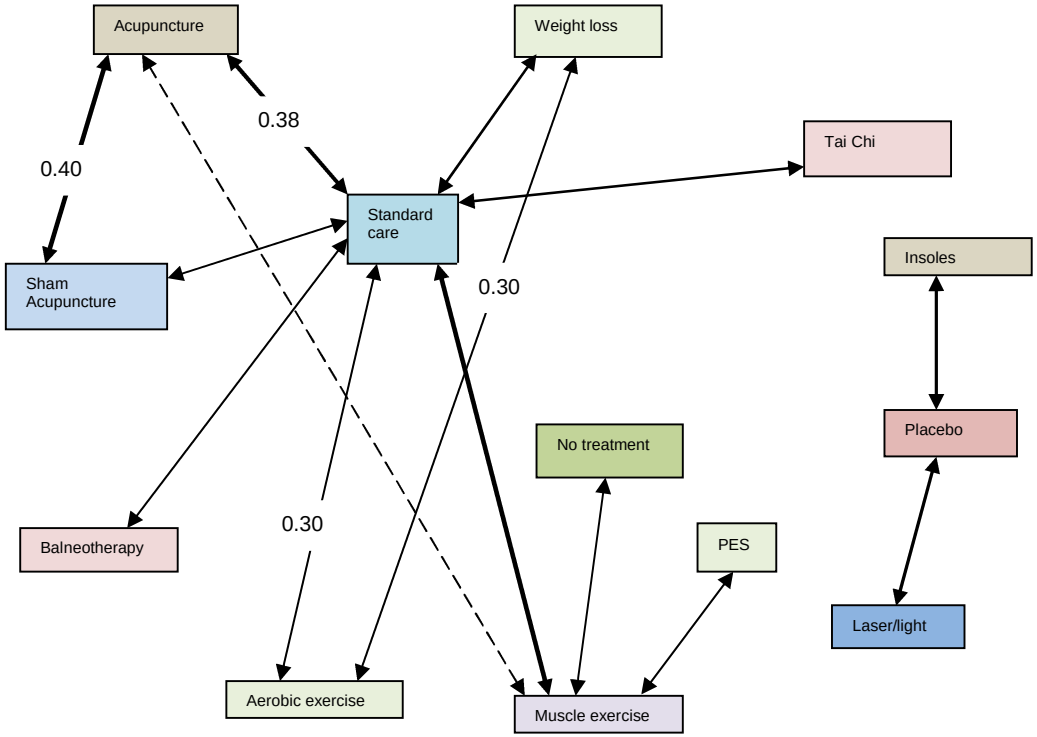


Figure 34: End of treatment analysis, higher-quality trials, therapy-only intervention set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

Table 29: Change in pain compared to standard care: end of treatment analysis, therapy-only intervention set (main and sensitivity analyses).

Intervention	Main analysis: Trials of any-quality				Sensitivity analysis (excluding extreme data)				Sensitivity analysis: Higher-quality trials only			
	Trials	SMD (95% Cr I)		Median rank (CrI)	Trials	SMD (95% Cr I)		Median rank (CrI)	Trials	SMD (95% Cr I)		Median rank (CrI)
Standard care (comparator)	39	-	-	17 (12,20)	36	-	-	17 (13,20)	12	-	-	8 (6,10)
Interferential (INT)	2	-1.06	(-2.05 to -0.07)	2 (1,14)	2	-1.11	(-2.02 to -0.21)	2 (1,12)		N/A		
Acupuncture (ACU)	18	-0.78	(-1.10 to -0.46)	3 (1,9)	17	-0.85	(-1.15 to -0.54)	3 (1,7)	8	-1.01	(-1.42 to -0.62)	2 (1,4)
Static magnets (MAG)	1	-0.78	(-1.92 to 0.36)	3 (1,20)	1	-0.87	(-1.88 to 0.14)	3 (1,18)		N/A		
PES (PES)	4	-0.70	(-1.35 to -0.05)	4 (1,14)	4	-0.77	(-1.34 to -0.19)	4 (1,12)	1	-1.57	(-2.56 to -0.57)	1 (1,4)
TENS (TENS)	11	-0.62	(-1.07 to -0.18)	6 (2,12)	10	-0.50	(-0.91 to -0.10)	7 (3,14)		N/A		
Aerobic exercise (AE EX)	5	-0.60	(-1.05 to -0.16)	6 (1,14)	4	-0.32	(-0.72 to 0.07)	11 (4,19)	1	0.12	(-0.61 to 0.85)	9 (4,10)
Balneotherapy (BAL)	8	-0.46	(-0.96 to 0.05)	8 (3,16)	8	-0.54	(-0.98 to -0.09)	7 (2,14)	1	-1.01	(-1.85 to -0.17)	2 (1,6)
Muscle exercise (MU EX)	25	-0.36	(-0.59 to -0.14)	10 (5,15)	24	-0.37	(-0.57 to -0.17)	10 (5,15)	8	-0.38	(-0.74 to -0.02)	5 (3,8)
Weight loss (WEI)	3	-0.36	(-0.89 to 0.16)	10 (2,19)	3	-0.29	(-0.74 to 0.16)	12 (4,20)	2	0.01	(-0.56 to 0.57)	8 (4,10)
Manual (MAN)	3	-0.30	(-0.91 to 0.32)	11 (2,20)	3	-0.33	(-0.87 to 0.22)	11 (3,20)		N/A		
Tai Chi (TAI)	3	-0.28	(-0.86 to 0.29)	12 (3,20)	2	-0.26	(-0.92 to 0.41)	12 (2,21)	2	-0.26	(-0.93 to 0.40)	6 (3,10)
Sham Acupuncture (SH ACU)	8	-0.26	(-0.73 to 0.21)	12 (4,20)	8	-0.33	(-0.75 to 0.10)	11 (4,19)	5	-0.47	(-0.98 to 0.05)	5 (3,8)
Laser (LAS)	3	-0.25	(-0.98 to 0.48)	12 (3,21)	3	-0.33	(-0.97 to 0.32)	11 (3,20)		N/A		
Ice/cooling (ICE)	3	-0.24	(-1.01 to 0.54)	12 (2,21)	3	-0.28	(-1.01 to 0.44)	12 (2,21)		N/A		
Braces (BRA)	1	-0.15	(-1.08 to 0.78)	14 (2,22)	1	-0.15	(-0.94 to 0.64)	14 (2,22)		N/A		
Heat treatment (HEA)	4	-0.04	(-0.66 to 0.57)	16 (6,21)	4	-0.06	(-0.60 to 0.49)	16 (7,21)		N/A		
Insoles (INS)	3	-0.01	(-0.72 to 0.69)	16 (5,22)	3	-0.07	(-0.69 to 0.54)	16 (6,21)		N/A		
PEMF (PEMF)	4	0.01	(-0.63 to 0.64)	17 (7,21)	4	-0.07	(-0.62 to 0.48)	16 (7,21)		N/A		
Placebo (PLA)	33	0.04	(-0.36 to 0.45)	17 (13,20)	32	-0.05	(-0.41 to 0.32)	17 (12,20)		N/A		
NMES (NMES)	1	0.46	(-0.64 to 1.55)	21 (5,22)	1	0.45	(-0.52 to 1.43)	21 (7,22)		N/A		
No intervention (NO INT)	4	0.59	(-0.03 to 1.22)	21 (17,22)	4	0.56	(0.01 to 1.13)	22 (18,22)	1	0.33	(-0.58 to 1.24)	10 (4,10)
	Data points: 100 Residual deviance: 116 %dd: 16% Between-study standard deviation: 0.43 (95% CrI:0.32 to 0.56)				Data points: 95 Residual deviance: 97.78 %dd: 2.9 % Between-study standard deviation: 0.35 (95% CrI:0.25 to 0.47)				Data points: 23 Residual deviance: 22.83 %dd: -0.7% Between-study standard deviation: 0.33 (95% CrI:0.15 to 0.60)			

SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. Results in order of mean treatment effectiveness

Table 30: Change in pain compared to acupuncture: end of treatment analysis, therapy-only intervention set (main analysis and sensitivity analyses)

Intervention	Main analysis: Trials of any-quality			Sensitivity analysis excluding extreme data			Sensitivity analysis: Higher-quality trials			
	No. Trials	SMD (95% Cr I)		Median rank (CrI)	No. Trials	SMD (95% Cr I)		Median rank (CrI)	No. Trials	SMD (95% Cr I)		Median rank (CrI)
Acupuncture (comparator)	18	-	-	3 (1,9)	17	-	-	3 (1,7)	8	-	-	2 (1,4)
Interferential (INT)	2	-0.28	(-1.31 to 0.75)	2 (1,14)	2	-0.26	(-1.20 to 0.68)	2 (1,12)		NA		
Static magnets (MAG)	1	0.00	(-1.17 to 1.17)	3 (1,20)	1	-0.02	(-1.07 to 1.01)	3 (1,18)		NA		
PES (PES)	4	0.08	(-0.62 to 0.78)	4 (1,14)	4	0.08	(-0.55 to 0.71)	4 (1,12)	1	-0.56	(-1.61 to 0.50)	1 (1,4)
TENS (TENS)	11	0.16	(-0.35 to 0.67)	6 (2,12)	10	0.31	(-0.21 to 0.83)	7 (3,14)		NA		
Aerobic exercise (AE EX)	5	0.18	(-0.38 to 0.72)	6 (1,14)	4	0.34	(-0.13 to 0.81)	11 (4,19)	1	1.14	(0.30 to 1.97)	9 (4,10)
Balneotherapy (BAL)	8	0.32	(-0.25 to 0.90)	8 (3,16)	8	0.48	(0.12 to 0.83)	7 (2,14)	1	0.00	(-0.93 to 0.95)	2 (1,6)
Muscle exercise (MU EX)	24	0.42	(0.04 to 0.79)	10 (5,15)	24	0.52	(-0.17 to 1.21)	10 (5,15)	8	0.64	(0.13 to 1.15)	5 (3,8)
Weight loss (WEI)	3	0.42	(-0.21 to 1.04)	10 (2,19)	3	0.52	(-0.10 to 1.14)	12 (4,20)	2	1.02	(0.33 to 1.72)	8 (4,10)
Manual therapy (MAN)	3	0.48	(-0.21 to 1.17)	11 (2,20)	3	0.53	(0.03 to 1.02)	11 (3,20)		NA		
Tai Chi (TAI)	3	0.50	(-0.17 to 1.16)	12 (3,20)	3	0.59	(-0.14 to 1.32)	12 (2,21)	2	0.75	(-0.01 to 1.53)	6 (3,10)
Sham Acupuncture (SH ACU)	8	0.52	(0.15 to 0.89)	12 (4,20)	8	0.56	(-0.19 to 1.33)	11 (4,19)	5	0.55	(0.16 to 0.94)	5 (3,8)
Laser (LAS)	3	0.53	(-0.24 to 1.31)	12 (3,21)	3	0.56	(0.01 to 1.09)	11 (3,20)		NA		
Ice/cooling (ICE)	3	0.54	(-0.26 to 1.35)	12 (2,21)	3	0.52	(0.20 to 0.85)	12 (2,21)		NA		
Braces (BRA)	1	0.63	(-0.36 to 1.62)	14 (2,22)	1	0.70	(-0.16 to 1.54)	14 (2,22)		NA		
Heat treatment (HEA)	4	0.74	(0.07 to 1.41)	16 (6,21)	4	0.78	(0.11 to 1.43)	16 (7,21)		NA		
Insoles (INS)	3	0.77	(0.01 to 1.52)	16 (5,22)	3	0.78	(0.16 to 1.38)	16 (6,21)		NA		
Standard care (ST CARE)	39	0.78	(0.46 to 1.10)	17 (12,20)	37	0.80	(0.36 to 1.24)	17 (13,20)		NA		8 (6,10)
PEMF (PEMF)	4	0.79	(0.10 to 1.47)	17 (7,21)	4	0.79	(0.18 to 1.39)	16 (7,21)	13	1.01	(0.62 to 1.42)	
Placebo (PLA)	33	0.82	(0.34 to 1.30)	17 (13,20)	32	0.85	(0.54 to 1.15)	17 (12,20)		NA		
NMES (NMES)	1	1.24	(0.09 to 2.38)	21 (5,22)	1	1.30	(0.28 to 2.32)	21 (7,22)		NA		
No intervention (NO INT)	4	1.37	(0.69 to 2.05)	21 (17,22)	4	1.41	(0.79 to 2.04)	22 (18,22)	1	1.35	(0.37 to 2.33)	10 (4,10)
	Data points: 100 Residual deviance: 116 %dd: 16% Between-study standard deviation: 0.43 (95% CrI:0.32 to 0.56)				Data points: 95 Residual deviance: 97.78 %dd: 2.9% Between-study standard deviation: 0.35 (95% CrI:0.25 to 0.47)				Data points: 23 Residual deviance: 22.83 %dd: -0.7% Between-study standard deviation: 0.33 (95% CrI:0.15 to 0.60)			

SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. Results in order of mean effectiveness

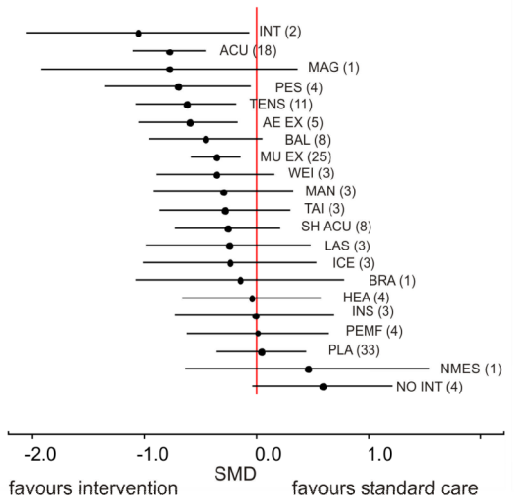


Figure 35: SMDs of each treatment compared to standard care for the analysis including studies of any-quality (no of studies in brackets).

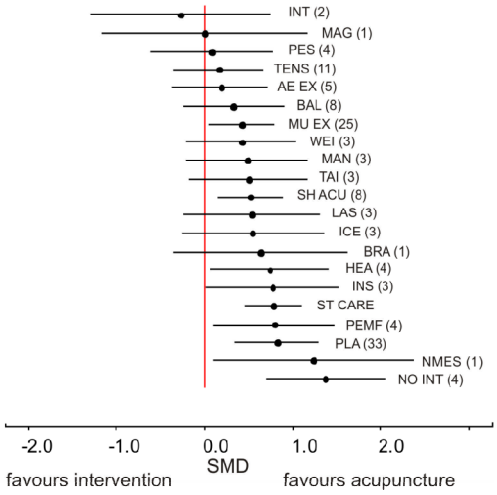


Figure 36: SMDs of each treatment compared to acupuncture for the analysis including studies of any-quality (no of studies in brackets).

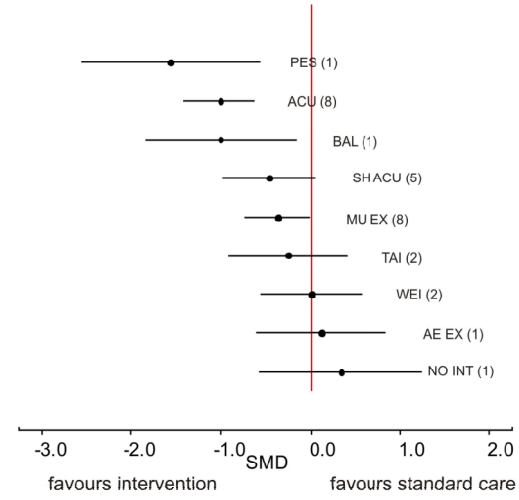


Figure 37: SMDs of each treatment compared to standard care for the analysis including higher-quality studies (no of studies in brackets).

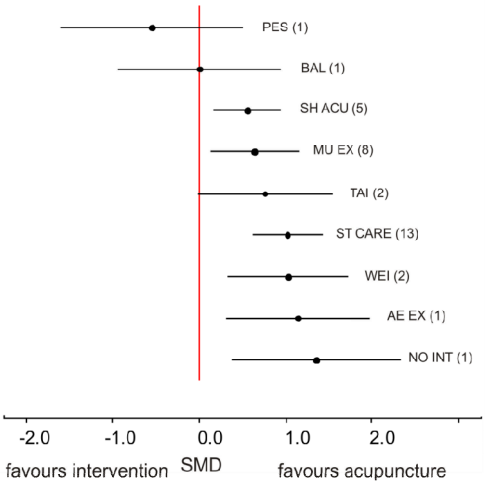


Figure 38: SMDs of each treatment compared to acupuncture for the analysis including higher-quality studies (no of studies in brackets).

Number of studies in brackets. ACU=acupuncture, AE EX= Aerobic exercise, BAL=Balneotherapy, BRA=Braces, HEA=Heat treatment, ICE=Ice/cooling treatment, INS=Insoles, INT=Interferential therapy, LAS=Laser therapy, MAG=Static magnets, MAN=Manual therapy, MU EX = muscle-strengthening exercise, NMES=Neuromuscular electrical stimulation, NO INT=No intervention, PEMF=pulsed electromagnetic fields, PES=Pulsed electrical stimulation, PLA=placebo, ST CARE=Standard care, SH ACU=Sham acupuncture, TAI=Tai Chi, TENS=Transcutaneous electrical nerve stimulation, WEI = Weight loss intervention

Table 31: WOMAC pain score difference results (back-transformed from SMDs) of network meta-analyses for comparisons with standard care

Intervention	Trials of any-quality			Higher-quality trials		
	No. of Trials	Change in WOMAC VAS 0-100 pain (95% Cr I)		No. of Trials	Change in WOMAC VAS 0-100 pain (95% Cr I)	
Standard care (comparator)	39	-	-	12	-	-
Interferential therapy (INT)	2	-17.43	(-33.82 to -1.12)		N/A	
Acupuncture (ACU)	18	-12.87	(-18.19 to -7.52)	8	-16.69	(-23.42 to -10.26)
Static magnets (MAG)	1	-12.86	(-31.68 to 5.99)		N/A	
PES (PES)	4	-11.58	(-22.20 to -0.85)	1	-25.89	(-42.21 to -9.42)
TENS (TENS)	11	-10.25	(-17.69 to -2.91)		N/A	
Aerobic exercise (AE EX)	5	-9.83	(-17.38 to -2.66)	1	2.02	(-10.10 to 13.99)
Balneotherapy (BAL)	8	-7.55	(-15.89 to 0.84)	1	-16.65	(-30.52 to -2.75)
Muscle exercise (MU EX)	24	-6.00	(-9.69 to -2.34)	7	-6.21	(-12.27 to -0.33)
Weight loss (WEI)	3	-5.93	(-14.73 to 2.69)	2	0.15	(-9.24 to 9.40)
Manual therapy (MAN)	3	-4.89	(-15.07 to 5.29)		N/A	
Tai Chi (TAI)	3	-4.70	(-14.22 to 4.86)	2	-4.30	(-15.27 to 6.60)
Sham Acupuncture (SH ACU)	8	-4.30	(-11.97 to 3.45)	5	-7.67	(-16.19 to 0.75)
Laser therapy (LAS)	3	-4.09	(-16.18 to 7.94)		N/A	
Ice/cooling treatment (ICE)	3	-3.93	(-16.64 to 8.87)		N/A	
Braces (BRA)	1	-2.42	(-17.81 to 12.91)			
Heat treatment (HEA)	4	-0.71	(-10.90 to 9.40)		N/A	
Insoles (INS)	3	-0.16	(-11.87 to 11.41)		N/A	
PEMF (PEMF)	4	0.13	(-10.35 to 10.56)		N/A	
Placebo (PLA)	33	0.69	(-6.00 to 7.36)		N/A	
NMES (NMES)	1	7.51	(-10.52 to 25.54)		N/A	
No intervention (NO INT)	4	9.71	(-0.51 to 20.05)	1	5.52	(-9.53 to 20.51)

6.3.3 **Grouped interventions set**

The network diagrams for the any-quality studies and for the higher-quality studies are given in Figures 39 and 40). Due to the grouping of interventions under 'Physiotherapy' these networks include some trials that were unable to be included in the Therapy-only intervention set network. When only higher-quality studies were included, the changes to the direct and indirect evidence available to estimate certain treatment effects resulted in changes to the treatment effect estimates beyond that which would be expected given the credible intervals of the estimates. For example the effect size for placebo vs. standard care switched from +0.07 (-0.30 to 0.45) for the any-quality studies analysis to -1.72 (-2.91 to -0.54) for higher-quality studies analysis. This suggests that either study quality had an effect on the outcome for at least one of the comparisons involving these interventions or that there is a significant difference in treatment effect between treatments lumped together. This raises suspicions about the results for a few treatment effects in the higher-quality analysis. The data for the higher-quality analysis is presented in section 10.4.3 NMA Appendix Table 5. The benefit of a greater number of trials from including poor quality trials seems likely to outweigh any potential bias associated with poor quality bias, and there is no evidence that poor quality trials as defined in this study are biased. Furthermore, the results are plausible given the other analyses in this study. The results of the analyses of any-quality studies are set out below and presented in Table 32.

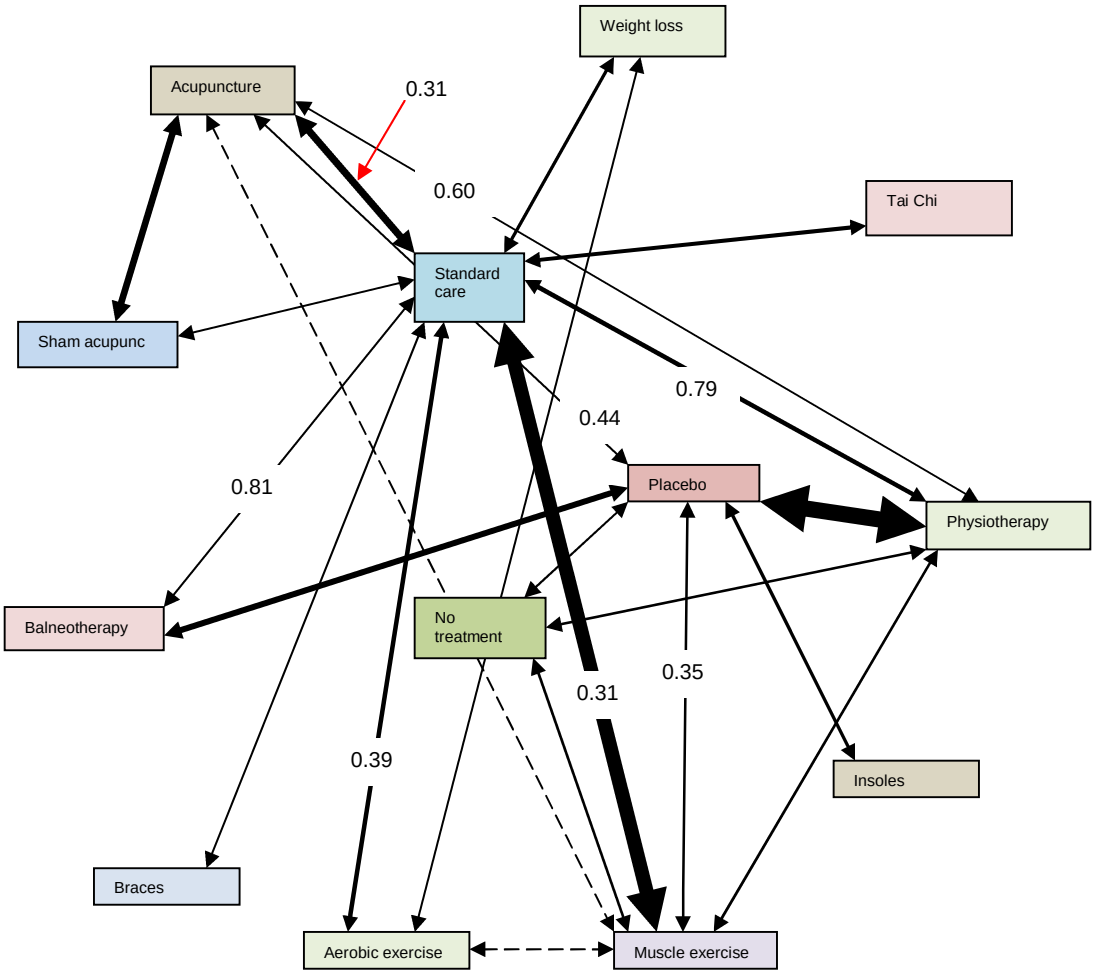


Figure 39: End of treatment analysis/any-quality trials/grouped interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

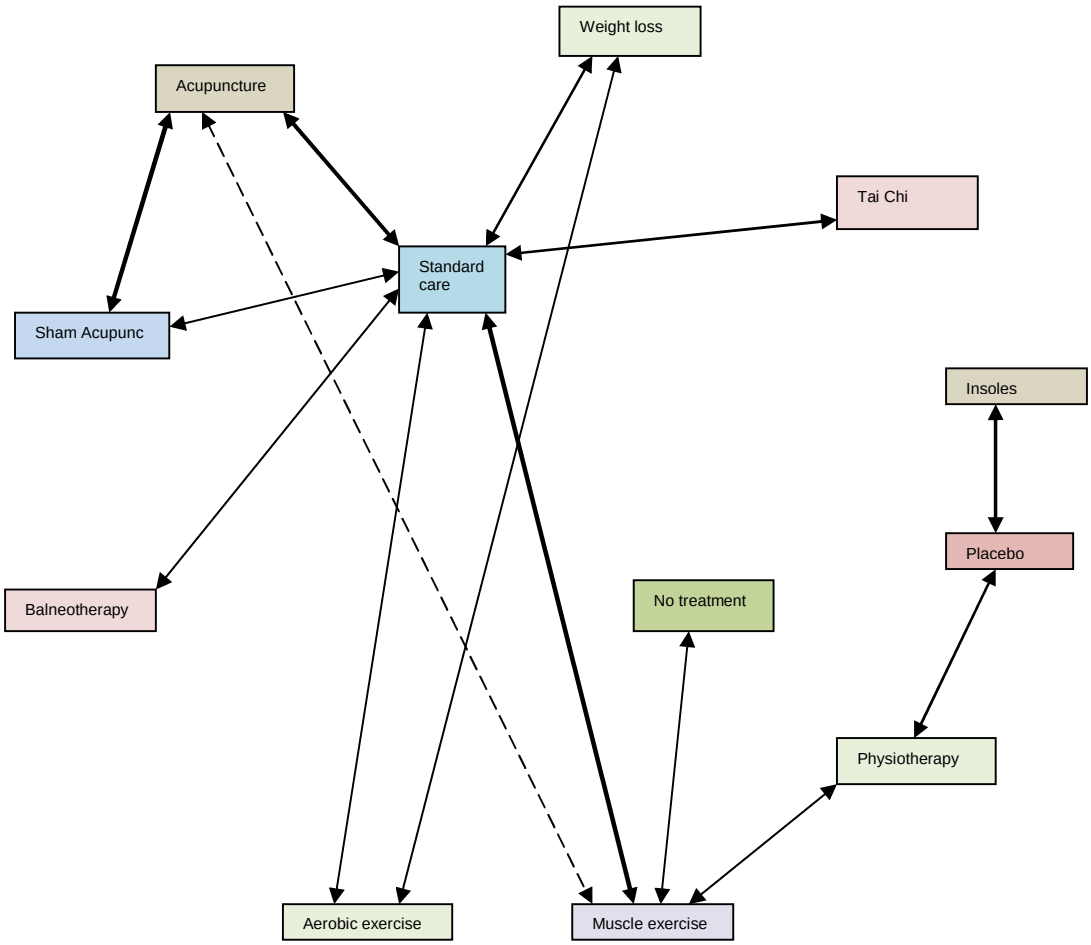


Figure 40: End of treatment analysis/higher-quality trials/grouped interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

6.3.3.1 *All studies (any-quality)*

All 13 interventions formed part of a connected network with acupuncture and the evidence was informed by 86 trials. The network is presented in Figure 39. Stable estimates were produced by the model for the SMDs and for the between-study SD of the random effects distribution, τ. The mean between-study SD across the comparisons was 0.46 SMD. In the network of all studies (any-quality trials) there was some evidence of inconsistency, but not complete inconsistency.

Table 32 presents the treatment effects compared with standard care in order of mean effectiveness. The median effectiveness rank and the uncertainty in the rank are also presented. All of the mean estimates favoured the intervention over standard care except for insoles, placebo and no intervention. Acupuncture, muscle-strengthening exercise, aerobic exercise, physiotherapy treatments and Tai Chi all showed a significant treatment benefit over standard care with the 95% credible intervals not crossing the line of no effect.

The analysis found that as well as showing a significant treatment benefit over standard care, acupuncture showing a significant treatment benefit over muscle-strengthening exercise (SMD -0.41, 95% CrI -0.80 to -0.02); insoles (SMD -0.78 , 95% CrI -1.53 to -0.01); and sham acupuncture (SMD -0.52 , 95% CrI -0.91 to -0.14); as well as placebo (SMD -0.83, 95% CrI -1.30 to -0.37) and no intervention (SMD -1.50, 95% CrI -2.21 to -0.81). Full results are given at http://www.york.ac.uk/inst/crd/Documents/FullResultsPain.xlsm

Table 32: Change in pain compared to standard care: end of treatment analysis, grouped intervention set (all-quality trial analyses)

Number of Interventions	13				
Number of trials	86				
Mean between study SD τ (SMD)	0.46 (95%CrI:0.35 to 0.58)				
Residual deviance (%dd)	112.7 (17.4%)				
Number of data points	96				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	Acupuncture	18	-0.76	(-1.09 to -0.43)	1 (1,5)
	Aerobic exercise	5	-0.61	(-1.09 to -0.15)	3 (1,8)
	Balneotherapy	8	-0.43	(-0.93 to 0.07)	5 (1,10)
	Weight loss	3	-0.36	(-0.92 to 0.19)	6 (1,12)
	Physiotherapy treatment	31	-0.35	(-0.70 to -0.01)	6 (2,9)
	Muscle exercise	24	-0.35	(-0.58 to -0.12)	6 (3,9)
	Tai Chi	3	-0.28	(-0.88 to 0.32)	7 (1,12)
	Sham Acupuncture	8	-0.24	(-0.72 to 0.25)	7 (2,12)
	Braces	1	-0.15	(-1.12 to 0.82)	8 (1,13)
	Insoles	3	0.01	(-0.70 to 0.72)	10 (2,13)
	Placebo	33	0.07	(-0.30 to 0.45)	11 (7,12)
	No intervention	4	0.74	(0.12 to 1.37)	13 (11,13)
	Standard care	39	0.00		10 (8,12)

SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. Results in order of mean treatment effectiveness

6.4 Pain - 3 months from the start of treatment

There were 75 trials with data that could be used in the 3 months from start of treatment analyses. The number of trials actually included in each analysis depended on how many of them formed part of a connected network with acupuncture.

There was considerable variation in the average treatment duration and the follow up time to the time point included in the analysis across the main interventions (see Appendix 10.4.3 Figure 2). In the Therapy-plus-adjunct interventions network of any-quality trials, for 42 out of 67 trials (63%) the 3 months from the start of treatment time point was the same as the end of treatment time point.

6.4.1 Analyses of the three networks: therapy-plus-adjunct intervention; therapy-only intervention; and grouped intervention

Details of the networks are given in Figures 41-46. Details of the analysis and results for all three levels of networks (Therapy-plus-adjunct Intervention, Therapy-only Intervention, and Grouped Intervention) using any-quality trials and the sensitivity analyses including only higher-quality trials are presented in Tables 33 and 34.

Across all three networks the analyses of all trials found consistent evidence of a significant beneficial effect of acupuncture, muscle-strengthening exercise and aerobic exercise over standard care. Across the Therapy-plus-adjunct Intervention and Therapy-only Intervention networks the analyses of all trials also provided evidence for a beneficial effect of interferential therapy, PES and balneotherapy. The beneficial effects of insoles and PEMF in the Therapy-plus-adjunct Intervention network, were not supported by the results from the other networks.

The sensitivity analyses which included only higher-quality trials generally reflected the main analyses, with the results for acupuncture remaining consistent across all analyses.

There was evidence from the Therapy-plus-adjunct and Therapy-only interventions networks that acupuncture was significantly better at reducing pain than muscle-strengthening exercise, heat treatment, insoles, PEMF, NMES, placebo, and no intervention at a 95% level of credibility. When several interventions, including PEMF and NMES, were grouped into the physiotherapy category, acupuncture did not show a statistically significant treatment effect over physiotherapy treatments at a 95% level of credibility. When only higher-quality studies were included, there was only one trial informing the physiotherapy estimate versus standard care. Full results are available http://www.york.ac.uk/inst/crd/Documents/FullResultsPain.xlsm

In general the results of the analysis of the 3 month from start of treatment end point reflect those from the equivalent End of Treatment analyses. In particular the results for acupuncture are consistent across the two time points. This is not surprising given that in many cases the two time points were the same.

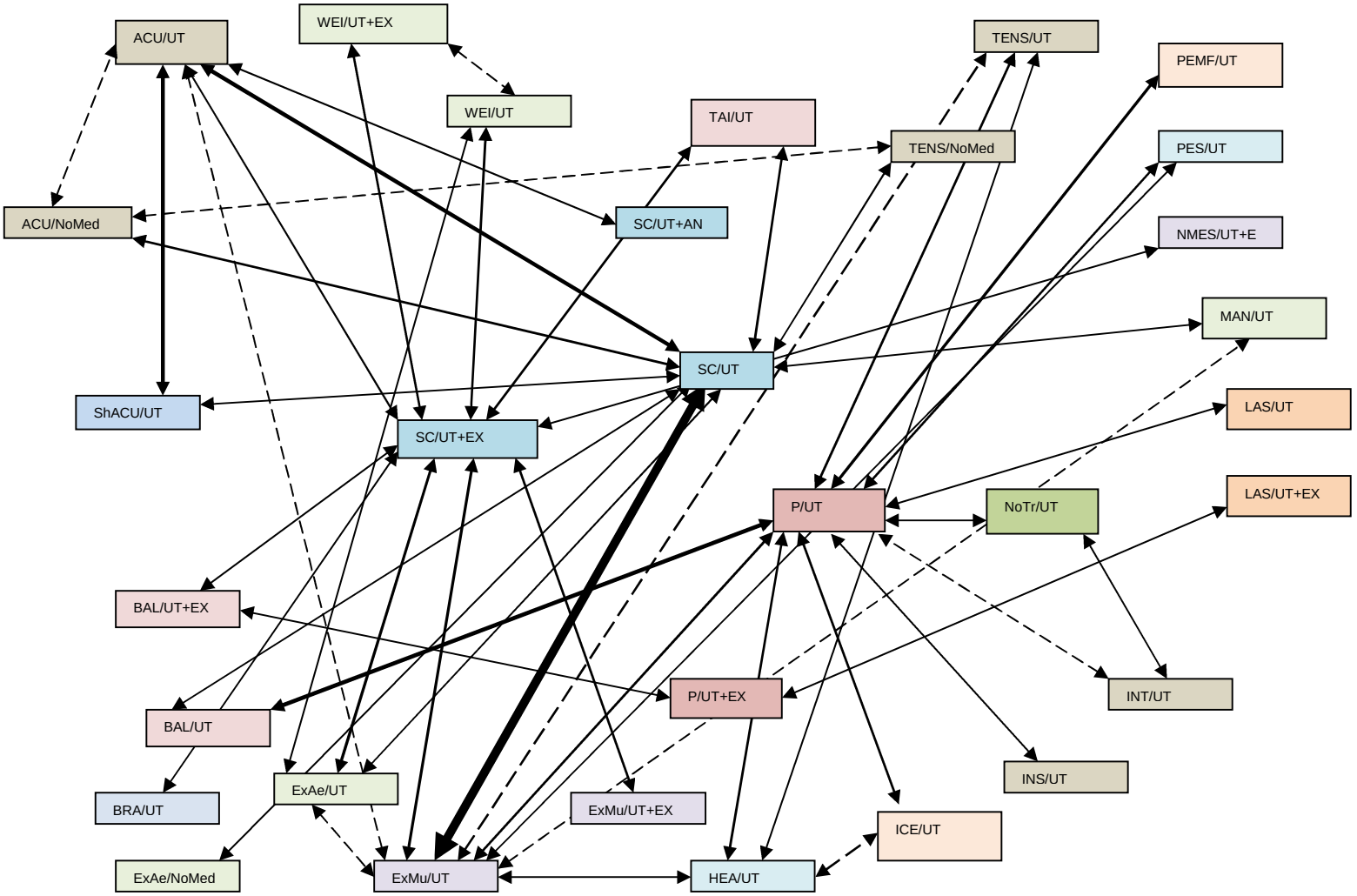


Figure 41: Three months from start of treatment analysis/all quality trials/therapy-plus-adjunct interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

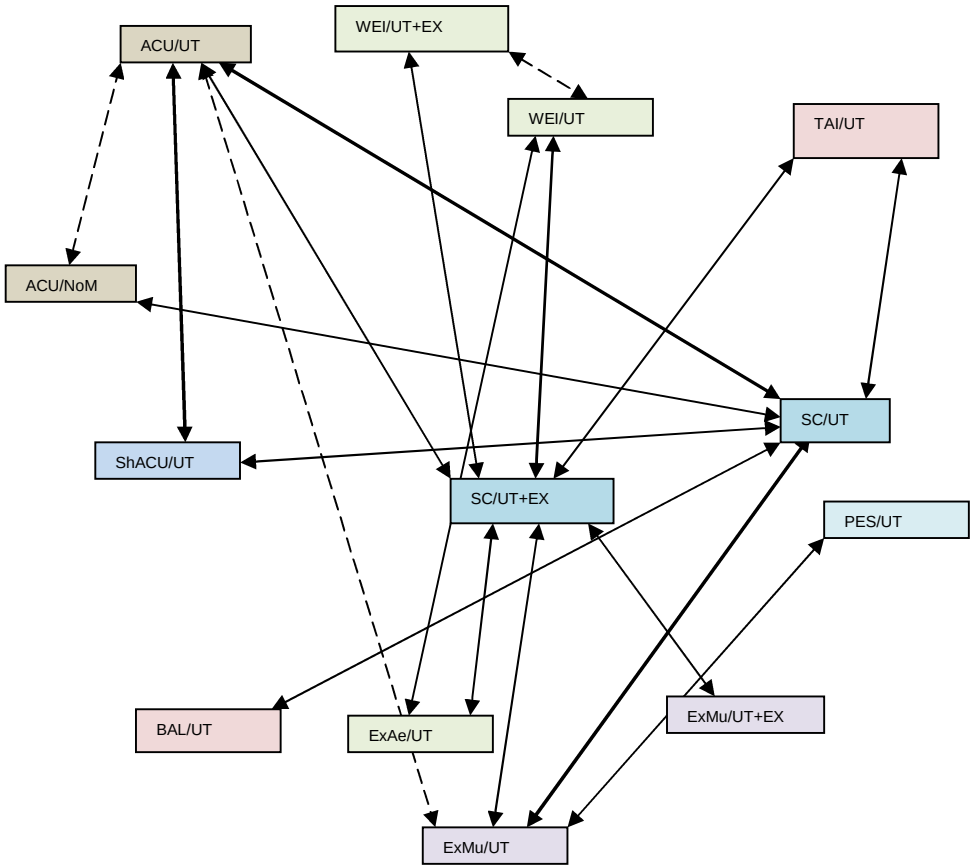


Figure 42: Three months from start of treatment analysis/higher-quality trials/therapy-plus-adjunct interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

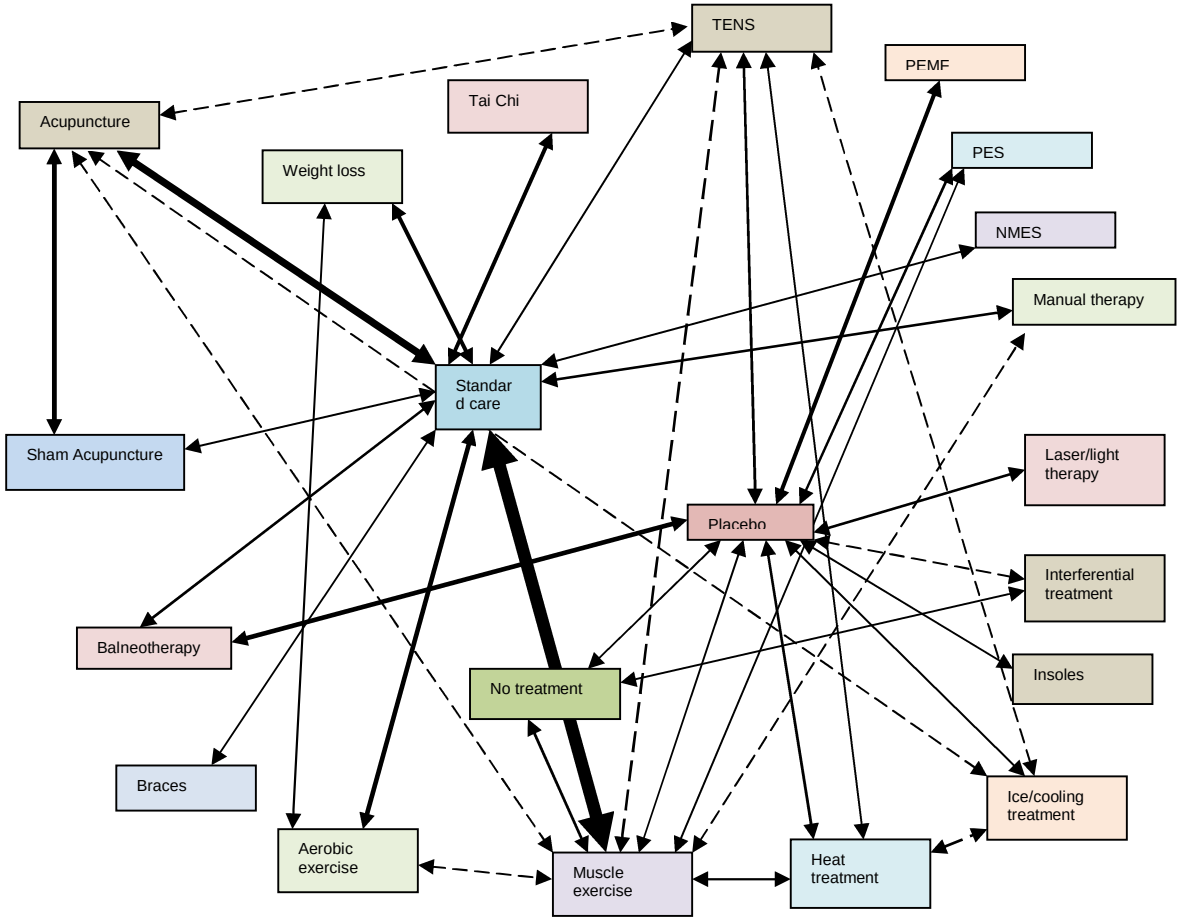


Figure 43: Three months from start of treatment analysis, any-quality trials, therapy-only intervention set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

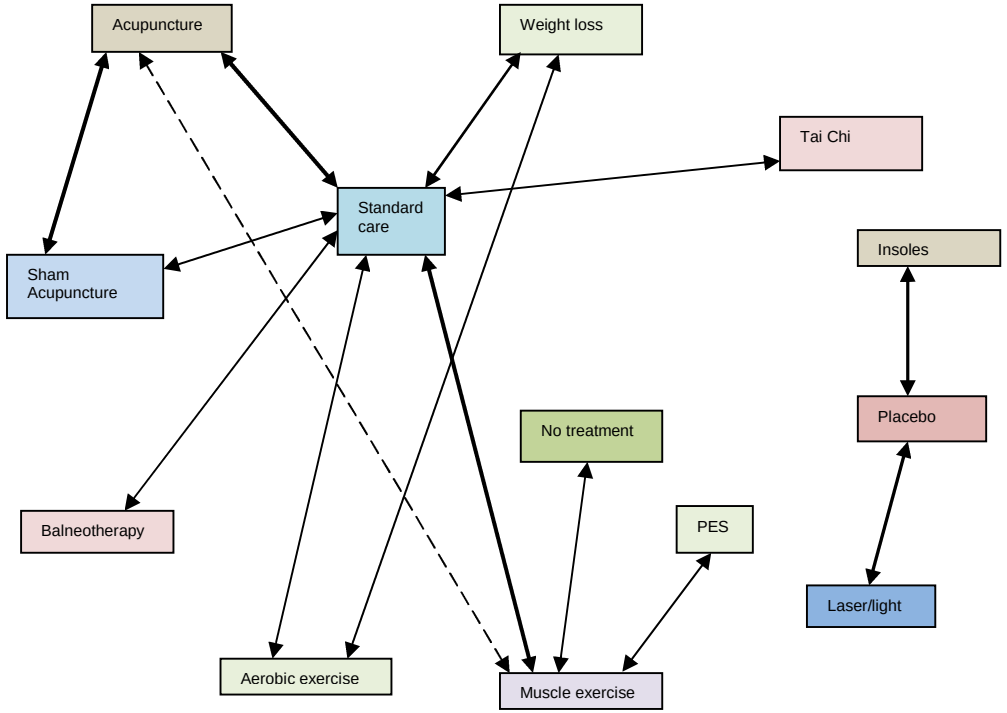


Figure 44: Three months from start of treatment analysis, higher-quality trials, therapy-only intervention set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

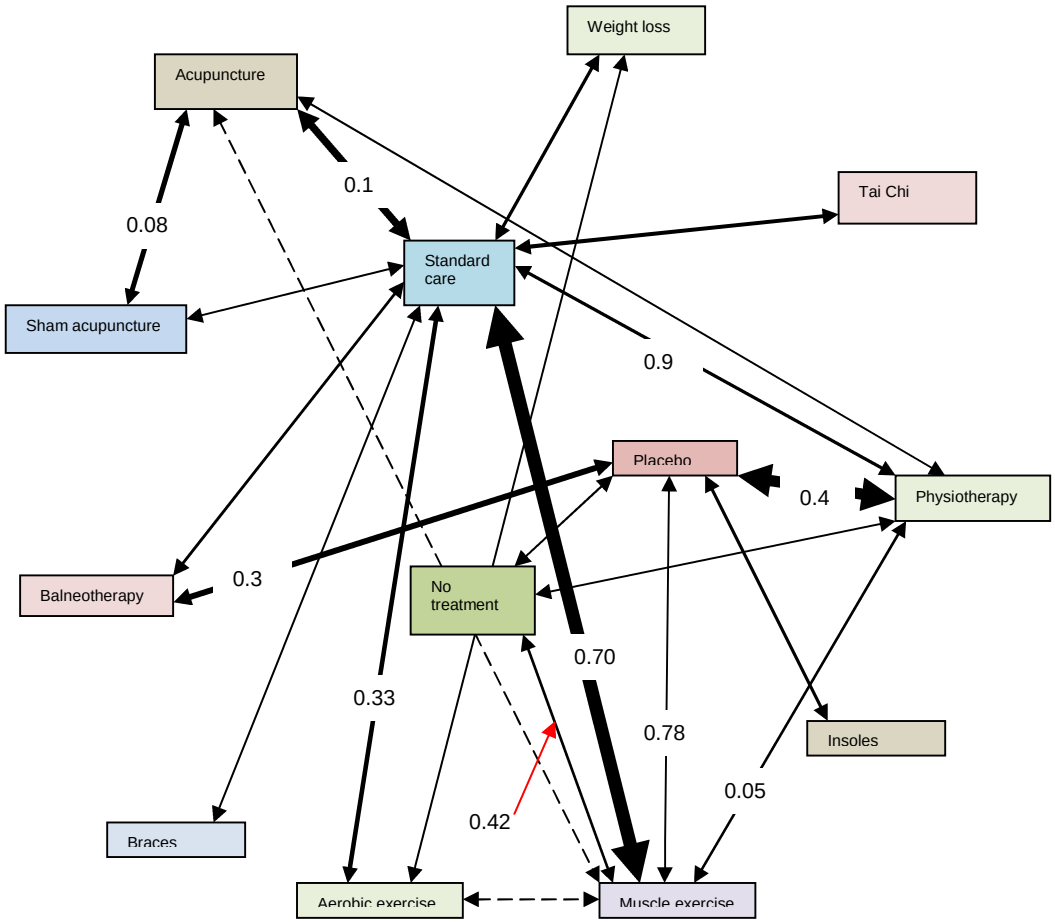


Figure 45: Three months from start of treatment analysis/any-quality trials/grouped interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

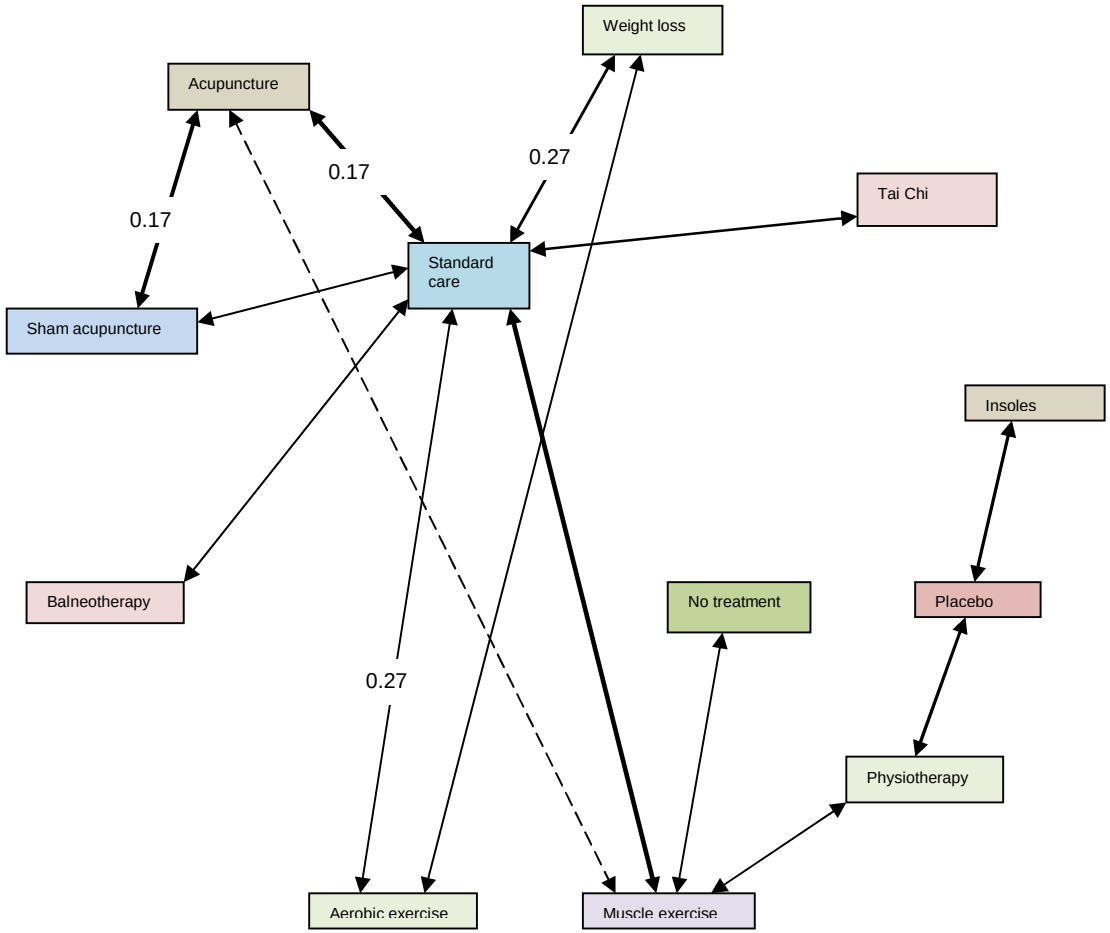


Figure 46: Three months from start of treatment analysis/higher-quality trials/grouped interventions set.

The numbers represent the inconsistency value for the direct and indirect estimates for the relevant comparison. 1 indicates complete inconsistency and 0 represents no inconsistency. The thickness of the lines represents the number of trials making the comparison. The dotted lines represent comparisons in 3-arm trials for which there is data point in the analyses.

Table 33: Change in pain compared to standard care: three months from start of treatment, all networks, any-quality analyses

	Therapy-plus-adjunct Intervention				**Therapy-only intervention**				**Grouped Intervention**				
Number of Intervention	31				21				13				
Number of trials	64				74				73				
Mean between study SD τ (SMD)	0.23 (95%CrI:0.12 to 0.36)				0.40 (95%CrI:0.29 to 0.53)				0.44 (95%CrI:0.33 to 0.58)				
Residual deviance (%dd)	76.91 (2.5%)				97.15 (17%)				95.39 (17.8%)				
Number of data points	75				83				81				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	ExAe/NoMed	(1)	-4.28	(-5.78 to -2.76)	1 (1,2)	Interferential	(2)	-1.45	(-2.42 to -0.47)	1 (1,8)	Acupuncture	(14)	-0.80	(-1.15 to -0.44)	1 (1,6)
	Int/UT	(1)	-2.30	(-4.12 to -0.49)	2 (1,16)	PES	(3)	-1.16	(-1.86 to -0.48)	2 (1,8)	Aerobic exercise	(5)	-0.60	(-1.07 to -0.16)	3 (1,10)
	Ins/UT	(1)	-1.46	(-2.46 to -0.46)	4 (2,17)	Acupuncture	(14)	-0.80	(-1.13 to -0.46)	4 (2,10)	Sham Acupuncture	(6)	-0.48	(-0.99 to 0.04)	5 (1,11)
	PES/UT	(3)	-1.31	(-1.89 to -0.73)	5 (2,11)	Ice/cooling treatment	(1)	-0.69	(-2.00 to 0.60)	6 (1,21)	Balneotherapy	(9)	-0.43	(-0.89 to 0.05)	5 (1,11)
	PEMF/UT	(3)	-1.14	(-1.78 to -0.50)	6 (3,15)	PEMF	(4)	-0.59	(-1.27 to 0.07)	7 (2,17)	Physiotherapy treatment	(4)	-0.40	(-0.81 to 0.00)	6 (2,10)
	SC/UT+AN	(1)	-1.13	(-1.95 to -0.30)	6 (2,22)	Aerobic exercise	(5)	-0.58	(-1.01 to -0.17)	7 (2,16)	Muscle exercise	(23)	-0.36	(-0.59 to -0.14)	7 (3,10)
	Ice/UT	(1)	-0.98	(-2.16 to 0.21)	8 (2,28)	Balneotherapy	(9)	-0.51	(-0.99 to -0.04)	8 (3,16)	Weight loss	(3)	-0.36	(-0.90 to 0.18)	7 (1,12)
	Acu/UT	(11)	-0.96	(-1.24 to -0.69)	8 (4,14)	Sham Acupuncture	(6)	-0.48	(-0.97 to 0.00)	9 (3,18)	Insoles	(3)	-0.29	(-1.03 to 0.44)	8 (1,13)
	Acu/NoMed	(2)	-0.94	(-1.89 to 0.03)	9 (2,27)	Insoles	(3)	-0.40	(-1.14 to 0.31)	11 (3,19)	Tai Chi	(3)	-0.28	(-0.88 to 0.31)	8 (1,12)
	Bal/UT	(6)	-0.85	(-1.33 to -0.36)	10 (5,19)	Muscle exercise	(23)	-0.40	(-0.61 to -0.18)	11 (6,16)	Placebo	(25)	-0.24	(-0.67 to 0.20)	9 (4,12)
	ShAcu/UT	(5)	-0.82	(-1.21 to -0.42)	11 (5,20)	Placebo	(25)	-0.36	(-0.83 to 0.11)	12 (6,17)	Braces	(1)	-0.15	(-1.11 to 0.81)	10 (1,13)
	ExMu/UT+EX	(2)	-0.65	(-1.17 to -0.13)	14 (6,24)	Weight loss	(3)	-0.35	(-0.86 to 0.14)	12 (4,19)	No intervention	(3)	0.46	(-0.23 to 1.18)	13 (9,13)
	P/UT	(17)	-0.62	(-1.09 to 0.14)	14 (9,24)	TENS	(5)	-0.34	(-0.96 to 0.28)	12 (4,19)					
	TENS/UT	(3)	-0.54	(-1.21 to -0.12)	16 (7,28)	Tai Chi	(3)	-0.28	(-0.83 to 0.26)	13 (4,20)	Standard care	(38)	0.00		11 (8,13)

	Therapy-plus-adjunct Intervention					Therapy-only intervention					Grouped Intervention	
	Las/UT	(1)	-0.53	(-1.45 to 0.10)	17 (5,30)	Manual	(1)	-0.24	(-1.40 to 0.92)	14 (1,21)		
	ExAe/UT	(4)	-0.49	(-0.89 to -0.09)	17 (10,25)	Braces	(1)	-0.15	(-1.03 to 0.73)	15 (3,21)		
	Wei/UT+EX	(2)	-0.46	(-1.00 to 0.06)	18 (8,27)	Laser	(3)	-0.14	(-0.88 to 0.59)	16 (5,21)		
	Wei/UT	(2)	-0.45	(-0.97 to 0.05)	18 (9,27)	Heat treatment	(4)	-0.11	(-0.74 to 0.51)	16 (6,21)		
	ExMu/UT	(19)	-0.44	(-0.62 to -0.26)	19 (13,24)	No intervention	(3)	0.21	(-0.47 to 0.91)	19 (10,21)		
	NoTr/UT	(1)	-0.44	(-1.86 to 0.98)	19 (3,31)	NMES	(1)	0.46	(-0.59 to 1.50)	21 (7,21)		
	Bra/UT	(1)	-0.39	(-1.07 to 0.30)	20 (7,29)							
	Tai/UT	(3)	-0.35	(-0.75 to 0.05)	21 (11,28)	Standard care	(38)	0.00		17 (13,20)		
	Hea/UT	(4)	-0.29	(-0.84 to 0.27)	22 (12,29)							
	Bal/UT+EX	(2)	-0.27	(-0.89 to 0.35)	23 (10,29)							
	Man/UT	(1)	-0.27	(-1.26 to 0.73)	23 (5,31)							
	SC/UT+EX	(13)	-0.24	(-0.55 to 0.08)	23 (18,27)							
	P/UT+EX	(2)	0.16	(-0.8 to 1.12)	28 (12,31)							
	TENS/NoMed	(1)	0.19	(-1.29 to 1.68)	28 (5,31)							
	NMES/UT+EX	(1)	0.22	(-0.68 to 1.12)	29 (14,31)							
	Las/UT+EX	(1)	0.38	(-0.79 to 1.54)	30 (12,31)							
	SC/UT	(24)	0.00		27 (23,30)							
SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. Results in order of mean treatment effectiveness For other abbreviations see Table 3												

Table 34: Change in pain compared to standard care: three months from start of treatment, all networks, higher-quality studies analyses

	Therapy-plus-adjunct Intervention					**Therapy-only intervention**				
Number of Intervention	13					10				
Number of trials	17					17				
Mean between study SD τ (SMD)	0.31 (95%CrI:0.03 to 0.63)					0.42 (95%CrI:0.20 to 0.77)				
Residual deviance (%dd)	22.94 (9.2%)					21.33 (6.7%)				
Number of data points	21					20				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	PES/UT	(1)	-1.71	(-2.69 to -0.80)	1 (1,5)	PES	(1)	-1.66	(-2.84 to -0.50)	1 (1,5)
	Acu/NoMed	(1)	-1.43	(-2.75 to -0.06)	2 (1,12)	Balneotherapy	(1)	-1.14	(-2.15 to -0.12)	2 (1,7)
	Bal/UT	(1)	-1.13	(-1.91 to -0.37)	3 (1,10)	Acupuncture	(7)	-0.96	(-1.45 to -0.48)	3 (1,5)
	Acu/UT	(6)	-0.97	(-1.36 to -0.58)	4 (2,8)	Sham Acupuncture	(4)	-0.67	(-1.31 to -0.03)	5 (2,8)
	ShAcu/UT	(3)	-0.92	(-1.44 to -0.40)	5 (2,10)	No intervention	(1)	-0.56	(-1.66 to 0.52)	5 (1,10)
	ExMu/UT+EX	(1)	-0.76	(-1.7 to 0.14)	6 (2,12)	Muscle exercise	(7)	-0.48	(-0.92 to -0.06)	6 (3,8)
	ExMu/UT	(5)	-0.52	(-0.95 to 0.17)	8 (4,12)	Tai Chi	(1)	-0.20	(-1.28 to 0.88)	7 (2,10)
	Tai/UT	(2)	-0.44	(-1.14 to 0.23)	8 (4,13)	Weight loss	(2)	0.00	(-0.70 to 0.69)	8 (4,10)
	Wei/UT	(2)	-0.40	(-1.22 to 0.37)	9 (4,13)	Aerobic exercise	(1)	0.12	(-0.79 to 1.01)	9 (4,10)
	SC/UT+EX	(7)	-0.37	(-0.98 to 0.21)	9 (6,12)					
	ExAe/UT	(2)	-0.27	(-1.18 to 0.60)	11 (4,13)					
	Wei/UT+EX	(1)	-0.24	(-1.16 to 0.63)	11 (4,13)					
	SC/UT	(10)	0.00		13 (8,13)	Standard care	(12)	0.00		8 (6,10)

SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. Results in order of mean treatment effectiveness For other abbreviations see Table 3

6.5 Pain - 3 months from the end of treatment time point

Although there were 21 trials with data available for this time point, no networks could be formed including acupuncture for the Therapy-plus-adjunct, the Therapy-only or the Grouped intervention set networks. There were two trials including acupuncture that reported an outcome at this time point and sham acupuncture was the comparator in both trials.

6.6 Summary of pain results

The main results are those for the end of treatment time point. There was no great difference in the results between the end of treatment analysis and the 3 months from start of treatment analysis. There was no network for the 3 months from the end of treatment analysis.

Overall, across all the analyses including any-quality studies, for most interventions the mean estimate of effectiveness favoured the intervention over standard care. However, the 95% credible intervals were wide and many crossed the line of no effect, i.e. they can be considered to be not statistically significant. There was overlap between the majority of credible intervals for both the mean effects and for the median ranks, indicating that caution must be exercised when interpreting these results in terms of which are 'best'.

A high level of inconsistency across the direct and indirect evidence was found for the treatment comparisons involving PES in both the Therapy-plus-adjunct and Therapy-only intervention set analyses, which suggests that there is bias or lack of exchangeability across the associated comparisons, and therefore the credibility interval estimates for PES may be underestimated. This implication may hold true for analyses including higher-quality trials only even though inconsistency for comparisons involving PES could not be evaluated due to a lack of triangles of evidence.

The results of the analyses of the Therapy-plus-adjunct intervention set provided no indication of a treatment effect difference between the majority of adjuncts. This suggests a lack of power in distinguishing between these treatment effects. Aerobic exercise with no medication was more effective than aerobic exercise with treatment as usual, but this lacks face validity. The analysis of any-quality studies for this set, found that PES, acupuncture, balneotherapy, sham acupuncture, laser/light treatment, static magnets and Tai Chi all showed a statistically significant treatment benefit over standard care, regardless of the adjunctive treatment. The sensitivity analysis of just higher-quality trials, showed a statistically significant treatment benefit over standard care for PES, acupuncture, balneotherapy, sham acupuncture, and muscle-strengthening regardless of the adjunctive treatment.

The results for the Therapy-only set (any-quality trials), found interferential therapy, acupuncture, PES, TENS, aerobic exercise, and muscle-strengthening exercise to have a statistically significant treatment benefit over standard care. The results of the sensitivity analysis (higher-quality trials only) generally reflected those of the main analysis except the credible intervals for aerobic exercise now crossed the line of no effect, the effect of balneotherapy became significant, and there were no higher-quality trials of interferential therapy. When acupuncture was compared with the other interventions in higher-quality trials it was found to be statistically significantly better at a 95% level of credibility than sham acupuncture, muscle-strengthening exercise, weight loss, aerobic exercise, and no intervention. Acupuncture's median rank was 2 (95% credible intervals 1-4).

The results for the Grouped interventions set (any-quality trials) found that acupuncture, muscle-strengthening exercise, aerobic exercise, physiotherapy treatments and Tai Chi all showed a significant treatment benefit over standard care, and acupuncture showed a significant treatment benefit over muscle-strengthening exercise, insoles, and sham acupuncture, as well as placebo and no intervention.

The results for acupuncture versus standard care across all the networks are collated in Table 35 and show a reasonable consistency.

Table 35: Summary of acupuncture versus standard care across all analyses (pain)

Analysis	SMD	CrI
End of Treatment Therapy-plus-adjunct Any Quality	-1.03	(-1.34 to -0.72)
End of Treatment Therapy-plus-adjunct Higher Quality	-1.07	(-1.38 to -0.75)
End of Treatment Therapy-only Any Quality	-0.78	(-1.10 to -0.46)
End of Treatment Therapy-only Higher Quality	-1.01	(-1.42 to -0.62)
End of Treatment Grouped Any Quality	-0.76	(-1.09 to -0.43)
End of Treatment Grouped Higher Quality	-1.01	(-1.40 to -0.64)
3 months from Start of Treatment Therapy-plus-adjunct Any Quality	-0.96	(-1.24 to -0.69)
3 months from Start of Treatment Therapy-plus-adjunct Higher Quality	-0.97	(-1.37 to -0.58)
3 months from Start of Treatment Therapy-plus-adjunct Any Quality	-0.80	(-1.13 to -0.46)
3 months from Start of Treatment Therapy-plus-adjunct Higher Quality	-0.96	(-1.45 to -0.48)
3 months from Start of Treatment Grouped Any Quality	-0.79	(-1.15 to -0.44)
3 months from Start of Treatment Grouped Higher Quality	-0.96	(-1.39 to -0.54)

6.7 WOMAC index – all outcome time points

There were few studies and few interventions included in a connected network for the analyses with a WOMAC index outcome. Details of the networks are given in Appendix 10.4.2. Details of the analysis and results for all three levels of networks (Therapy-plus-adjunct Intervention, Therapy-only Intervention, and Grouped Intervention), using all quality trials and the sensitivity analyses including only higher-quality trials are presented in Tables 37-40. No results are presented for the sensitivity analyses with higher-quality studies only for the Therapy-plus-adjunct intervention set since no stable between-study standard deviation was estimated. Most of the results presented here are compared to standard care. The full results are available at: http://www.york.ac.uk/inst/crd/Documents/FullResultsWOMAC.xlsm

Across the analyses of all available trials the results consistently indicate that acupuncture compared to standard care has a beneficial effect on the WOMAC index (mean SMD around -1.0) which is statistically significant (Tables 36-40). For other treatments a statistically significant beneficial effect could be demonstrated only in the Therapy-plus-adjunct Interventions analyses: weight loss, muscle strengthening exercise and Tai Chi (Tai Chi was only statistically significant at the end of treatment time point). Standard care plus home exercise was also found to have a beneficial effect on the WOMAC index.

The sensitivity analysis including only higher-quality trials included only acupuncture, muscle strengthening exercise, sham acupuncture and Tai Chi. Stable estimates could not be generated from the model for the sensitivity analysis 3 months post start of treatment. Also for three other sensitivity analyses models, the estimate for the between-study SD of the random effects distribution was sensitivive to the upper limit of the prior distribution for the between-study SD. Stable estimates were obtained at the increased upper limits of the prior distribution for the SD (from 2 until 12/40, depending on the analysis).

The results were consistent across the Therapy-only Intervention and Grouped intervention networks and across the time points and mean estimates of effect were similar to those from the any-quality analyses, but the results were no longer statistically significant (at the 95% level of credibility).

For the end of treatment analyses, the mean estimate of effectiveness consistently favoured acupuncture over muscle-strengthening exercise, Tai Chi, heat treatment and sham acupuncture, but there was no consistent evidence that acupuncture was statistically significantly more effective than any of the other main interventions at a 95% level of credibility. Full results are available at: http://www.york.ac.uk/inst/crd/Documents/FullResultsWOMAC.xlsm

For the 3 months from the end of treatment analyses, the mean estimate of effectiveness favoured acupuncture versus all the other interventions included in all of the analyses, except for weight loss in the Therapy-plus-adjunct Intervention set with any-quality studies. There was consistent evidence for the Therapy-only and Grouped intervention sets that acupuncture was more effective than sham acupuncture. For the Grouped intervention set, acupuncture was significantly more effective than physiotherapy treatments, balneotherapy and placebo. For the Therapy-only intervention set, acupuncture was more effective than heat treatment, one of the physiotherapy interventions. These

results were not reproduced in the higher-quality studies only sensitivity analyses due to low power to estimate treatment effect differences. Full results are available at: http://www.york.ac.uk/inst/crd/Documents/FullResultsWOMAC.xlsm

Table 36: Summary of acupuncture versus standard care across all analyses (WOMAC Index)

Analysis	SMD	CrI
End of Treatment Therapy-plus-adjunct Any Quality	-1.17	(-1.59 to -0.75)
End of Treatment Therapy-plus-adjunct Higher Quality	-1.13	(-2.07 to -0.18)
End of Treatment Therapy-plus-adjunct Any Quality	-1.01	(-1.78 to -0.24)
End of Treatment Therapy-plus-adjunct Higher Quality	-0.97	(-2.03 to 0.09)
End of Treatment Grouped Any Quality	-1.02	(-1.72 to -0.31)
End of Treatment Grouped Higher Quality	-0.97	(-2.02 to 0.10)
3 months from Start of Treatment Therapy-plus-adjunct Any Quality	-1.12	(-1.45 to -0.75)
3 months from Start of Treatment Therapy-plus-adjunct Higher Quality	-1.21	(-2.63 to 0.16)
3 months from Start of Treatment Therapy-plus-adjunct Any Quality	-0.97	(-1.35 to -0.56)
3 months from Start of Treatment Therapy-plus-adjunct Higher Quality	-0.97	(-1.93 to 0.03)
3 months from Start of Treatment Grouped Any Quality	-0.97	(-1.33 to -0.59)
3 months from Start of Treatment Grouped Higher Quality	-0.97	(-1.93 to 0.03)

Table 37: Change in WOMAC index compared to standard care: end of treatment, all networks, any-quality studies analyses

	Therapy-plus-adjunct Intervention					Therapy-only intervention					Grouped Intervention				
Number of Intervention	9					12					10				
Number of trials	11					17					17				
Mean between study SD τ (SMD)	0.29 (95%CrI:0.01 to 0.90)					0.87 (95%CrI:0.48 to 1.54)					0.79 (95%CrI:0.45 to 1.34)				
Residual deviance (%dd)	14.38 (2.7%)					19.92 (-0.4%)					19.83 (-0.9%				
Number of data points	14					20					20				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	Wei/UT	(1)	-1.61	(-2.79 to -0.42)	1 (1,6)	Insoles	(1)	-1.30	(-4.85 to 2.21)	3 (1,12)	Acupuncture	(8)	-1.02	(-1.72 to -0.31)	3 (1,7)
	Acu/UT	(6)	-1.17	(-1.59 to -0.74)	3 (1,6)	PES	(1)	-1.13	(-4.60 to 2.34)	4 (1,12)	Muscle str. exercise	(2)	-0.90	(-2.28 to 0.42)	3 (1,8)
	ExMu/UT	(2)	-1.14	(-2.14 to -0.16)	3 (1,7)	TENS	(2)	-1.07	(-3.30 to 1.15)	4 (1,11)	Weight loss	(1)	-0.78	(-2.49 to 0.94)	4 (1,10)
	Acu/NoMed	(1)	-0.88	(-2.50 to 0.77)	5 (1,9)	Acupuncture	(8)	-1.01	(-1.78 to -0.22)	4 (1,9)	Insoles	(1)	-0.75	(-3.49 to 1.90)	4 (1,10)
	SC/UT+EX	(3)	-0.83	(-1.59 to -0.04)	5 (2,8)	Placebo	(5)	-0.85	(-3.81 to 2.04)	5 (2,11)	Physiotherapy treatment	(4)	-0.54	(-2.15 to 1.00)	5 (1,9)
	Tai/UT	(3)	-0.71	(-1.42 to -0.01)	6 (2,8)	Weight loss	(1)	-0.78	(-2.69 to 1.11)	6 (1,12)	Placebo	(5)	-0.30	(-2.35 to 1.66)	6 (1,9)
	ShAcu/UT	(2)	-0.53	(-1.17 to 0.26)	7 (3,9)	Muscle str.exercise	(2)	-0.65	(-2.34 to 1.04)	6 (1,11)	Tai Chi	(2)	-0.24	(-1.48 to 1.00)	6 (1,10)
	Hea/UT	(1)	-0.44	(-1.82 to 0.92)	7 (2,9)	Tai Chi	(2)	-0.24	(-1.62 to 1.14)	8 (2,12)	Sham Acupuncture	(3)	-0.17	(-1.27 to 0.97)	7 (2,10)
						Sham Acupuncture	(3)	-0.15	(-1.37 to 1.08)	8 (2,12)	Balneotherapy	(2)	0.48	(-1.90 to 2.79)	10 (2,10)
						Balneotherapy	(2)	-0.06	(-3.29 to 3.16)	9 (2,12)					
						Heat treatment	(1)	0.05	(-2.51 to 2.62)	9 (1,12)					
	SC/UT	(6)	0.00		9 (7,9)	Standard care	(9)			9 (4,12)	Standard care	(9)			8 (4,10)
SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. For other abbreviations see Table 3 Results in order of mean treatment effectiveness															

Table 38: Change in WOMAC Index compared to standard care: End of Treatment, All networks, Higher-quality studies analyses

	Therapy-plus-adjunct Intervention				Therapy-only intervention				Grouped Intervention			
Number of Intervention	6				5				5			
Number of trials	5				6				6			
Mean between study SD τ (SMD)					The estimate for the between study SD of the random effects distribution increased as the upper limit of the prior distribution for the between study SD increased from 2 until 15. The increase in the between-study SD was 12% over this range. The results reported here are for a prior upper limit of 15. 0.81 (95%CrI:0.27 to 2.27)				The estimate for the between study SD of the random effects distribution increased as the upper limit of the prior distribution for the between study SD increased from 2 until 12. The increase in the between-study SD was 12% over this range. The results reported here are for a prior upper limit of 12. 0.81 (95%CrI:0.26 to 2.29)			
Residual deviance (%dd)					8.05 (0.6%) (for a prior upper limit of 15)				8.06 (0.8%) (for a prior upper limit of 12)			
Number of data points					8 (for a prior upper limit of 15)				8 (for a prior upper limit of 12)			
Results	Treatment	No. trials	SMD	(95% CrI) Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI) Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI) Median rank (CrI)
	Acu/UT	(3)			Acupuncture	(4)	-0.97	(-2.07 to 0.14) 1 (1,4)	Acupuncture	(4)	-0.97	(-2.07 to 0.14) 1 (1,4)
	ExMu/UT	(1)			Muscle str. exercise	(1)	-0.60	(-2.42 to 1.16) 2 (1,5)	Muscle str. exercise	(1)	-0.60	(-2.41 to 1.19) 2 (1,5)
	SC/UT+EX	(2)			Sham Acupuncture	(2)	-0.31	(-1.81 to 1.18) 3 (1,5)	Sham Acupuncture	(2)	-0.31	(-1.82 to 1.21) 3 (1,5)
	ShAcu/UT	(1)			Tai Chi	(2)	-0.24	(-1.64 to 1.16) 3 (1,5)	Tai Chi	(2)	-0.24	(-1.63 to 1.18) 3 (1,5)
	Tai/UT	(2)										
	SC/UT	(3)			Standard care	(5)	0.00	4 (2,5)	Standard care	(5)	0.00	4 (2,5)
No results are presented for the sensitivity analysis with higher-quality studies only for the Therapy-plus-adjunct intervention set since no stable between-study standard deviation was estimated. SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. For other abbreviations see Table 3 Results in order of mean treatment effectiveness												

Table 39: Change in WOMAC index compared to standard care: three months from start of treatment, all networks, any-quality studies analyses

	Therapy-plus-adjunct Intervention					**Therapy-only intervention**					**Grouped Intervention**				
Number of Intervention	10					13					10				
Number of trials	12					19					19				
Mean between study SD τ (SMD)	0.23 (95%CrI:0.01 to 0.71)					0.38 (95%CrI:0.14 to 0.74)					0.35 (95%CrI:0.13 to 0.65)				
Residual deviance	14.19 (-5.4%)					21.43 (-2.6%)					21.34 (-3%)				
Number of data points	15					22					22				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	Wei/UT	(1)	-1.48	(-2.47 to -0.45)	1 (1,5)	Acupuncture	(8)	-0.97	(-1.35 to -0.56)	2 (1,8)	Acupuncture	(8)	-0.97	(-1.33 to -0.59)	1 (1,3)
	Acu/UT	(6)	-1.12	(-1.45 to -0.74)	2 (1,5)	Laser	(1)	-0.86	(-2.96 to 1.27)	3 (1,11)	Weight loss	(1)	-0.77	(-1.67 to 0.12)	2 (1,7)
	ExMu/UT	(2)	-0.93	(-1.78 to -0.07)	3 (1,8)	Weight loss	(1)	-0.78	(-1.73 to 0.17)	4 (1,11)	Muscle str. exercise	(3)	-0.43	(-1.01 to 0.09)	3 (1,6)
	SC/UT+EX	(3)	-0.71	(-1.35 to -0.02)	5 (2,8)	Insoles	(1)	-0.52	(-2.69 to 1.69)	5 (1,13)	Sham Acupuncture	(3)	-0.25	(-0.82 to 0.37)	4 (2,9)
	Wei/UT	(1)	-0.61	(-1.16 to 0.07)	6 (2,9)	Muscle str. exercise	(3)	-0.35	(-0.98 to 0.23)	7 (2,12)	Tai Chi	(2)	-0.24	(-0.93 to 0.45)	4 (2,9)
	Tai/UT	(3)	-0.56	(-1.17 to 0.09)	6 (2,9)	PES	(1)	-0.34	(-2.47 to 1.81)	7 (1,13)	Physiotherapy treatment	(5)	0.48	(-0.48 to 1.36)	7 (4,9)
	Acu/NoMed	(1)	-0.54	(-2.09 to 0.94)	6 (1,10)	TENS	(2)	-0.32	(-1.89 to 1.28)	7 (2,12)	Insoles	(1)	0.51	(-1.04 to 1.96)	8 (2,10)
	ExMu/NoMed	(1)	-0.28	(-0.93 to 0.38)	8 (3,10)	Balneotherapy	(2)	-0.26	(-2.27 to 1.78)	8 (2,13)	Balneotherapy	(2)	0.76	(-0.54 to 1.97)	9 (4,10)
	Hea/UT	(1)	0.25	(-0.94 to 1.46)	10 (5,10)	Tai Chi	(2)	-0.24	(-0.96 to 0.48)	8 (2,12)	Placebo	(6)	0.95	(-0.16 to 1.98)	10 (7,10)
						Sham Acupuncture	(3)	-0.24	(-0.85 to 0.41)	8 (2,12)					
						Placebo	(6)	-0.06	(-1.95 to 1.86)	9 (4,13)					
						Heat treatment	(1)	0.82	(-0.38 to 1.98)	13 (6,13)					
	SC/UT	(7)	0.00		9 (7,10)	Standard care	(3)	0.00		10 (5,12)	Standard care	(10)	0.00		6 (4,9)
SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. For other abbreviations see Table 3 Results in order of mean treatment effectiveness															

Table 40: Reduction in WOMAC index compared to standard care: three months from start of treatment, all networks, higher-quality studies analyses

	Therapy-plus-adjunct Intervention					Therapy-only intervention					Grouped Intervention				
Number of Intervention	7					5					5				
Number of trials	5					6					6				
Mean between study SD τ (SMD)	Stable estimates were produced by the model for the standardised mean differences but not for the between study SD of the random effects distribution τ. The results are not presented.					0.76 (95%CrI:0.24 to 2.16)					The estimate for the between study SD of the random effects distribution increased as the upper limit of the prior distribution for the between study SD increased from 2 until 30. The increase in the between-study SD was 11% over this range. The results reported here are for a prior upper limit of 30. 0.76 (95%CrI:0.24 to 2.16)				
Residual deviance (%dd)						8.26 (3.3%)					8.26 (3.3%) (for a prior upper limit of 30)				
Number of data points						8					8 (for a prior upper limit of 30)				
Results	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)	Treatment	No. trials	SMD	(95% CrI)	Median rank (CrI)
	Acu/UT	(3)				Acupuncture	(4)	-0.97	(-1.97 to 0.08)	1 (1,4)	Acupuncture	(4)	-0.97	(-1.97 to 0.07)	1 (1,3)
	ExMu/UT	(1)				Muscle str. exercise	(2)	-0.38	(-1.65 to 0.87)	3 (1,5)	Muscle str. exercise	(2)	-0.38	(-1.64 to 0.86)	3 (1,5)
	Tai/UT	(1)				Sham Acupuncture	(2)	-0.31	(-1.71 to 1.13)	3 (1,5)	Sham Acupuncture	(2)	-0.31	(-1.71 to 1.13)	3 (1,5)
	SC/UT+EX	(2)				Tai Chi	(1)	-0.09	(-1.99 to 1.82)	4 (1,5)	Tai Chi	(1)	-0.09	(-2.00 to 1.82)	4 (1,5)
	ShAcu/UT	(1)													
	ExMu/NoMed	(1)													
	Sc/UT	(3)				Standard care	(5)	0.00		4 (2,5)	Standard care	(5)	0.00		4 (2,5)
No results are presented for the sensitivity analysis with higher-quality studies only for the Therapy-plus-adjunct intervention set since no stable between-study standard deviation was estimated. SMD Standardised mean difference; CrI Credible intervals; %dd % deviance difference. For other abbreviations see Table 3 Results in order of mean treatment effectiveness															

6.8 Comparison of network meta-analysis SMDs with pair wise meta-analysis SMDs

Table 41 compares the SMD estimates using NMA with the estimates using pair wise meta-analysis. There are two factors which may contribute to differences in credible/confidence intervals between the NMA and MA results. Firstly, in an NMA the between-study variance is shared by all the interventions. If the between-study variance is greater in the NMA than in the MA for a particular intervention, this will tend to make the credible interval from the NMA wider than the confidence interval from the MA, and vice-versa.

Secondly, the indirect evidence has an impact on the estimate; it can add power making the credible interval more precise. The more trials there are informing the indirect evidence, the more power the indirect evidence will add. The more direct evidence there is, the less of an impact the indirect evidence will have. But if the mean estimate from the indirect evidence is different to the estimate from the direct evidence (i.e. there is some inconsistency) then this may widen the credible interval.

Table 41: Comparison of the end of treatment pain (therapy-only) network meta-analysis SMDs with the pair wise meta-analysis SMDs (for comparisons with standard care, any-quality)

Intervention	No. of trials representing the intervention	NMA SMD (95% Cr I)	No. of trials representing the intervention	Pair wise meta-analysis SMD (95% CI)
Standard care (comparator)	39	-	-	-
Interferential therapy	2	-1.06 (-2.05 to -0.07)	-	Not estimable
Acupuncture	18	-0.78 (-1.10 to -0.46)	9	-0.75 (-1.13 to -0.37)
Static magnets	1	-0.78 (-1.92 to 0.36)	-	Not estimable
PES	4	-0.70 (-1.35 to -0.05)	-	Not estimable
TENS	11	-0.62 (-1.07 to -0.18)	1	-1.29 (-2.58 to 0.01)
Aerobic exercise	5	-0.60 (-1.05 to -0.16)	5	-0.70 (-1.28 to -0.12)
Balneotherapy	8	-0.46 (-0.96 to 0.05)	1	-1.01 (-1.48 to -0.54)
Muscle str. exercise	24	-0.36 (-0.59 to -0.14)	17	-0.35 (-0.45 to -0.26)
Weight loss	3	-0.36 (-0.89 to 0.16)	3	-0.06 (-0.22 to 0.10)
Manual therapy	3	-0.30 (-0.91 to 0.32)	2	-0.14 (-0.55 to 0.28)
Tai Chi	3	-0.28 (-0.86 to 0.29)	3	-0.30 (-0.54 to -0.06)
Laser therapy	3	-0.25 (-0.98 to 0.48)	-	Not estimable
Ice/cooling treatment	3	-0.24 (-1.01 to 0.54)	-	Not estimable
Braces	1	-0.15 (-1.08 to 0.78)	-	Not estimable
Heat treatment	4	-0.04 (-0.66 to 0.57)	-	Not estimable
Insoles	3	-0.01 (-0.72 to 0.69)	-	Not estimable
PEMF	4	0.01 (-0.63 to 0.64)	-	Not estimable
NMES	1	0.46 (-0.64 to 1.55)	1	0.46 (-0.23 to 1.14)

7. DISCUSSION

7.1 Statement of principal findings

We have conducted a rigorous systematic review and network meta-analysis to evaluate the efficacy of acupuncture, and its relevant comparators, for alleviating knee OA pain. The main results are those for the end of treatment time point as this was reported by most studies and there was no great difference in results between this time point and the 3 months from start of treatment time point, given that two-thirds of the data used for the two time points were the same. Our results, therefore, relate only to short-term benefit, due to the limited treatment durations in most of the trials.

The standard meta-analysis provided some evidence, when all studies were considered, of a benefit on OA of the knee pain with acupuncture, muscle strengthening exercise, aerobic exercise, balneotherapy, TENS, static magnets, braces, NMES, and interferential therapy. However the quality of most trials was poor and sample sizes small and interpretation of the results across the interventions was difficult. These analyses identified four trials as potential sources of significant heterogeneity in the network meta-analyses.

The NMA results for the Therapy-only set (any-quality trials) indicates that a number of treatments appear to have a beneficial mean effect and that acupuncture is a worthwhile treatment option for treating knee pain due to osteoarthritis. The results of the main analysis were supported by the sensitivity analyses of higher-quality studies in that acupuncture was significantly better than standard care, sham acupuncture, and muscle-strengthening exercise. In the sensitivity analysis of only higher-quality trials, only PES had a greater mean treatment effect estimate and higher median rank than acupuncture, but, as less evidence informed the PES estimate (1 trial) than the acupuncture estimate (8 trials), the 95% credible interval of the treatment effect ranks were the same for PES and acupuncture (95% Crl rank: 1-4). This analysis also indicated that muscle-strengthening exercise and balneotherapy have pain-alleviating effects significantly better than standard care (although the credible interval for the sham acupuncture barely crosses the line of no effect). Again, there was a large difference in the number of trials informing the network for each of these interventions (seven trials for muscle-strengthening exercise, but only one small study for balneotherapy).

There was no evidence to demonstrate a statistically significant difference between many of the treatments in the Therapy-only analyses and between different adjunct therapies in the Therapy-plus-adjunct analyses. These findings are perhaps not surprising, as there were only a limited number of small trials investigating many of the interventions studied (inevitably resulting in wider credible intervals), and where the evidence was indirect the uncertainty in the treatment effect was greater.

Our data on acupuncture were found to be reasonably consistent across the three groupings of interventions (Therapy-only, Therapy-plus-adjunct and Grouped) and two levels of trial quality (any-quality and higher-quality) as well as between the pair-wise (direct) comparisons and the network meta-analyses (direct and indirect comparisons) for those networks that were stable.

A high level of inconsistency across the direct and indirect evidence was found for the treatment comparisons involving PES in both the Therapy-plus-adjunct and Therapy-only intervention set analyses, which suggests that there is bias or lack of exchangeability across the associated comparisons, and therefore the credibility interval estimates for PES may be underestimated; there may be an effect modifier unacounted for.

There were few studies and few interventions available to form connected networks for the WOMAC index outcome. However, across the analyses of all available trials the results consistently indicate that acupuncture, compared to standard care, has a beneficial effect (mean SMD around -1.0).

Numerous systematic reviews (some summarised in a review of reviews[162]) have been conducted on interventions (or classes of interventions) included in this review; many have produced effect sizes derived from direct comparisons. A Cochrane review evaluating acupuncture versus waiting list control reported a statistically significant, clinically relevant short-term improvement in pain, similar to our findings (SMD -0.96, 95% CI: -1.19 to -0.72).[163]

Although, for many of the interventions in our review most of the trials informing the network were of poor quality, resulting in a lack of reliable evidence to support their efficacy, the possibility of benefit from these treatments cannot be ruled out. The exceptions to this conclusion appear to be insoles (without ankle support) and laser/light therapy: although they had no comparators to connect them to the main network, pair-wise meta-analyses of studies of satisfactory quality suggested no worthwhile benefit compared with placebo.

Various quantifications of the clinical relevance of knee OA pain change scores exist. Two such estimates have been reported by one study;[164] the 'minimal clinically important change' (MCIC) was -15mm (on a VAS 0-100 scale, and derived from a prior Delphi exercise[165]), and the 'minimal perceptible clinical improvement' (MPCI, the smallest change detectable by the patient) was -9.7mm (on a WOMAC VAS 0-100 scale). Another study estimated the 'minimal clinically important improvement' (MCII), although only for pain on movement, as being -19.9mm on a VAS 0-100 scale; this figure varied by baseline pain score, with patients with less pain having a smaller MCII (10.8mm) and patients with severe pain having a larger MCII (36.6mm).[166] In the context of these studies, our results indicate that acupuncture produces both a MPCI, and quite possibly a MCIC (we could not locate a MCIC specifically for WOMAC pain), but may only yield a MCII for patients with low levels of pain (see Table 31).

Another useful method for interpreting results is the comparison of effect sizes with those of other relevant available treatments which are generally considered to have clinically meaningful benefits.[167] Other reviews have reported SMDs (versus placebo) for pain relief from paracetamol, oral, and topical NSAIDs of 0.14 (95% CI 0.08 to 0.23); 0.29 (95% CI: 0.22 to 0.35); and 0.44 (95% CI: 0.27 to 0.62) respectively.[168] The SMDs for acupuncture from our results compare favourably with these pharmacological treatments, indicating that acupuncture has a credible role to play in the management of knee OA pain. Further factors to consider when interpreting results are safety, the rapidity of onset - and durability - of treatment benefit, patient choice, and the convenience, cost, and likelihood of patient adherence to treatment; these factors would clearly differ when comparing acupuncture with pharmacological treatments.[167]

7.2 Strengths and limitations of the review

Our aim was not to replicate earlier reviews, but to use network meta-analysis methods – which allow inclusion of all relevant direct and indirect evidence - to investigate the relative efficacies of treatments. When trials comparing all the alternative treatments of interest do not exist for decision-makers, a network meta-analysis enables decisions to be made regarding the options that might be considered regarding the choice of treatment, and provides a basis for considering the need for further research comparing treatments. A network meta-analysis allows for comparisons on the basis of synthesising all the available evidence in a consistent framework, rather than making such decisions by subjective inferences from disparate data. We believe our study is the first network meta-analysis of physical treatments for OA of the knee. As such we encountered significant methodological challenges.

A network meta-analysis retains the benefits of randomisation by entering the trial treatment effects in the model. In a random effects model a common between-study variance across the comparisons, and a normal distribution, are assumed. It is possible that between-study variances are not equal across comparisons, but the assumption allows a between-study variance to be estimated for comparisons with only one or two trials. It is also assumed that direct and indirect evidence is exchangeable for any one comparison. This requires the interventions and populations to be sufficiently similar across the trials providing the direct and indirect evidence, and for there to be unbiased outcomes due to well-controlled, designed trials. In this study, the main cause for concern regarding heterogeneity between comparisons, and thereby between direct and indirect evidence (inconsistency), involved the placebo treatment group. This assumed that the placebo effects were the same for different therapies, and that the placebos were equally plausible. In addition, the elimination of poor quality studies for the Grouped intervention set for the end of treatment analysis produced significantly diverging results for placebo and also, although less so, for insoles and physiotherapy treatment. This suggests that either study quality had an effect on the outcome for at least one of the comparisons involving these interventions or that there is a significant difference in treatment effect between treatments lumped together.

Our review was extensive in its eligibility criteria, encompassing a large number of interventions, with various adjunct therapies, placebos and populations, which meant a certain amount of clinical heterogeneity was inevitable. Most studies recruited general knee OA populations, although it was acknowledged that within this categorisation there will have been variation in characteristics. Where populations were noticeably different (e.g. patients awaiting knee surgery), in order to assess the effect on the results and model fit, we performed a sensitivity analysis with them removed. Furthermore, heterogeneity was also explored using standard meta-analysis and trials which were clearly a source of heterogeneity were removed from analyses; treatment duration and intensity sometimes varied considerably within a single intervention, and varied enormously between types of intervention. This variation between interventions was to be expected, considering the diversity of treatments included. Occasionally an individual study used different doses or types of the same intervention; in such cases we pooled the results. This maximised the amount of data available for our analyses, and also avoided being restrictive in the regimens studied. However, the results we used for these trials can only be taken as a generalisable estimate of efficacy.

Our comprehensive and rigorous search strategy minimised the risk of missing eligible trials. However, although our review included 134 studies, limitations and differences in the reporting of data restricted the data available for our analyses, such that only 87 trials were included in the standard and network meta-analyses. Furthermore, the lack of long term data limits the interpretation of the results to only the short term effects of therapy.

To enable our analysis to be thorough, we examined the different ways in which treatments could be defined, or grouped. Many interventions were given in addition to standard, or adjunctive, care and it was recognised that adjunctive care varies between trials ranging from no treatment at all, to care bordering on being an additional active intervention. In order to allow a thorough assessment of the possible effects of this variation either within or between studies we classified these adjunct treatments for each trial arm. There was no evidence for any differences amongst the adjunct treatments for any of the treatment classes. This analysis had greater power than the analysis with more Therapy-plus-adjunct treatment definitions, and the exchangeability assumption was not compromised. However, our analyses of these studies did indicate that the type of adjunctive treatment used did not appear to have a noticeable impact on effect size and heterogeneity (this finding was supported by examination of the standard meta-analysis forest plots).

A more general concern, which will have been relevant to all reviews of these physical therapies, not just a network meta-analysis, relates to the poor quality of a large majority of the studies. Despite adequate reporting of randomisation and allocation concealment procedures not being necessary in order for studies to achieve a satisfactory quality rating, more than three-quarters of the studies in our review were still classified as being of poor quality. However, a major strength of our review is that trials of a diverse range of interventions have been evaluated equally, using the same quality assessment tool; this allows for a fair comparison in terms of assessing the strength of the evidence base for each intervention.

Study quality issues for the interventions we studied have been noted previously. A study published in 2003 compared over 100 RCTs (published in the preceding 10 years) of non-pharmacological and pharmacological treatments for hip or knee osteoarthritis found that the non-pharmacological articles scored lower than reports of pharmacological treatments for study quality assessment using the Jadad scale and the Delphi list.[169]

Our review evaluated only studies where the active interventions were delivered in isolation, rather than when combined with other physical interventions; in practice, interventions may be given in combination. It should therefore be noted, when interpreting our results, that trials investigating such mixed interventions were excluded. This decision was taken to facilitate homogeneity of interventions, as it was thought likely that combination therapies might vary considerably, which might then preclude meaningful pooling of data.

It was hoped and anticipated that the majority of studies meeting the eligibility criteria in our review would use the WOMAC pain subscale to assess pain. However, details on which type of pain was recorded (e.g. walking, stair climbing, resting) were frequently absent; when they were provided, it was clear that several different types of pain were assessed. The lack of detail in the reporting of pain in many papers may have contributed to heterogeneity between trials. Lamentably, only around a

quarter of studies reported WOMAC pain scores in a useable format. It is unclear why so many studies did not use WOMAC scores, which were recommended as a primary measure of efficacy in knee OA trials at a key consensus meeting in 1994.[170] WOMAC has been shown to be both a reliable and valid tool in studies of osteoarthritis of the knee and hip.[171] Its limited use, coupled with the wide range of other tools used in studies in our review, has produced another possible source of heterogeneity between studies.

This restricted use of WOMAC was accompanied by limited reporting of how scores were calculated. Ideally the scores relating to each WOMAC question for each participant would simply be summed to produce a total score for each subscale (and a total overall score), the (study population) mean of which would be reported. However, it became apparent that many studies performed transformations of the scores, e.g. standardising to a 0-10 or 0-100 range, or (rarely) a more unorthodox range. Occasionally such transformations were explained by authors, but more commonly they were not. Sometimes, in studies of patients with low baseline pain scores, it was not even clear whether a Likert or a VAS 0-10 scale was used. In such cases we had to exercise judgement in assigning a type of pain scale to the trial in question. In light of these concerns about the reporting of scores, the most appropriate method of preparing the data for synthesis was conversion to standardised mean differences, as this would obviate the uncertainties. By reporting outcomes only as changes in from baseline, medians, percentage changes, or just p-values, data from many studies were not suitable for pooling.

8. CONCLUSIONS

The first network meta-analysis of physical interventions for knee pain due to osteoarthritis, indicates that acupuncture is one of a number of physical treatments that produces a clinically-relevant effect in alleviating pain in the short-term. Moreover, acupuncture compared favourably with the other treatments. Although further research is needed to substantiate these conclusions, acupuncture should nevertheless be considered as an evidence-based treatment option for relieving pain due to osteoarthritis of the knee.

8.1 Implications for service provision

Acupuncture can be considered as an evidence-based treatment option for relieving pain due to osteoarthritis of the knee. Although our review did not evaluate the cost-effectiveness of the interventions, it is worth noting that our main results on effectiveness do not concur with the NICE guidance for osteoarthritis management[6] which states that TENS, insoles, braces, manual therapy, and heat or cold (thermotherapy) should be considered as adjuncts to core treatment. For these interventions our end of treatment analyses found no evidence (of significant differences from standard care) to support this guidance, other than for TENS where the evidence was equivocal: all the TENS studies in our analyses were of poor quality, raising concerns about the reliability of the evidence. We have provided evidence on the effectiveness of acupuncture that NICE may want to consider when revising their guidance.

8.2 Implications for research

Larger, more robust RCTs with longer treatment periods, which also examine the effectiveness of re-treatment following treatment cessation (to evaluate durability and attenuation effects) are needed in order to comprehensively assess the value of many of these interventions. This is particularly true for TENS, where the studies conducted so far have been of unreliable quality, and PES and balneotherapy, which although our results highlight them as being promising treatments, were both represented by only one small higher-quality trial. The adoption of consistent study designs across trials would facilitate improved indirect analysis in the future to substantiate our findings.

As our review identified specific interventions that had some beneficial effect, the next step might be to test whether combinations of the more promising treatments might result in an even greater effect. However, it cannot be assumed that combining interventions will be beneficial: it has been suggested that acupuncture and muscle-strengthening exercise are effective via similar mechanisms,[172] so an additive effect may not be possible. An RCT investigating such a mixed intervention found no significant WOMAC pain differences between the acupuncture and exercise/advice and exercise/advice alone groups at 6 weeks, 6 months, and 12 months.[173] The true value of this mixed intervention remains somewhat uncertain since the delivery of acupuncture in the study may have been sub-optimal (6 sessions over 3 weeks), and pain was not assessed at the end of acupuncture treatment (at 3 weeks). The optimum timing and parameters of treatment for both acupuncture and muscle-strengthening exercise also need to be more clearly defined from future studies.

For example, for acupuncture it is unclear what might be the optimum number and frequency of sessions, which style of acupuncture is more effective, and what level of training would be ideal.

Results from our pairwise meta-analysis of higher-quality insoles studies suggested that insoles (without ankle support) were unlikely to be effective; there appears to be little value in conducting further research (as a recently published trial confirms).[174]

If adding or subtracting trials causes results to change more than expected, given the credible intervals around the estimates, this indicates unquantified uncertainty and unreliable results. Research could be conducted to develop a statistic to measure the stability or instability of the results, given change in the evidence base.

9. REFERENCES

1. *Common uses of acupuncture*. NHS Choices, 2010. http://www.nhs.uk/Conditions/Acupuncture/Pages/What-is-it-used-for.aspx.
2. Sun Y, Gan TJ, Dubose JW, Habib AS. Acupuncture and related techniques for postoperative pain: a systematic review of randomized controlled trials. *Br J Anaesth* 2008;101:151-60.
3. Peat G, McCartney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of health care. *Ann Rheum Dis* 2001;60:91-7.
4. Lethbridge-Çejku M, Scott WW, Reichle R, Ettinger WH, Zonderman A, Costa P, et al. Association of radiographic features of osteoarthritis of the knee with knee pain: data from the Baltimore longitudinal study of aging. *Arthritis Care Res* 1995;8:182-88.
5. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JWJ, Dieppe P, et al. EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a task force of the standing committee for international clinical studies including therapeutic trials (ESCISIT). *Ann Rheum Dis* 2003;62:1145-55.
6. National Collaborating Centre for Chronic Conditions. *Osteoarthritis: national clinical guideline for care and management in adults*. London: Royal College of Physicians, 2008.
7. Tramer MR, Moore RA, Reynolds DJ, McQuay HJ. Quantitative estimation of rare adverse events which follow a biological progression: a new model applied to chronic NSAID use. *Pain* 2000;85:169-82.
8. Pound P, Britten N, Morgan M, Yardley L, Pope C, Daker-White G, et al. Resisting medicines: a synthesis of qualitative studies of medicine taking. *Soc Sci Med* 2005;61:133-55.
9. Arthritis Care. *Osteoarthritis nation: the most comprehensive UK report of people with osteoarthritis*. London: Arthritis Care, 2004.
10. Devos-Comby L, Cronan T, Roesch SC. Do exercise and self-management interventions benefit patients with osteoarthritis of the knee? A metaanalytic review. *J Rheumatol* 2006;33:744-56.
11. Gilchrist I. Anterior knee pain. *Hands On* 2004;5:1-5.
12. Cummings M. Why recommend acupuncture for low back pain but not for osteoarthritis? A commentary on recent NICE guidelines. *Acupunct Med* 2009;27:128-9.
13. Centre for Reviews and Dissemination. *Systematic reviews: CRD's guidance for undertaking reviews in health care*. York: Centre for Reviews and Dissemination, 2009.
14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
15. Lipsey M, Wilson D. *Practical meta-analysis*. Newbury Park, CA: Sage, 2001.
16. [017] Is the standardised mean difference a suitable measure of treatment effect? 7th Cochrane Colloquium; 1999 5-9 October Rome, Italy.
17. Scholten RJ, de Beurs E, Bouter LM. From effect size into number needed to treat. *Lancet* 1999;354:598.
18. Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.2 [updated September 2009]. Oxford: The Cochrane Collaboration, 2009.
19. Whitehead A. *Meta-analysis of controlled clinical trials*. Chichester: John Wiley & Sons, 2002.
20. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44.
21. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
22. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in mixed treatment comparisons: Application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. *Stat Med* 2009;28:1861-81.
23. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008;26:753-67.
24. Ades AE, Welton N, Lu G. *Mixed treatment comparisons*. Community Based Medicine, University of Bristol, https://www.bris.ac.uk/cobm/research/mpes/mtc.html.
25. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7:434-55.
26. Lim B-W, Hinman RS, Wrigley TV, Sharma L, Bennell KL. Does knee malalignment mediate the effects of quadriceps strengthening on knee adduction moment, pain, and function in medial knee osteoarthritis? A randomized controlled trial. *Arthritis Care Res* 2008;59:943-51.

27. Fischer G, Pelka RB, Barovic J. [Adjuvant treatment of osteo arthritis of the knee with weak pulsing magnetic fields: results of a prospective, placebo-controlled trial]. *Z Orthop Ihre Grenzgeb* 2005;143:544-50.
28. Rapp W, Boeer J, Albrich C, Heitkamp HC. [Efficiency of vibration or strength training for knee stability in osteoarthritis of the knee] *Aktuel Rheumatol* 2009;34:240-45.
29. Molsberger A, Bowing G, Jensen KU, Lorek M. [Acupuncture treatment for the relief of gonarthrosis pain: a controlled clinical trial] *Schmerz* 1994;8:37-42.
30. Arman MI. [Therapeutischer Effekt der Kryotherapie bei der aktivierten Gonarthrose. Eine untersucherblinde, kontrollierte Studie]. *Z Phys Med Baln Med Klim* 1988;17:368-69.
31. Wu Z-h, Bao F. [Observation on therapeutic effects of electroacupuncture for the treatment of knee osteoarthritis]. *Zhongguo Gu Shang* 2008;21:170-2.
32. Miranda-Filloy JA, Barbazan AC, Monteagudo Sanchez B, Grana GJ, Galdo Fernandez F. [Effect of transcutaneous electrical quadriceps muscle stimulation in knee osteoarthritis symptomatology]. *Rehabilitacion* 2005;39:167-70.
33. 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. *Chin Med J* 2010;123:165-72.
34. Williamson L, Wyatt MR, Yein K, Melton JTK. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. *Rheumatology (Oxford)* 2007;46:1445-9.
35. 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). *Chin Med* 2008;3:2.
36. Ng MML, Leung MCP, Poon DMY. The effects of electro-acupuncture and transcutaneous electrical nerve stimulation on patients with painful osteoarthritic knees: a randomized controlled trial with follow-up evaluation. *J Altern Complement Med* 2003;9:641-9.
37. Yurtkuran M, Kocagil T. TENS, electroacupuncture and ice massage: comparison of treatment for osteoarthritis of the knee. *Am J Acupunct* 1999;27:133-40.
38. 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. *Acupunct Med* 2008;26:17-26.
39. 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. *Acupunct Med* 2008;26:69-78.
40. 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. *Evid Based Complement Alternat Med* 2009;doi:10.1093/ecam/nen080.
41. 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 Musculoskelet Disord* 2009;10:130.
42. 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;38:346-54.
43. Tukmachi E, Jubb R, Dempsey E, Jones P. The effect of acupuncture on the symptoms of knee osteoarthritis - an open randomised controlled study. *Acupunct Med* 2004;22:14-22.
44. 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 Rheum* 2006;54:3485-93.
45. 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;366:136-43.
46. 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;329:1216.
47. 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 Complement Altern Med* 2002;2:3.
48. Bao F, Wu Z. Observation on therapeutic effect of knee osteoarthritis treated by electroacupuncture. *International Journal of Clinical Acupuncture* 2007;16:191-5.
49. 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;3:112-15.

50. Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AMK, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial *Ann Intern Med* 2004;141:901-10.
51. Takeda W, Wessel J. Acupuncture for the treatment of pain of osteoarthritic knees. *Arthritis Care Res* 1994;7:118-22.
52. 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. *J Am Geriatr Soc* 2007;55:1541-7.
53. Keogan F, Gilsenan C, Hussey J, O'Connell P. Open or closed chain quadriceps exercises in treatment of osteoarthritis of the knee; which is more effective? A blinded randomised controlled trial. *Physiotherapy Ireland* 2007;28:47-8.
54. Borjesson M, Robertson E, Weidenhielm L, Mattsson E, Olsson E. Physiotherapy in knee osteoarthrosis: effect on pain and walking. *Physiother Res Int* 1996;1:89-97.
55. Gur H, Cakin N, Akova B, Okay E, Kucukoglu S. Concentric versus combined concentric-eccentric isokinetic training: effects on functional capacity and symptoms in patients with osteoarthrosis of the knee. *Arch Phys Med Rehabil* 2002;83:308-16.
56. Huang MH, Lin YS, Lee CL, Yang RC. Use of ultrasound to increase effectiveness of isokinetic exercise for knee osteoarthritis. *Arch Phys Med Rehabil* 2005;86:1545-51.
57. Jan MH, Lin JJ, Liau JJ, Lin YF, Lin DH. Investigation of clinical effects of high- and low-resistance training for patients with knee osteoarthritis: a randomized controlled trial. *Phys Ther* 2008;88:427-36.
58. Rogind H, Bibow-Nielsen B, Jensen B, Moller HC, Frimodt-Moller H, Bliddal H. The effects of a physical training program on patients with osteoarthritis of the knees. *Arch Phys Med Rehabil* 1998;79:1421-7.
59. Lund H, Weile U, Christensen R, Rostock B, Downey A, Bartels EM, et al. A randomized controlled trial of aquatic and land-based exercise in patients with knee osteoarthritis. *J Rehabil Med* 2008;40:137-44.
60. Topp R, Woolley S, Hornyak J, Khuder S, Kahaleh B. The effect of dynamic versus isometric resistance training on pain and functioning among adults with osteoarthritis of the knee. *Arch Phys Med Rehabil* 2002;83:1187-95.
61. Baker KR, Nelson ME, Felson DT, Layne JE, Sarno R, Roubenoff R. The efficacy of home based progressive strength training in older adults with knee osteoarthritis: a randomized controlled trial. *J Rheumatol* 2001;28:1655-65.
62. Ettinger WH, Burns R, Messier SP, Applegate W, Rejeski WJ, Morgan T, et al. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). *JAMA* 1997;277:25-31.
63. Trans T, Aaboe J, Henriksen M, Christensen R, Bliddal H, Lund H. Effect of whole body vibration exercise on muscle strength and proprioception in females with knee osteoarthritis. *Knee* 2009;16:256-61.
64. Hay EM, Foster NE, Thomas E, Peat G, Phelan M, Yates HE, et al. Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomised trial. *BMJ* 2006;333:995.
65. Callaghan MJ, Oldham JA, Hunt J. An evaluation of exercise regimes for patients with osteoarthritis of the knee: a single-blind randomized controlled trial. *Clin Rehabil* 1995;9:213-18.
66. Peloquin L, Bravo G, Gauthier P, Lacombe G, Billiard JS. Effects of a cross-training exercise program in persons with osteoarthritis of the knee: a randomized controlled trial. *J Clin Rheumatol* 1999;5:126-36.
67. Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. *Arch Phys Med Rehabil* 1999;80:1293-9.
68. McCarthy CJ, Mills PM, Pullen R, Richardson G, Hawkins N, Roberts CR, et al. Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: a randomised controlled trial and health economic analysis. *Health Technol Assess* 2004;8:1-61.
69. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E, et al. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. *Arthritis Care Res* 2007;57:1211-9.
70. Abrahams S, Demetriou P. A comparison of the benefits of physiotherapy and anti-inflammatory drugs for osteoarthritis of the knee. *Journal Of Orthopaedic Medicine* 2002;24:79-85.
71. An B, Dai K, Zhu Z, Wang Y, Hao Y, Tang T, et al. Baduanjin alleviates the symptoms of knee osteoarthritis. *J Altern Complement Med* 2008;14:167-74.

72. Kuptniratsaikul V, Tosayanonda O, Nilganuwong S, Thamalikitkul V. The efficacy of a muscle exercise program to improve functional performance of the knee in patients with osteoarthritis. *J Med Assoc Thai* 2002;85:33-39.
73. Rosemffet MG, Schneeberger EE, Citera G, Sgobba ME, Laiz C, Schmulevich H, et al. Effects of functional electrostimulation on pain, muscular strength, and functional capacity in patients with osteoarthritis of the knee. *J Clin Rheumatol* 2004;10:246-9.
74. Cheing GLY, Hui-Chan CWY, Chan KM. Does four weeks of TENS and/or isometric exercise produce cumulative reduction of osteoarthritic knee pain? *Clin Rehabil* 2002;16:749-60.
75. Bezalel T, Carmeli E, Katz-Leurer M. The effect of a group education programme on pain and function through knowledge acquisition and home-based exercise among patients with knee osteoarthritis: a parallel randomised single-blind clinical trial. *Physiotherapy* 2010;96:137-43.
76. Durmus D, Alayli G, Canturk F. Effects of quadriceps electrical stimulation program on clinical parameters in the patients with knee osteoarthritis. *Clin Rheumatol* 2007;26:674-8.
77. Hasegawa R, Islam MM, Nasu E, Tomiyama N, Lee SC, Koizumi D, et al. Effects of combined balance and resistance exercise on reducing knee pain in community-dwelling older adults. *Phys Occup Ther Geriatr* 2010;28:44-56.
78. Lin DH, Lin CHJ, Lin YF, Jan MH. Efficacy of 2 non-weight-bearing interventions, proprioception training versus strength training, for patients with knee osteoarthritis: a randomized clinical trial. *J Orthop Sports Phys Ther* 2009;39:450-7.
79. Schilke JM, Johnson GO, Housh TJ, O'Dell JR. Effects of muscle-strength training on the functional status of patients with osteoarthritis of the knee joint. *Nurs Res* 1996;45:68-72.
80. Bilgici A, Akdeniz O, Kuru O, Ulusoy H. The effect of a home-based exercise therapy versus an aerobic exercise programme on pain and functional disability in patients with knee osteoarthritis. *Ann Rheum Dis* 2004;63:364-65.
81. Aglamis B, Toraman NF, Yaman H. Change of quality of life due to exercise training in knee osteoarthritis: SF-36 and WOMAC. *J Back Musculoskelet Rehabil* 2009;22:43-48.
82. Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. *Arthritis Rheum* 2004;50:1501-10.
83. Thorstensson CA, Roos EM, Petersson IF, Ekdahl C. Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis: a randomized controlled trial [ISRCTN20244858]. *BMC Musculoskelet Disord* 2005;6:27.
84. Dias RC, Dias JMD, Ramos LR. Impact of an exercise and walking protocol on quality of life for elderly people with OA of the knee. *Physiother Res Int* 2003;8:121-30.
85. Keefe FJ, Blumenthal J, Baucom D, Affleck G, Waugh R, Caldwell DS, et al. Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. *Pain* 2004;110:539-49.
86. Kovar PA, Allegrante JP, MacKenzie CR, Peterson MG, Gutin B, Charlson ME. Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. *Ann Intern Med* 1992;116:529-34.
87. Yip YB, Sit JW, Fung KKY, Wong DYS, Chong SYC, Chung LH, et al. Impact of an arthritis self-management programme with an added exercise component for osteoarthritic knee sufferers on improving pain, functional outcomes, and use of health care services: an experimental study. *Patient Educ Couns* 2007;65:113-21.
88. Wang C, Schmid CH, Hibberd PL, Kalish R, Roubenoff R, Rones R, et al. Tai chi is effective in treating knee osteoarthritis: a randomized controlled trial. *Arthritis Care Res* 2009;61:1545-53.
89. Ni GX, Song L, Yu B, Huang CH, Lin JH. Tai Chi improves physical function in older Chinese women with knee osteoarthritis. *J Clin Rheumatol* 2010;16:64-67. Retraction in: Schumacher HR. *J Clin Rheumatol* 2010;16:357.
90. Lee HJ, Park HJ, Chae Y, Kim SY, Kim SN, Kim ST, et al. Tai Chi Qigong for the quality of life of patients with knee osteoarthritis: a pilot, randomized, waiting list controlled trial. *Clin Rehabil* 2009;23:504-11.
91. Brismee JM, Paige RL, Chyu MC, Boatright JD, Hagar JM, McCaleb JA, et al. Group and home-based tai chi in elderly subjects with knee osteoarthritis: a randomized controlled trial. *Clin Rehabil* 2007;21:99-111.
92. Christensen R, Astrup A, Bliddal H. Weight loss: the treatment of choice for knee osteoarthritis? A randomized trial. *Osteoarthritis Cartilage* 2005;13:20-7.
93. Jenkinson CM, Doherty M, Avery AJ, Read A, Taylor MA, Sach TH, et al. Effects of dietary intervention and quadriceps strengthening exercises on pain and function in overweight people with knee pain: randomised controlled trial. *BMJ* 2009;339:b3170.

94. Miller GD, Nicklas BJ, Davis C, Loeser RF, Lenchik L, Messier SP. Intensive weight loss program improves physical function in older obese adults with knee osteoarthritis. *Obesity* 2006;14:1219-30.
95. Balint GP, Buchanan WW, Adam A, Ratko I, Poor L, Balint PV, et al. The effect of the thermal mineral water of Nagybaracska on patients with knee joint osteoarthritis - a double blind study. *Clin Rheumatol* 2007;26:890-4.
96. Fioravanti A, Iacoponi F, Bellisai B, Cantarini L, Galeazzi M. Short- and long-term effects of spa therapy in knee osteoarthritis. *Am J Phys Med Rehabil* 2010;89:125-32.
97. Sukenik S, Flusser D, Codish S, Abu-Shakra M. Balneotherapy at the Dead Sea area for knee osteoarthritis. *Isr Med Assoc J* 1999;1:83-5.
98. Cantarini L, Leo G, Giannitti C, Cevenini G, Barberini P, Fioravanti A. Therapeutic effect of spa therapy and short wave therapy in knee osteoarthritis: a randomized, single blind, controlled trial. *Rheumatol Int* 2007;27:523-9.
99. Yurtkuran M, Yurtkuran M, Alp A, Nasircilar A, Bingol U, Altan L, et al. Balneotherapy and tap water therapy in the treatment of knee osteoarthritis. *Rheumatol Int* 2006;27:19-27.
100. Kovacs I, Bender T. The therapeutic effects of Cserkeszolo thermal water in osteoarthritis of the knee: a double blind, controlled, follow-up study. *Rheumatol Int* 2002;21:218-21.
101. Sherman G, Zeller L, Avriel A, Friger M, Harari M, Sukenik S. Intermittent balneotherapy at the Dead Sea area for patients with knee osteoarthritis. *Isr Med Assoc J* 2009;11:88-93.
102. Karagulle M, Karagulle MZ, Karagulle O, Donmez A, Turan M. A 10-day course of SPA therapy is beneficial for people with severe knee osteoarthritis. A 24-week randomised, controlled pilot study. *Clin Rheumatol* 2007;26:2063-71.
103. Nguyen M, Revel M, Dougados M. Prolonged effects of 3 week therapy in a spa resort on lumbar spine, knee and hip osteoarthritis: follow-up after 6 months. A randomized controlled trial. *Br J Rheumatol* 1997;36:77-81.
104. Mahboob N, Sousan K, Shirzad A, Amir G, Mohammad V, Reza M, et al. The efficacy of a topical gel prepared using Lake Urmia mud in patients with knee osteoarthritis. *J Altern Complement Med* 2009;15:1239-42.
105. Forestier R, Desfour H, Tessier JM, Francon A, Foote AM, Genty C, et al. Spa therapy in the treatment of knee osteoarthritis: a large randomised multicentre trial. *Ann Rheum Dis* 2010;69:660-5.
106. Wigler I, Elkayam O, Paran D, Yaron M. Spa therapy for gonarthrosis: a prospective study. *Rheumatol Int* 1995;15:65-8.
107. Flusser D, Abu-Shakra M, Friger M, Codish S, Sukenik S. Therapy with mud compresses for knee osteoarthritis: comparison of natural mud preparations with mineral-depleted mud. *J Clin Rheumatol* 2002;8:197-203.
108. Tishler M, Rosenberg O, Levy O, Elias I, Amit-Vazina M. The effect of balneotherapy on osteoarthritis. Is an intermittent regimen effective? *Eur J Intern Med* 2004;15:93-96.
109. Barrios JA, Crenshaw JR, Royer TD, Davis IS. Walking shoes and laterally wedged orthoses in the clinical management of medial tibiofemoral osteoarthritis: a one-year prospective controlled trial. *Knee* 2009;16:136-42.
110. Rodrigues PT, Ferreira AF, Pereira RMR, Bonfa E, Borba EF, Fuller R. Effectiveness of medial-wedge insole treatment for valgus knee osteoarthritis. *Arthritis Care Res* 2008;59:603-8.
111. Baker K, Goggins J, Xie H, Szumowski K, LaValley M, Hunter DJ, et al. A randomized crossover trial of a wedged insole for treatment of knee osteoarthritis. *Arthritis Rheum* 2007;56:1198-203.
112. Toda Y, Tsukimura N. Influence of concomitant heeled footwear when wearing a lateral wedged insole for medial compartment osteoarthritis of the knee. *Osteoarthritis Cartilage* 2008;16:244-53.
113. Maillefert JF, Hudry C, Baron G, Kieffert P, Bourgeois P, Lechevalier D, et al. Laterally elevated wedged insoles in the treatment of medial knee osteoarthritis: a prospective randomized controlled study. *Osteoarthritis Cartilage* 2001;9:738-45.
114. Nigg BM, Emery C, Hiemstra LA. Unstable shoe construction and reduction of pain in osteoarthritis patients. *Med Sci Sports Exerc* 2006;38:1701-8.
115. Chen CY, Chen CL, Hsu SCC, Chou SW, Wang KC. Effect of magnetic knee wrap on quadriceps strength in patients with symptomatic knee osteoarthritis. *Arch Phys Med Rehabil* 2008;89:2258-64.
116. Hinman MR, Ford J, Heyl H. Effects of static magnets on chronic knee pain and physical function: a double-blind study. *Altern Ther Health Med* 2002;8:50-5.
117. Wolsko PM, Eisenberg DM, Simon LS, Davis RB, Walleczek J, Mayo-Smith M, et al. Double-blind placebo-controlled trial of static magnets for the treatment of osteoarthritis of the knee: results of a pilot study. *Altern Ther Health Med* 2004;10:36-43.
118. Auerbach B, Melzer C. [Cross-linked hyaluronic acid in the treatment of osteoarthritis of the knee - results of a prospective randomized trial]. *Zentralbl Chir* 2002;127:895-9.

119. Kirkley A, Webster-Bogaert S, Litchfield R, Amendola A, MacDonald S, McCalden R, et al. The effect of bracing on varus gonarthrosis. *J Bone Joint Surg Am* 1999;81:539-48.
120. Brouwer RW, van Raaij TM, Verhaar JAN, Coene LNJEM, Bierma-Zeinstra SMA. Brace treatment for osteoarthritis of the knee: a prospective randomized multi-centre trial. *Osteoarthritis Cartilage* 2006;14:777-83.
121. Lewis B, Lewis D, Cumming G. The analgesic efficacy of transcutaneous electrical nerve stimulation (TENS) compared with a non-steroidal anti-inflammatory drug (naprosyn) in painful osteoarthritis (OA) of the knee. *Aust N Z J Med* 1988;18 Suppl 1:224.
122. Taylor P, Hallett M, Flaherty L. Treatment of osteoarthritis of the knee with transcutaneous electrical nerve stimulation. *Pain* 1981;11:233-40.
123. Law PPW, Cheing GLY. Optimal stimulation frequency of transcutaneous electrical nerve stimulation on people with knee osteoarthritis. *J Rehabil Med* 2004;36:220-5.
124. Cheing GLY, Tsui AYY, Lo SK, Hui-Chan CWY. Optimal stimulation duration of tens in the management of osteoarthritic knee pain. *J Rehabil Med* 2003;35:62-8.
125. Grimmer K. A controlled double blind study comparing the effects of strong burst mode TENS and High Rate TENS on painful osteoarthritic knees. *Aust J Physiother* 1992;38:49-56.
126. Pietrosimone BG, Hart JM, Saliba SA, Hertel J, Ingersoll CD. Immediate effects of transcutaneous electrical nerve stimulation and focal knee joint cooling on quadriceps activation. *Med Sci Sports Exerc* 2009;41:1175-81.
127. Burch FX, Tarro JN, Greenberg JJ, Carroll WJ. Evaluating the benefits of patterned stimulation in the treatment of osteoarthritis of the knee: a multi-center, randomized, single-blind, controlled study with an independent masked evaluator. *Osteoarthritis Cartilage* 2008;16:865-72.
128. Selfe TK, Bourguignon C, Taylor AG. Effects of noninvasive interactive neurostimulation on symptoms of osteoarthritis of the knee: a randomized, sham-controlled pilot study. *J Altern Complement Med* 2008;14:1075-81.
129. Kang RW, Lewis PB, Kramer A, Hayden JK, Cole BJ. Prospective randomized single-blinded controlled clinical trial of percutaneous neuromodulation pain therapy device versus sham for the osteoarthritic knee: a pilot study. *Orthopedics* 2007;30:439-45.
130. Smith CR, Lewith GT, Machin D. TNS and osteo-arthritic pain. Preliminary study to establish a controlled method of assessing transcutaneous nerve stimulation as a treatment for the pain caused by osteo-arthritis of the knee. *Physiotherapy* 1983;69:266-8.
131. Lewis B, Lewis D, Cumming G. The comparative analgesic efficacy of transcutaneous electrical nerve stimulation and a non-steroidal anti-inflammatory drug for painful osteoarthritis. *Br J Rheumatol* 1994;33:455-60.
132. Alcidi L, Beneforti E, Maresca M, Santosuosso U, Zoppi M. Low power radiofrequency electromagnetic radiation for the treatment of pain due to osteoarthritis of the knee. *Reumatismo* 2007;59:140-5.
133. Lewis D, Lewis B, Sturrock RD. Transcutaneous electrical nerve stimulation in osteoarthrosis: a therapeutic alternative? *Ann Rheum Dis* 1984;43:47-9.
134. Bai L, Cheng L, Wang J, Wang G, Wang Y, Lv D, et al. [Application of arthroscopic RIGIDfix cross pins and Intrafix screw for posterior cruciate ligament reconstruction]. *Chinese Journal of Reparative & Reconstructive Surgery* 2008;22:234-8.
135. Fukuda TY, Ovanessian V, Cunha RAD, Filho ZJ, Cazarini C, Rienzo FA, et al. Pulsed short wave effect in pain and function in patients with knee osteoarthritis. *J Appl Res* 2008;8:189-98.
136. Zizic TM, Hoffman KC, Holt PA, Hungerford DS, O'Dell JR, Jacobs MA, et al. The treatment of osteoarthritis of the knee with pulsed electrical stimulation. *J Rheumatol* 1995;22:1757-61.
137. Garland D, Holt P, Harrington JT, Caldwell J, Zizic T, Cholewczynski J. A 3-month, randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of a highly optimized, capacitively coupled, pulsed electrical stimulator in patients with osteoarthritis of the knee. *Osteoarthritis Cartilage* 2007;15:630-7.
138. Callaghan MJ, Whittaker PE, Grimes S, Smith L. An evaluation of pulsed shortwave on knee osteoarthritis using radioleucoscintigraphy: a randomised, double blind, controlled trial. *Joint Bone Spine* 2005;72:150-55.
139. Efficacy of pulsed electromagnetic therapy in painful knee osteoarthritis. Proceedings of the 62nd Annual Meeting of the American College of Rheumatology; 1998; San Diego. American College of Rheumatology.
140. Jacobson JI, Gorman R, Yamanashi WS, Saxena BB, Clayton L. Low-amplitude, extremely low frequency magnetic fields for the treatment of osteoarthritic knees: a double-blind clinical study. *Altern Ther Health Med* 2001;7:54-64.

141. Pipitone N, Scott DL. Magnetic pulse treatment for knee osteoarthritis: a randomised, double-blind, placebo-controlled study. *Curr Med Res Opin* 2001;17:190-6.
142. Thamsborg G, Florescu A, Oturai P, Fallentin E, Tritsaris K, Dissing S. Treatment of knee osteoarthritis with pulsed electromagnetic fields: a randomized, double-blind, placebo-controlled study. *Osteoarthritis Cartilage* 2005;13:575-81.
143. Trock DH, Bollet AJ, Markoll R. The effect of pulsed electromagnetic fields in the treatment of osteoarthritis of the knee and cervical spine. Report of randomized, double blind, placebo controlled trials. *J Rheumatol* 1994;21:1903-11.
144. Talbot LA, Gaines JM, Ling SM, Metter EJ. A home-based protocol of electrical muscle stimulation for quadriceps muscle strength in older adults with osteoarthritis of the knee. *J Rheumatol* 2003;30:1571-8.
145. Young SL, Woodbury MG, Fryday-Field K, Donovan T, Bellamy N, Haddad R. Efficacy of interferential current stimulation alone for pain reduction in patients with osteoarthritis of the knee: a randomized placebo control clinical trial. *Phys Ther* 1991;71 Suppl:852.
146. Defrin R, Ariel E, Peretz C. Segmental noxious versus innocuous electrical stimulation for chronic pain relief and the effect of fading sensation during treatment. *Pain* 2005;115:152-60.
147. Adedoyin RA, Olaogun MOB, Fagbeja OO. Effect of interferential current stimulation in management of osteo-arthritic knee pain. *Physiotherapy* 2002;88:493-99.
148. Mazzuca SA, Page MC, Meldrum RD, Brandt KD, Petty-Saphon S. Pilot study of the effects of a heat-retaining knee sleeve on joint pain, stiffness, and function in patients with knee osteoarthritis. *Arthritis Care Res* 2004;51:716-21.
149. Clarke GR, Willis LA, Stenner L, Nichols PJR. Evaluation of physiotherapy in the treatment of osteoarthrosis of the knee. *Rheum Rehab* 1974;13:190-97.
150. Rattanachaiyanont M, Kuptniratsaikul V. No additional benefit of shortwave diathermy over exercise program for knee osteoarthritis in peri-/post-menopausal women: an equivalence trial. *Osteoarthritis Cartilage* 2008;16:823-8.
151. Stelian J, Gil I, Habot B, Rosenthal M, Abramovici I, Kutok N, et al. Improvement of pain and disability in elderly patients with degenerative osteoarthritis of the knee treated with narrow-band light therapy. *J Am Geriatr Soc* 1992;40:23-6.
152. Tascioglu F, Armagan O, Tabak Y, Corapci I, Oner C. Low power laser treatment in patients with knee osteoarthritis. *Swiss Med Wkly* 2004;134:254-8.
153. Nivbrant B, Friberg S. Therapeutic laser treatment in gonarthrosis. *Acta Orthop Scand* 1989;60 Suppl 231:33.
154. Gur A, Cosut A, Sarac AJ, Cevik R, Nas K, Uyar A. Efficacy of different therapy regimes of low-power laser in painful osteoarthritis of the knee: a double-blind and randomized-controlled trial. *Lasers Surg Med* 2003;33:330-8.
155. 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 Med Sci* 2009;24:129-36.
156. Bulow PM, Jensen H, Danneskiold-Samsoe B. Low power Ga-Al-As laser treatment of painful osteoarthritis of the knee. A double-blind placebo-controlled study. *Scand J Rehabil Med* 1994;26:155-9.
157. Kang T, Kim S, Kim S. The efficacy of low power laser therapy in patients with knee osteoarthritis. *Ann Rheum Dis* 2006;65 Suppl 2:233.
158. Perlman AI, Sabina A, Williams AL, Njike VY, Katz DL. Massage therapy for osteoarthritis of the knee: a randomized controlled trial. *Arch Intern Med* 2006;166:2533-8.
159. Tucker M, Brantingham JW, Myburg C. Relative effectiveness of a non-steroidal anti-inflammatory medication (Meloxicam) versus manipulation in the treatment of osteo-arthritis of the knee. *European Journal of Chiropractic* 2003;50:163-83.
160. Pollard H, Ward G, Hoskins W, Hardy K. The effect of a manual therapy knee protocol on osteoarthritic knee pain: a randomised controlled trial. *J Can Chiropr Assoc* 2008;52:229-42.
161. Moss P, Sluka K, Wright A. The initial effects of knee joint mobilization on osteoarthritic hyperalgesia. *Man Ther* 2007;12:109-18.
162. Jamtvedt G, Dahm K, Christie A, Moe R, Haavardsholm E, Holm I, et al. Physical therapy interventions for patients with osteoarthritis of the knee: An overview of systematic reviews. *Phys Ther* 2008;88:123-36.
163. 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.
164. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis

index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000;27:2635-41.
165. Bellamy N, Carette S, Ford PM, Kean WF, le Riche NGH, Lussier A, et al. Osteoarthritis antirheumatic drug trials. III. Setting the delta for clinical trials - results of a consensus development (Delphi) exercise. *J Rheumatol* 1992;19:451-7.
166. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005;64:29-33.
167. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009;146:238-44.
168. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, et al. OARSI recommendations for the management of hip and knee osteoarthritis. Part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage* 2010;18:476-99.
169. Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA* 2003;290:1062-70.
170. Dieppe PA. Recommended methodology for assessing the progression of osteoarthritis of the hip and knee joints. *Osteoarthritis Cartilage* 1995;3:73-77.
171. Sun Y, Sturmer T, Gunther KP, Brenner H. Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee - a review of the literature. *Clin Rheumatol* 1997;16:185-98.
172. Andersson S, Lundeberg T. Acupuncture-from empiricism to science: functional background to acupuncture effects in pain and disease. *Med Hypotheses* 1995;45:271-81.
173. 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;335:436.
174. Bennell KL, Bowles KA, Payne C, Cicuttini F, Williamson E, Forbes A, et al. Lateral wedge insoles for medial knee osteoarthritis: 12 month randomised controlled trial. *BMJ* 2011;342:d2912

10. APPENDICES

10.1 Literature search strategy

The base search strategy to locate clinical trials was designed for Ovid MEDLINE and translated for all other databases searched.

1 Osteoarthritis, Knee/
2 (gonarthrosis or gonarthritis).ti,ab.
3 1 or 2
4 Osteoarthritis/
5 (Osteoarthriti$ or OA or osteo arthriti$ or osteoarthros$ or osteo arthros$ or arthropath$ or arthrosis or arthroses).ti,ab.
6 degenerative arthriti$.ti,ab.
7 degenerative joint disease.ti,ab.
8 4 or 5 or 6 or 7
9 Knee/
10 Knee Joint/
11 (knee$ or patella$ or knee cap$ or kneecap$ or femorotibial or femoro tibial or tibiofemoral or tibio femoral or patellofemoral or patello femoral).ti,ab.
12 9 or 10 or 11
13 8 and 12
14 Arthralgia/
15 (arthralgi$ or (joint$ adj3 pain$)).ti,ab.
16 chronic pain$.ti,ab.
17 14 or 15 or 16
18 12 and 17
19 3 or 13 or 18
20 Acupuncture/
21 exp Acupuncture Therapy/
22 acupuncture$.ti,ab.
23 (electroacupuncture$ or electro acupuncture$).ti,ab.
24 (osteopuncture$ or osteo puncture$).ti,ab.
25 (perioste$ adj3 (stimulati$ or therap$ or needling)).ti,ab.
26 exp Physical Therapy Modalities/
27 (physiotherap$ or physio therap$ or physical therap$ or manual therap$).ti,ab.
28 (massage$ or acupressure or shiatsu or shiatzu or zhi ya or chih ya).ti,ab.
29 Chiropractic/
30 Traction/
31 (chiropractic or manipulat$ or traction or kinesiolog$ or mobilis$ or mobiliz$).ti,ab.
32 Osteopathic medicine/
33 osteopath$.ti,ab.
34 (hydrotherap$ or hydro therap$ or water therap$ or pool therap$).ti,ab.
35 exp Exercise/
36 exp Sports/
37 Physical Fitness/
38 (exercise$ or workout$ or work out$ or train$ or physical$ activ$ or kinesiotherap$ or keep$ fit or aerobics).ti,ab.
39 (muscle$ adj3 (stretch$ or strengthen$)).ti,ab.
40 (walk$ adj3 (fitness or aerobic or program$ or intervention$ or session$ or regime$)).ti,ab.
41 pedometer$.ti,ab.
42 (bicycl$ or cycle$ or cycling).ti,ab.
43 (run$ or jog$ or treadmill$).ti,ab.
44 (swim$ or water sport$ or aquatic$ or water aerobic$ or aqua aerobic$ or water gymnastics).ti,ab.
45 (tai ji or taiji or taijiquan or tai chi or t ai chi or taichi or shadow boxing).ti,ab.
46 (yoga or yogic or pilates or danc$).ti,ab.
47 (qigong or qi gong or chi kung or chikung or ch i kung).ti,ab.
48 (CPM or (passive adj (motion or movement))).ti,ab.

49 vibration/
50 (vibrati$ or mechanical stimul$).ti,ab.
51 (balneology or balneotherap$ or balneo therap$ or bath$ or crenobalneotherap$ or thalassotherap$ or spa or spas).ti,ab.
52 (thermotherap$ or thermo therap$ or hypertherm$ or hyper therm$ or diatherm$ or short wave or shortwave or ultrasonic or cryotherap$ or cryo therap$).ti,ab.
53 (heat or hot or ice or cold).ti,ab.
54 exp Cryotherapy/
55 exp Orthotic Devices/
56 (brace$ or bracing or orthotic$ or orthoses).ti,ab.
57 (insert$ or insole$).ti,ab.
58 (TENS or ALTENS).ti,ab.
59 (transcutaneous adj2 nerve stimulation).ti,ab.
60 (electroanalgesia or electro analgesia).ti,ab.
61 (electric$ nerve stimulation or electrostimulation or electro stimulation).ti,ab.
62 EMS.ti,ab.
63 ((muscle or electric$) adj3 stimulat$).ti,ab.
64 (neuromodulation or neuro modulation or neurostimulation or neuro stimulation).ti,ab.
65 interferential.ti,ab.
66 Electromagnetic Fields/
67 Magnetic Field Therapy/
68 PEMF.ti,ab.
69 ((electromagnetic$ or magnetic$) adj3 field$).ti,ab.
70 (biomagnetic$ or bio magnetic$ or pulsed signal).ti,ab.
71 Laser Therapy, Low-Level/
72 laser.ti,ab.
73 phototherapy/
74 (light or phototherap$ or photo therap$).ti,ab.
75 or/20-74
76 19 and 75
77 exp Obesity/
78 Overweight/
79 body mass index/
80 (obese or obesity).ti,ab.
81 (overweight or over weight).ti,ab.
82 Weight Loss/
83 Weight Gain/
84 (weight adj3 los$).ti,ab.
85 (weight adj3 reduc$).ti,ab.
86 (weight adj3 decreas$).ti,ab.
87 (weight adj3 gain$).ti,ab.
88 (weight adj3 increas$).ti,ab.
89 (weight adj3 chang$).ti,ab.
90 (BMI or body mass index).ti,ab.
91 or/77-90
92 Bariatrics/
93 exp Diet/
94 exp Diet Therapy/
95 (diet$ or slim or slimming).ti,ab.
96 (weight adj3 control$).ti,ab.
97 (weight adj3 manage$).ti,ab.
98 low calorie$.ti,ab.
99 calorie control$.ti,ab.
100 (calorie adj3 count$).ti,ab.
101 (caloric adj3 restrict$).ti,ab.
102 (calorie$ adj3 restrict$).ti,ab.
103 (energy adj3 restrict$).ti,ab.
104 (protein adj3 restrict$).ti,ab.
105 (weight watchers or weightwatchers or slimfast or nutrition class$ or meal replacement$).ti,ab.
106 or/92-105

107 exp Anti-Obesity Agents/
108 ((anti obes$ or antiobes$) adj3 (agent$ or drug$ or therap$ or medicine$)).ti,ab.
109 appetite suppressant$.ti,ab.
110 appetite depressant$.ti,ab.
111 (orlistat or xenical).ti,ab,rn.
112 phentermine.ti,ab,rn.
113 Phentermine/
114 (sibutramine or reductil).ti,ab,rn.
115 (rimonabant or acomplia).ti,ab,rn.
116 or/107-115
117 106 or 116
118 19 and 91 and 117
119 76 or 118
120 randomized controlled trial.pt.
121 controlled clinical trial.pt.
122 randomized.ab.
123 placebo.ab.
124 drug therapy.fs.
125 randomly.ab.
126 trial.ab.
127 groups.ab.
128 120 or 121 or 122 or 123 or 124 or 125 or 126 or 127
129 animals/ not (animals/ and humans/)
130 128 not 129
131 119 and 130

Full details of all the databases searched, search strategies and results can be found at: www.york.ac.uk/inst/crd/Documents/OAKSearchStrategiesWebLink.docx

10.2 Quality assessment

Full results are available at:
http://www.york.ac.uk/inst/crd/Documents/StudyQualityAssessmentResults.docx

1 Was the number of patients randomised to treatment stated?	1.1 Yes 1.2 No
2 Method of randomisation appropriate?	2.1 Yes 2.2 No 2.3 Unclear/not stated
3 Appropriate allocation concealment?	3.1 Yes 3.2 No 3.3 Unclear/not stated
4 Appropriate type of placebo used?	4.1 Yes 4.2 No (give brief details) 4.3 Unclear/not stated 4.4 Not applicable
5 Group baseline characteristics comparable?	5.1 Yes 5.2 No

	5.3 Unclear/not stated
6 Described as double blind?	6.1 Yes 6.2 No 6.3 Unclear/not stated
7 Treatment giver blinded?	7.1 Yes 7.2 No 7.3 Unclear/not stated
8 Patient blinded?	8.1 Yes 8.2 No 8.3 Unclear/not stated
9 Outcome assessor blinded?	9.1 Yes 9.2 No 9.3 Unclear/not stated
10 Use of a power calculation reported?	10.1 Yes 10.2 No
11 Eligibility criteria adequately reported?	11.1 Yes 11.2 No
12 Clear reporting of losses to follow up?	12.1 Yes 12.2 No
13 Intention-to-treat data reported (analysed)?	13.1 Yes 13.2 No 13.3 Unclear
14 At least 90% full follow up achieved?	14.1 Yes 14.2 No 14.3 Unclear
15 Overall study quality (see method for rating below)	15.1 Excellent 15.2 Good 15.3 Satisfactory 15.4 Poor

Method for rating overall study quality:
Excellent: answers Yes for 1-3, 5, 6, 8-14 and Yes or Not Applicable for 4
Good : answers Yes for 1, 5, 8, 9, 11-14 and Yes or unclear/not stated for questions 2, 3, 4 (or Not applicable for 4)
Satisfactory: answers Yes for 1, 5, 11, 12, 13 and Not ‘No’ for 4)
Poor: The answer is not Yes for one of the criteria required for ‘Satisfactory’.

10.3 Statistical Formulae

Derivation of standardised mean differences

The Hedges-g standardised mean difference (SMD) between treatment A treatment B and was derived as follows (Reference Lipsey):

$$SMD = \left(1 - \frac{3}{4N - 9}\right) d$$

where

$$d = \frac{\overline{X}_A - \overline{X}_B}{S_P}$$

$\overline{X}_A$ is the treatment effect estimate of treatment A, $\overline{X}_B$ is the treatment effect estimate of treatment B, S_P is the pooled standard deviation across the trial arms.

10.3.1 Derivation of standard deviations

From the standard error:

$$s = se\sqrt{n}$$

10.3.2 Derivation of standard errors

From a 95% confidence interval:

$$se = \frac{I_u - \overline{X}}{t}$$

Where *t* is the student's t-value for the sample size, I_u is the upper limit of the confidence interval, and $\overline{X}$ is the estimate of the mean.

10.3.3 Derivation of pooled standard deviations

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_m - 1)s_m^2}{n_1 + n_2 + \cdots + n_m - m}}$$

Where *m* is the number of trial arms.

10.4 Network meta-analysis appendix

10.4.1 Table of all trials included in NMA

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
Acupuncture											
Bao 2007 [48]	Standard care	40 Y	China	63	62	NR	NR	General	Clinical	NR/UC	Poor
Berman 1999 [42]	Standard care	73 Y	USA	60	65	NR	mean 32	General	C+R	2 or higher	Poor
Itoh 2008[35]	TENS Standard care	18 Y	Japan	66	range 62-83	NR	NR	General	C+R	2 or higher	Poor
Itoh 2008 [38]	Sham acupuncture	24 Y	Japan	77	73	NR	NR	General	C+R	2 or higher	Poor
Lansdown 2009 [41]	Standard care	30 Y	UK	60	64	NR	NR	General	Clinical	NR/UC	Poor
Lu 2010[33]	Sham acupuncture	20 Y	China	NR	64	mean 66	NR	Both knees affected	C+R	2 or 3	Satisfactory
Miller 2009 [40]	Sham acupuncture Placebo Sham acupuncture	55 Y	Israel	69	71	NR	NR	General	NR/UC	NR/UC	Satisfactory
Ng 2003 [36]	TENS Standard care	14(imputed) Y (only vs TENS)	China	96	85	NR	NR	General	Clinical	NR/UC	Poor
Petrou 1988 [49]	Sham acupuncture	31 Y	Hungary	74	62	mean 80	NR	General	Clinical	NR/UC	Poor
Takeda 1994[51]	Sham acupuncture	40 Y	Canada	50	62	mean 90	Mean 33	General	C+R	NR/UC	Poor
Tukmachi 2004 [43]	Standard care	29 Y	UK	83	61	NR	NR	General	C+R	2 or 3	Good
Vas 2004 [46]	Sham acupuncture	88 Y	Spain	84	67	NR	mean 33	General	C+R	Ahlback grade 1 or higher	Good
Weiner 2007 [52]	Sham acupuncture	88 Y	USA	55	71	NR	mean 32	General	C+R	2 or higher	Good
Williamson 2007[34]	Exercise - MSE Standard care	181 Y	UK	54	71	NR	mean 32	Awaiting knee surgery	C+R	NR/UC	Satisfactory
Witt 2005[45]	Sham acupuncture Standard care	285 Y	Germany	66	64	NR	mean 29	General	C+R	2 or higher	Satisfactory
Witt 2006 [44]	Standard care	342 Y	Germany	60	61	NR	NR	General	C+R	NR/UC	Satisfactory
Wu 2008[31]	Standard care	34(imputed) Y	China	63	62	NR	NR	General	C+R	NR/UC	Poor
Yurtkuran 1999[37]	Ice/cooling treatment	100 Y	Turkey	91	58	NR	NR	General	C+R	NR/UC	Poor Please edit

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
	TENS Placebo TENS										
Exercise (strengthening								General			
An 2008 [71]	Standard care	21 Y	China	100	65	NR	26	General	Clinical	ACR Criteria	Poor
Baker 2001[61]	Placebo	38 Y	USA	78	69	NR	32	General	C+R	NR/UC median 3	Poor
Bezalel 2010 [75]	Heat treatment	50 Y	Israel	74	74	NR	NR	General	Unclear/NR	NR/UC	Poor
Borjesson 1996 [54]	Standard care	68 Y	Sweden	50	64	83	NR	Awaiting knee surgery	C+R	Ahlback grade I-III	Satisfactory
Cheing 2002[74]	TENS Placebo	47 Y	China	89	63	67	28	General	C+R	2 or higher	Poor
Durmus 2007 [76]	PES	50 Y	Turkey	100	55	NR	33	General	C+R	3 or lower	Satisfactory
Ettinger 1997[62]	Exercise - Aerobic Standard care	364 Y	USA	70	69	NR	53% >30kg/m2	General	C+R	NR/UC	Poor
Gur 2002[55]	No treatment	23 Y	Turkey	NR	56	79	NR	Both knees affected	Radiological	2 or 3	Poor
Hasegawa 2010[77]	Standard care	28 Y	Japan	64	77	55	24	General	Clinical	NR/UC	Poor
Huang 2005 [56]	Standard care	98 Y	Taiwan	81	62	NR	NR	Both knees affected	C+R	Altman grade II	Poor
Hurley 2007[69]	Standard care	53 N*	UK	70	67	81	30	General	Clinical	NR/UC	Satisfactory
Jan 2008 [57]	No treatment	98 Y	Taiwan	81	63	63	NR	Both knees affected	C+R	3 or lower	Satisfactory
Kuptniratsaikul 2002[72]	Standard care	366 Y	Tailand	78	68	NR	11% were obese	General	Radiological	2 or 3	Poor
Lin 2009 [78]	Standard care	72 Y	Taiwan	69	63	62	NR		C+R	3 or lower	Satisfactory
Lund 2008 [59]	Standard care	79 Y	Denmark	78	68	75	NR	General	C+R	Lequesne (1-26)score. Mean score = 11.	Poor
Maurer 1999[67]	Standard care	98 Y	USA	42	65	85	NR	General	C+R	3 or lower	Poor
McCarthy 2004 [68]	Standard care	172 Y	UK	58	65	NR	30	General	C+R		Satisfactory
Peloquin 1999[66]	Standard care	124 Y	Canada	70	66	NR	30	General	C+R	3 or lower	Poor
Schilke 1996[79]	Standard care	Y 20	USA	85	66	NR	NR	General	Clinical	NR/UC	Poor
Topp 2002 [60]	Standard care	102 Y	USA	73	63	89	NR	General	C+R	NR/UC	Poor
Trans 2009[63]	Standard care	52 N*	Denmark	100	60	81	30	General	C+R	NR/UC	Satisfactory

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
Williamson 2007[34]	Acupuncture Standard care	181 Y	UK	54	71	NR	32	Awaiting surgery	C+R	NR/UC	Satisfactory
Aerobic Exercise											
Aglamis 2009 [81]	Standard care	25 Y	Turkey	100	56	NR	33	Both knees affected	C+R	2 or higher	Poor
Ettinger 1997[62]	Exercise - MSE Standard care	364 Y	USA	70	69	NR	53% >30kg/m2	General	C+R	NR/UC	Poor
Keefe 2004 [85]	Standard care	30(imputed) Y	USA	50	59	NR	NR	General	Clinical	NR/UC	Poor
Kovar 1992 [86]	Standard care	92 Y	USA	83	69	mean 77	NR	General	C+R	NR/UC	Poor
Messier 2004[82]	Weight loss Standard care	240 Y	USA	72	69	Mean 94	34	Overweight/obese	C+R	3 or lower	Satisfactory
								General			
Tai chi studies											
Brismee 2007[91]	Standard care	N=41 Y	USA	83	70	73	28	General	Clinical	NR	Satisfactory
Lee 2009[90]	Standard care	N=44 Y	Korea	93	69	61	26	General	C+R	2 or higher	Satisfactory
Yip 2007[87]	Standard care	N=182 N*	Hong Kong	84	65	NR	NR	General	Clinical	NR	Poor
Weight loss studies											
Jenkinson 2009[93]	Standard care	389 Y	UK	66	61	Median 93	Median 33	General	Clinical	4 or lower 41% had a K&L score of 0	Satisfactory
Messier 2004[82]	Exercise : Aerobic Standard care	240 Y	USA	72	69	94	34	General	C+R	3 or lower	Satisfactory
Miller 2006[94]	Standard care	74 Y	USA	26	70 (all >/= 60)	98	35	Overweight/obese	Clinical	NR/UC	Poor
Balneotherapy											
Balint 2007 [95]	Placebo	N=52 Y	Hungary	63	50 to 75*	NR	NR	Both knees affected	Clinical	NR	Poor
Fioravanti 2010 [96]	Standard care	N=80 Y	Italy	75	70	NR	*26*	Both knees affected	C+R	3 or lower	Satisfactory
Flusser 2002[107]	Placebo	N=58 Y	Israel	85	65	*76*	NR	General	C+R	2 or 3	Poor
Forestier 2010 [105]	Standard care	N=309 N*	France	47	64	NR	*30*	General	C+R	1 or higher	Poor
Mahoob 2009[104]	Placebo	N=50 Y	Iran	100	44 to 79++	NR	NR	General	Clinical	NR	Poor

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
Sherman 2009 [101]	Placebo	N=44 Y	Israel	80	67	NR	NR	General	C+R	1 or higher	Poor
Sukenik 1999 [97]	Balneotherapy *3* groups	N=36 Y	Israel	89	63	NR	NR	General	C+R	Lequesne index of severity	Poor
Wigler 1995[106]	Balneotherapy *Two 'doses'*	N=33 Y	Israel	Specify 88	Specify Mean 65	NR	NR	General	C+R	NR	Poor
Yurtkuran 2006 [99]	Placebo	*N=52 Y*	*Turkey*	97	54	*77*	*32*	General	C+R	2 or 3	Poor
Insoles											
Barrios 2009[109]	Placebo	66 Y	USA	56	62	NR	33	Varus or valgus malalignment	C+R	2 or higher	Satisfactory
Maillefert 2001 [113]	Placebo	156 Y	France	74	65	NR	29	General	C+R	2 or higher	Satisfactory
Rodrigues 2008 [110]	Placebo	30 Y	Brazil	100	62	NR	30	Varus or valgus malalignment	C+R	2 or higher	Good
Static magnets											
Hinman 2002[116]	Placebo	N=43 (only one study)	USA	60	63	NR	NR	General	Clinical	NR	Poor
Braces											
Brouwer 2006 [120]	Standard care	N=117 N*	Netherlands	83	50	59	NR	Varus or valgus / malalignment	C+R	Ahlbäck score >0	Poor
TENS											
Alcidi 2007[132]	Heat treatment	N=40 Y	Italy	85	66	NR	NR	Overweight/obese	C+R	NR	Poor
Cheing 2002[74]	Exercise - MSE Placebo	N=47 Y	China	89	63	67	28	General	C+R	2 or higher	Poor
Cheing 2003 [124]	Placebo	N=38 Y	China	89	66	66	NR	General	C+R	2 or higher	Poor
Grimmer 1992 [125]	Different intensities of TENS Placebo	N=60 Y	Australia	62	67	NR	NR	General	C+R	NR	Poor
Itoh 2008[35]	Acupuncture, standard care	N=12 Y (only one study)	Japan	66	62-83[+]	NR	NR	General	C+R	2 or higher	Poor
Kang 2007 [129]	Placebo	N=63 Y	USA	71	57	NR	NR	General	C+R	NR	Poor
Lewis 1994 [131]	Standard care	N=56 N*	Australia	58	66	NR	NR	General	C+R	NR	Poor
Ng 2003 [36]	Electro-	N=14	China	96	85	NR	NR	General	Clinical	NR	Poor

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
	acupuncture, Standard care	N*									
Pietrosimone 2009 [126]	Ice/cooling treatment No treatment	N=33 Y	USA	48	56	88	30	General	Radiological	NR	Poor
Selfe 2008 [128]	Placebo	N=37 Y	USA	68	67	NR	31	General	C+R	NR	Poor
Yurtkuran 1999[37]	Acupuncture, ice/cooling, Placebo	N=100 Y	Turkey	91	58	NR	NR	General	C+R	NR	Poor
PES											
Callaghan 2005 [138]	Placebo	21 Y	UK	48	60	NR	28	General	Radiological	3 or higher	Poor
Durmus 2007 [76]	Exercise - MSE	50 Y***	Turkey	100	55	NR	33	General	C+R	3 or lower	Satisfactory
Garland 2007[137]	Placebo	58 Y	USA	66	66	NR	31	General	C+R	3 or higher	Poor
Miranda-Filloy 2005[32]	Placebo	18(imputed) Y	Spain	80	Older than 40 years	NR	NR	General	C+R	3 or higher	Poor
PEMF											
Fischer 2005[27]	Placebo	69 Y	Slovenia	72	60	NR	29.3	General	Radiological	NR/UC	Poor
Jacobson 2001 [140]	Placebo	176 N*	USA	NR	NR	NR	NR	Unclear	NR/UC	NR/UC	Poor
Perrot 1998 [139]	Placebo	40 Y	France	80	69	NR	NR	General	NR/UC	ACR OA criteria.	Poor
Thamsborg 2005 [142]	Placebo	83 Y	Denmark	54	60	NR	27	General	C+R	1 or higher	Poor
NMES											
Talbot 2003[144]	Standard care	34 Y***	USA	79	mean 71	NR	30	General	C+R	1 or higher	Poor
Interferential therapy											
Adedoyin 2002 [147]	Placebo	26 (imputed) Y	Nigeria	67	59	mean 78	28	General	C+R	NR/UC	Poor
Defrin 2005 [146]	No treatment	55 (imputed) Y	Israel	NR	67	NR	NR	General	Clinical	NR/UC	Poor
Heat treatment											
Alcidi [132]	TENS	N=40 Y (only one study)	Italy	*85*	*66*	NR	NR	General	C+R	NR	Poor
Bezalel [75]	Exercise - MSE	N=50 Y (only one study)	Israel	*74*	*74*	NR	NR	General	NR	NR	Poor
Clarke [149]	Ice/cooling treatment, placebo	N=30 Y	UK	*69*	*61*	NR	NR	General	Radiological	NR	Poor

Author	Comparators	N (EOT*)	Country	% Female	Age (in years: mean, unless stated)	Weight (in kg: mean, median, or range)	BMI (Mean, median, or range)	Type of knee OA population	Method of Diagnosis	Kellgren & Lawrence score	Overall study quality
Mazzuca [148]	Placebo	N=51 Y	USA	*77*	*63*	NR	NR	General	C+R	2 or higher	Poor
Ice/cooling											
Clarke 1974[149]	Heat treatment	N=45 Y (only study)	UK	69	61	NR	NR	General	Radiological	NR	Poor
Pietrosimone 2009 [126]	TENS	N=33 Y	USA	48	56	88	30	General	Radiological	NR	Poor
Yurtkuran 1999[37]	TENS	N=100 Y	Turkey	91	58	NR	NR	General	C+R	NR	Poor
Laser/light Therapy											
Gur 2003[154]	Placebo	N=90 N*	Turkey	80	60	NR	30	General	C+R	2 or higher	Satisfactory
Shen 2009 [155]	Placebo	N=40 Y	China	90	58	NR	NR	General	C+R	2 or higher	Satisfactory
Stelian 1992[151]	Placebo	N=50 Y	Israel	68	68	NR	NR	Both knees affected	NR	NR	Poor
Tascioglu 2004[152]	Placebo	N=60 Y	Turkey	70	62	NR	29	Both knees affected	C+R	2 or 3	Satisfactory
Manual therapy											
Pollard 2008 [160]	Placebo	N=43 Y (only study)	Australia	NR	56	NR	NR	General	C+R	NR	Poor
Tucker 2003 [159]	Standard care	N=103 Y	South Africa	63	59	82	NR	General	C+R	NR	Poor

* No Means, ** Only change from baseline scores reported

10.4.2 **Network tables (for networks for which diagrams are not presented (WOMAC Index))**

Network Table 1: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, therapy-plus-adjunct intervention set, any-quality analysis.

	SC/UT	SC/UT+ EX	Acu/NoMed	Acu/UT	ExMu/UT	Hea/UT	Tai/UT	Wei/UT	ShAcu/UT
SC/UT			1	4			2		1
SC/UT+ EX				1	1		1	1	
Acu/NoMed	1								
Acu/UT	4	1							1
ExMu/UT		1				1			
Hea/UT					1				
Tai/UT	2	1							
Wei/UT		1							
ShAcu/UT	1			1					

See Table 3 for abbreviations

Network Table 2: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, therapy-plus-adjunct intervention set, higher-quality analysis.

	SC/UT	SC/UT+ EX	Acu/UT	ExMu/UT	Tai/UT	ShAcu/UT
SC/UT			2		1	1
SC/UT+ EX			1	1	1	
Acu/UT	2	1				
ExMu/UT		1				
Tai/UT	1	1				
ShAcu/UT	1					

See Table 3 for abbreviations

Network Table 3: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, grouped intervention set, any-quality analysis.

	Standard care	Balneotherapy	Muscle str. exercise	Physiotherapy treatment	Insoles	Tai Chi	Weight loss	Acupuncture	Placebo	Sham Acupuncture
Standard care			1	1		2	1	6		1
Balneotherapy									2	
Muscle str. exercise	1			1						
Physiotherapy treatment	1		1						2	
Insoles									1	
Tai Chi	2									
Weight loss	1									
Acupuncture	6									2
Placebo		2		2	1					
Sham Acupuncture	1							2		

Network Table 4: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, grouped intervention set, higher-quality analysis.

	Standard care	Muscle str. exercise	Tai Chi	Acupuncture	Sham Acupuncture
Standard care		1	2	3	1
Muscle str. exercise	1				
Tai Chi	2				
Acupuncture	3				1
Sham Acupuncture	1			1	

Network Table 5: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, therapy-only intervention set, any-quality analysis.

	Standard care	Balneotherapy	Muscle str. exercise	Heat treatment	Insoles	PES	Tai Chi	TENS	Weight loss	Acupuncture	Placebo	Sham Acupuncture
Standard care			1				2	1	1	6		1
Balneotherapy											2	
Muscle str. exercise	1			1								
Heat treatment			1									
Insoles											1	
PES											1	
Tai Chi	2											
TENS	1										1	
Weight loss	1											
Acupuncture	6											2
Placebo		2			1	1		1				
Sham Acupuncture	1									2		

Network Table 6: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the end of treatment, therapy-only intervention set, higher-quality analysis.

	Standard care	Muscle str. exercise	Tai Chi	Acupuncture	Sham Acupuncture
Standard care		1	2	3	1
Muscle str. exercise	1				
Tai Chi	2				
Acupuncture	3				1
Sham Acupuncture	1			1	

Network Table 7: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, therapy-plus-adjunct intervention set, any-quality analysis.

	SC/UT	SC/UT+ EX	Acu/NoMed	Acu/UT	ExMu/NoMed	ExMu/UT	Hea/UT	Tai/UT	Wei/UT	ShAcu/UT
SC/UT			1	4	1			2		1
SC/UT+ EX				1		1		1	1	
Acu/NoMed	1									
Acu/UT	4	1								1
ExMu/NoMed	1									
ExMu/UT		1					1			
Hea/UT						1				
Tai/UT	2	1								
Wei/UT		1								
ShAcu/UT	1			1						

See Table 3 for abbreviations

Network Table 8: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, therapy-plus-adjunct intervention set, higher-quality analysis.

	SC/UT	SC/UT+ EX	Acu/UT	ExMu/NoMed	ExMu/UT	Tai/UT	ShAcu/UT
SC/UT			2	1			1
SC/UT+ EX			1		1	1	
Acu/UT	2	1					
ExMu/NoMed	1						
ExMu/UT		1					
Tai/UT		1					
ShAcu/UT	1						

See Table 3 for abbreviations

Network Table 9: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, grouped intervention set, any-quality analysis.

	Standard care	Balneotherapy	Muscle str. exercise	Physiotherapy treatment	Insoles	Tai Chi	Weight loss	Acupuncture	Placebo	Sham Acupuncture
Standard care			2	1		2	1	6		1
Balneotherapy									2	
Muscle str. exercise	2			1						
Physiotherapy treatment	1		1						3	
Insoles									1	
Tai Chi	2									
Weight loss	1									
Acupuncture	6									2
Placebo		2		3	1					
Sham Acupuncture	1							2		

Network Table 10: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, grouped intervention set, higher-quality analysis.

	Standard care	Muscle str. exercise	Tai Chi	Acupuncture	Sham Acupuncture
Standard care		2	1	3	1
Muscle str. exercise	2				
Tai Chi	1				
Acupuncture	3				1
Sham Acupuncture	1			1	

Network Table 11: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, therapy-only intervention set, any-quality analysis.

	Standard care	Balneotherapy	Muscle str. exercise	Heat treatment	Insoles	Laser	PES	Tai Chi	TENS	Weight loss	Acupuncture	Placebo	Sham Acupuncture
Standard care			2					2	1	1	6		1
Balneotherapy												2	
Muscle str. exercise	2			1									
Heat treatment			1										
Insoles												1	
Laser												1	
PES												1	
Tai Chi	2												
TENS	1											1	
Weight loss	1												
Acupuncture	6												2
Placebo		2			1	1	1		1				
Sham Acupuncture	1										2		

Network Table 12: The intervention comparisons included in the connected network with acupuncture for WOMAC index, the three months from start of treatment, therapy-only intervention set, any-quality analysis.

	Standard care	Muscle str. exercise	Tai Chi	Acupuncture	Sham Acupuncture
Standard care		2	1	3	1
Muscle str. exercise	2				
Tai Chi	1				
Acupuncture	3				1
Sham Acupuncture	1			1	

10.4.3 **Network meta-analysis appendix tables**

NMA Appendix Table 1: The number of trials with each of the specified scales that would be used in the analyses for each treatment out of the 91 trials with adequate data

	Acu	Bal	Bra	Ex-Ae	Ex-Mu	Hea	Ice	Ins	Int	Las	Man	NMES	PES	PEMF	Mag	Tai	TENS	Wei	ShaAcu
AIMS VAS 0 - 10				2	1														
AIMS2 pain subscale (0-10 scale)					1														
AIMS-4 item				1															
Borg Scale 0-10					1														
KSS 10 pt Likert	1																		1
Likert						1	1												
Likert 1-6				1	1														
Likert Four graded pain scale (1-none,4-severe)	2						1										1		1
Lysholm pain scale (no pain =25, constant pain =0)	1																		
Lysholm scores	1																		
McGill												1							
NR VAS 0 - 10	1								1								1		
NRS VAS 0 - 10					1														
Numerical Pain Rating Scale and VAS 0-100											1								
Scale 1 (minimal) to 10 (maximal) and pain diary														1					
VAS 0 – 10	2	2	1		5			1	1	2	2		1				3		1
VAS 0 – 100	2	2			2	1	1							2		1	5		1
WOMAC																			
Yes - Likert 5	5	1		1	6	2				2				1		1		3	2
Yes - VAS 10	1	1		1	1								1		1		1		1
Yes - VAS 100	2	2			4			2					1						1
Yes - std mean (VAS 100)					2								1						
Yes - Not specified		1														2			
Total	18	9	1	6	25	4	3	3	2	4	3	1	4	4	1	4	11	3	8

NMA Appendix Table 2: The mean standard deviations across the all the trials and across trials of good or satisfactory quality only, and the mean standard deviations by main intervention, for the VAS 0-10 scale.

VAS 0-10 scale

	Number of trials	Mean SD Satisfactory + quality trials	Number of trials	Mean SD All trials
Acu	2	1.67	2	1.67
Bal			2	0.96
ExAe			1	1
ExMu	2	1.42	5	1.5
Ins	1	2.4	1	2.4
Int			2	2.16
Las			2	1.76
Man			2	1.62
PES	1	0.51	2	1.74
TENS			3	1.86
SC			3	2.67
P	1	2.7	12	1.84
ShAcu	1	1.2	1	1.2
Mean across treatments		1.62		1.75

See Table 3 for abbreviations

NMA Appendix Table 3: The mean standard deviations across the all the trials and across trials of good or satisfactory quality only, and the mean standard deviations by main intervention, for the Likert 5 scale.

Likert 5 scale

	Number of trials	Mean SD Satisfactory + trials	Number of trials	Mean SD All trials
Acu	3	3.64	5	3.49
Bal			1	4.13
ExAe	1	4.2	1	4.2
ExMu	4	3.47	6	3.44
Hea			2	4.34
Las	2	3.19	2	3.19
PEMF			1	3.11
Wei	2	4.2	3	3.63
SC	6	3.8	11	3.46
P	2	3.82	5	4.06
NoTr	1	3.4	1	3.4
ShAcu	2	4.53	2	4.53
Mean across treatments		3.77		3.67

See Table 3 for abbreviations

NMA Appendix Table 4: Sensitivity analyses on including or excluding the correlation between comparisons in multiple-arm trials and halving or not halving the comparator sample size for one of the therapy-plus-adjunct intervention analyses

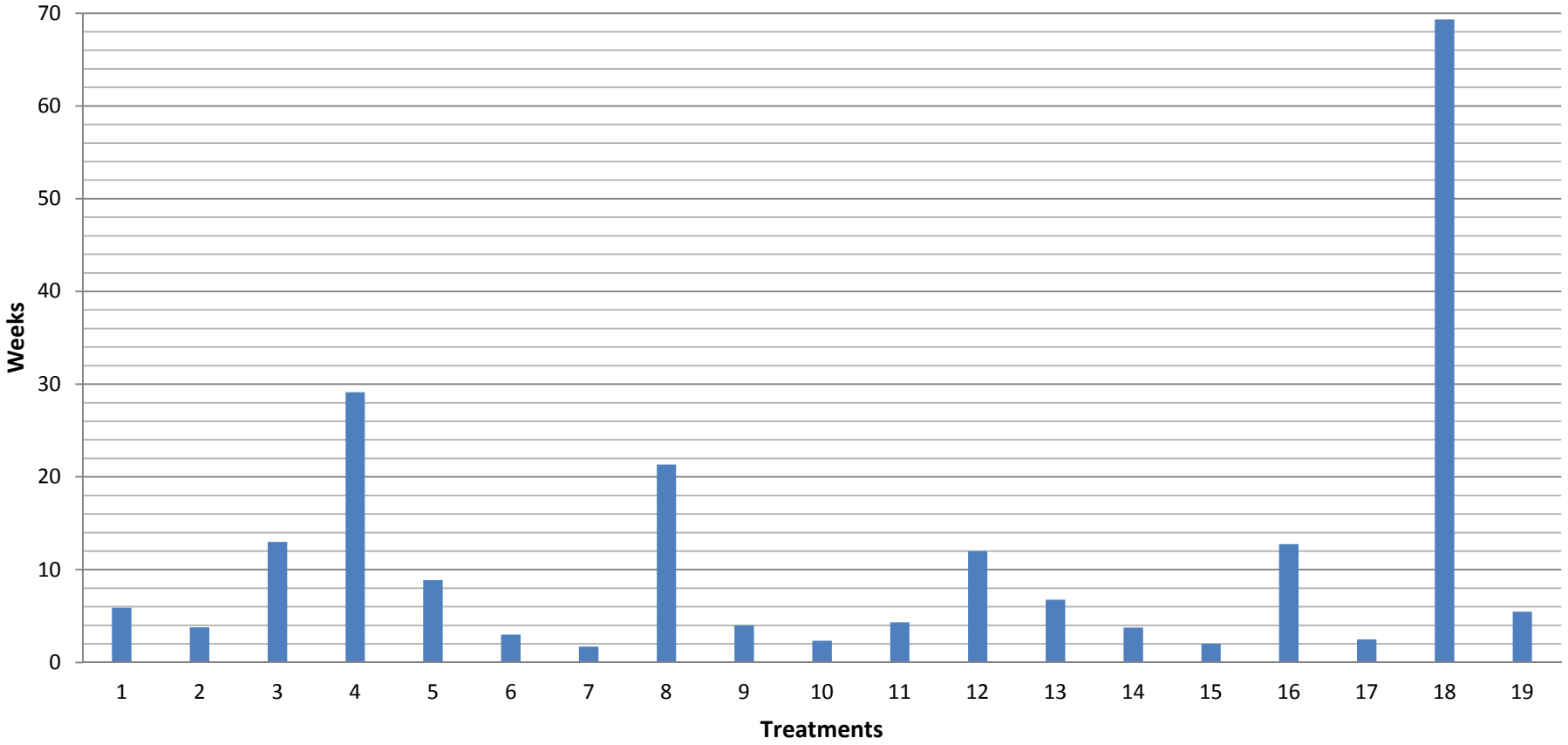
Comparator: standard care; standardised mean differences						
	Including correlation; full N		Including correlation; halve N		Excluding correlation; halve N	
Intervention	mean	sd	mean	sd	mean	sd
ExAe/NoMed	-0.24	0.21	-0.24	0.21	-0.26	0.21
Int/UT	-1.03	0.30	-1.03	0.31	-1.07	0.32
ExMu/NoMed	-0.56	0.64	-0.55	0.65	-0.54	0.69
ShAcu/NoMed	-0.21	0.21	-0.23	0.22	-0.12	0.24
Acu/NoMed	-0.44	0.72	-0.41	0.79	-0.34	0.82
TENS/NoMed	-0.01	0.67	-0.03	0.67	0.08	0.69
Ins/UT	-1.23	0.47	-1.25	0.55	-1.28	0.55
Mag/UT	-1.03	0.16	-1.02	0.17	-1.11	0.17
SC/UT+AN	-0.54	0.26	-0.56	0.26	-0.47	0.28
Acu/UT	-0.39	0.46	-0.39	0.46	-0.41	0.48
TENS/UT	-4.28	0.82	-4.29	0.81	-4.30	0.83
PES/UT	-0.49	0.25	-0.50	0.26	-0.48	0.28
Las/UT	-1.62	0.81	-1.59	0.88	-1.55	0.91
Ice/NoMed	-0.41	0.12	-0.42	0.12	-0.39	0.12
Man/UT	-0.66	0.35	-0.65	0.36	-0.67	0.37
Tai/UT	-0.25	0.29	-0.28	0.30	-0.19	0.32
ExMu/UT+EX	-0.88	0.83	-0.84	0.94	-0.74	1.05
Ice/UT	-0.59	0.38	-0.64	0.44	-0.56	0.48
Bal/UT	-1.05	0.57	-1.07	0.57	-0.96	0.59
P/NoMed	-1.86	0.80	-1.89	0.81	-1.78	0.92
ShAcu/UT	-0.85	0.41	-0.87	0.41	-0.76	0.43
Wei/UT+EX	-0.77	0.31	-0.82	0.32	-0.79	0.34
ExAe/UT	0.21	0.55	0.22	0.54	0.20	0.57
Wei/UT	-0.90	0.31	-0.91	0.31	-0.82	0.33
PEMF/NoMed	-0.45	0.75	-0.44	0.76	-0.43	0.81
NoTr/NoMed	-0.29	0.33	-0.31	0.34	-0.20	0.36
ExMu/UT	-1.03	0.53	-1.05	0.53	-0.95	0.56
Bra/UT	-0.32	0.39	-0.32	0.39	-0.32	0.42
Tai/NoMed	-0.79	0.32	-0.79	0.32	-0.81	0.33
PEMF/UT	-1.20	0.44	-1.19	0.46	-1.18	0.49
Hea/UT	-0.90	0.28	-0.94	0.29	-0.82	0.32
SC/UT+EX	-0.47	0.33	-0.47	0.34	-0.45	0.37
P/UT	-0.49	0.35	-0.50	0.35	-0.50	0.38
NoTr/UT	-1.28	0.70	-1.31	0.75	-1.34	0.77
NMES/UT+EX	-0.52	0.24	-0.50	0.24	-0.60	0.25
	Average	0.44	Average	0.46	Average	0.49

See Table 3 for abbreviations

NMA Appendix Table 5: Results for the grouped intervention higher-quality trials pain network – end of treatment time point

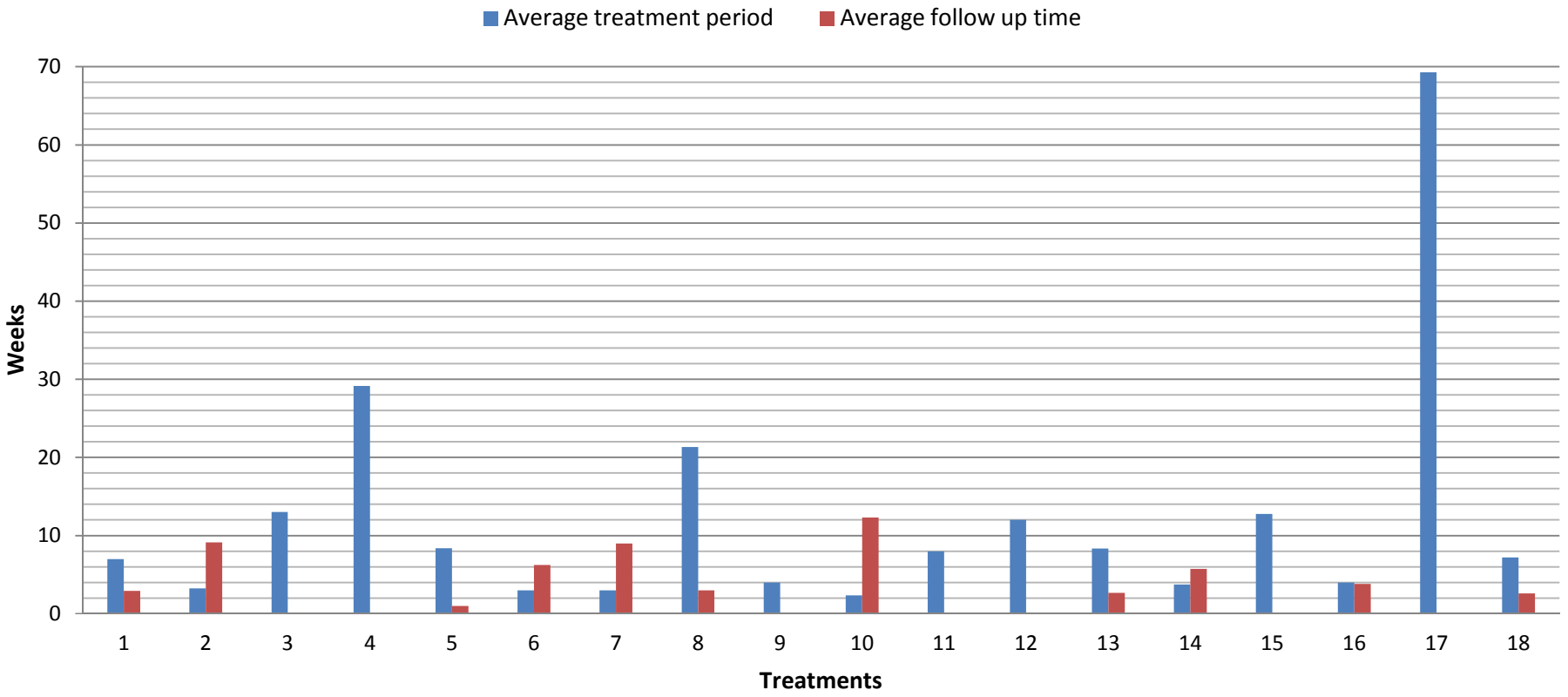
Intervention	Comparators																							
	Standard care		Placebo		No intervention		Acupuncture		Balneotherapy		Aerobic exercise		Muscle exercise		Physio treatment		Insoles		Tai Chi		Weight loss		Sham Acupuncture	
	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval	Mean	Credible interval
Standard care	0	0	1.72	(0.58 to 2.86)	0.33	(-1.20 to 0.54)	1.01	(0.64 to 1.40)	1.02	(0.20 to 1.82)	0.13	(-0.81 to 0.57)	0.38	(0.03 to 0.73)	1.57	(0.62 to 2.53)	1.73	(0.51 to 2.98)	0.26	(-0.38 to 0.91)	0.02	(-0.51 to 0.49)	0.47	(-0.02 to 0.96)
Placebo	-1.72	(-2.86 to -0.58)	0.00	(0.00 to 0.00)	2.06	(-3.40 to -0.71)	0.71	(-1.90 to 0.49)	0.71	(-2.11 to 0.70)	1.85	(-3.18 to -0.52)	1.35	(-2.43 to -0.25)	0.15	(-0.77 to 0.48)	0.01	(-0.46 to 0.51)	1.46	(-2.77 to -0.15)	1.74	(-2.98 to -0.50)	1.26	(-2.48 to -0.03)
No intervention	0.33	(-0.54 to 1.20)	2.06	(0.71 to 3.40)	0.00	(0.00 to 0.00)	1.35	(0.41 to 2.28)	1.35	(0.16 to 2.53)	0.21	(-0.91 to 1.31)	0.71	(-0.09 to 1.51)	1.91	(0.70 to 3.10)	2.07	(0.65 to 3.52)	0.60	(-0.48 to 1.68)	0.32	(-0.69 to 1.31)	0.80	(-0.19 to 1.78)
Acupuncture	-1.01	(-1.40 to -0.64)	0.71	(-0.49 to 1.90)	1.35	(-2.28 to -0.41)	0.00	(0.00 to 0.00)	0.00	(-0.90 to 0.89)	1.14	(-1.93 to -0.35)	0.64	(-1.13 to -0.15)	0.56	(-0.45 to 1.58)	0.72	(-0.55 to 2.01)	0.75	(-1.50 to -0.01)	1.03	(-1.65 to -0.40)	0.55	(-0.93 to -0.17)
Balneotherapy	-1.02	(-1.82 to -0.20)	0.71	(-0.70 to 2.11)	1.35	(-2.53 to -0.16)	0.00	(-0.89 to 0.90)	0.00	(0.00 to 0.00)	1.14	(-2.20 to -0.07)	0.64	(-1.51 to 0.25)	0.56	(-0.69 to 1.83)	0.72	(-0.74 to 2.22)	0.75	(-1.78 to 0.29)	1.03	(-1.98 to -0.07)	0.55	(-1.50 to 0.41)
Aerobic exercise	-0.13	(-0.57 to 0.81)	1.85	(0.52 to 3.18)	-0.21	(-1.31 to 0.91)	1.14	(0.36 to 1.93)	1.14	(0.07 to 2.20)	0.00	(0.00 to 0.00)	0.50	(-0.26 to 1.28)	1.70	(0.53 to 2.88)	1.86	(0.46 to 3.29)	0.39	(-0.55 to 1.33)	0.11	(-0.57 to 0.80)	0.59	(-0.26 to 1.44)
Muscle exercise	-0.38	(-0.73 to -0.03)	1.35	(0.25 to 2.43)	-0.71	(-1.51 to 0.09)	0.64	(0.15 to 1.13)	0.64	(-0.25 to 1.51)	0.50	(-1.28 to 0.26)	0.00	(0.00 to 0.00)	1.20	(0.30 to 2.09)	1.36	(0.17 to 2.56)	0.11	(-0.85 to 0.61)	0.39	(-1.00 to 0.21)	0.09	(-0.50 to 0.67)
Physio treatment	-1.57	(-2.53 to -0.62)	0.15	(-0.48 to 0.77)	-1.91	(-3.10 to -0.70)	0.56	(-1.58 to 0.45)	0.56	(-1.83 to 0.69)	1.70	(-2.88 to -0.53)	1.20	(-2.09 to -0.30)	0.00	(0.00 to 0.00)	0.16	(-0.61 to 0.96)	1.31	(-2.47 to -0.16)	1.59	(-2.67 to -0.51)	1.11	(-2.18 to -0.05)
Insoles	-1.73	(-2.98 to -0.51)	0.01	(-0.51 to 0.46)	-2.07	(-3.52 to -0.65)	0.72	(-2.01 to 0.55)	0.72	(-2.22 to 0.74)	1.86	(-3.29 to -0.46)	1.36	(-2.56 to -0.17)	0.16	(-0.96 to 0.61)	0.00	(0.00 to 0.00)	1.47	(-2.88 to -0.09)	1.75	(-3.10 to -0.43)	1.27	(-2.60 to 0.04)
Tai Chi	-0.26	(-0.91 to 0.38)	1.46	(0.15 to 2.77)	-0.60	(-1.68 to 0.48)	0.75	(0.01 to 1.50)	0.75	(-0.29 to 1.78)	0.39	(-1.33 to 0.55)	0.11	(-0.61 to 0.85)	1.31	(0.16 to 2.47)	1.47	(0.09 to 2.88)	0.00	(0.00 to 0.00)	0.28	(-1.09 to 0.54)	0.20	(-0.61 to 1.01)
Weight loss	-0.02	(-0.49 to 0.51)	1.74	(0.50 to 2.98)	-0.32	(-1.31 to 0.69)	1.03	(0.40 to 1.65)	1.03	(0.07 to 1.98)	0.11	(-0.80 to 0.57)	0.39	(-0.21 to 1.00)	1.59	(0.51 to 2.67)	1.75	(0.43 to 3.10)	0.28	(-0.54 to 1.09)	0.00	(0.00 to 0.00)	0.48	(-0.22 to 1.18)
Sham Acupuncture	-0.47	(-0.96 to 0.02)	1.26	(0.03 to 2.48)	-0.80	(-1.78 to 0.19)	0.55	(0.17 to 0.93)	0.55	(-0.41 to 1.50)	0.59	(-1.44 to 0.26)	0.09	(-0.67 to 0.50)	1.11	(0.05 to 2.18)	1.27	(-0.04 to 2.60)	0.20	(-1.01 to 0.61)	0.48	(-1.18 to 0.22)	0.00	(0.00 to 0.00)

Appendix Figure 1: The average treatment duration for each of the main treatments with evidence available for the end of treatment time point analysis. The treatments are coded in the table with the number of trials for each treatment.



Trials	18	9	1	6	20	4	3	3	2	3	3	1	4	4	1	4	11	3	7
	Acu	Bal	Bra	Ex-Ae	Ex-Mu	Hea	Ice	Ins	Int	Las	Man	NMES	PES	PEMF	Mag	Tai	TENS	Wei	ShaAcu
Code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

Appendix Figure 2: The average treatment duration and average follow up time for each of the main treatments that would be entered into the 3 months from start of follow up time point analysis. The treatments are coded in the table with the number of trials for each treatment.



Trials	14	9	1	7	20	4	1	3	2	3	1	1	3	4	4	5	3	5
	Acu	Bal	Bra	Ex-Ae	Ex-Mu	Hea	Ice	Ins	Int	Las	Man	NMES	PES	PEMF	Tai	TENS	Wei	ShaAcu
Code	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

10.4.4 **Network meta-analysis: Lists of (reference numbers of the) studies included in each analysis**

10.4.4.1 *Pain outcome analyses*

End of treatment, therapy-plus-adjunct intervention set, any-quality studies

140	54	78	144	151	60	91	129
120	86	85	55	35	79	26a	149
131	159	52	87	38	28	26b	148
128	34	81	59	97	27	66	139
132	101	40	43	124	90	61	44
76	49	68	126	125	72	155	106
142	107	48	31	93	33	62	116
37	42	63	94	57	104	96	146
160	41	137	45	71	77	32	74
110	56	138	51	82	75	67	

End of treatment, therapy-plus-adjunct intervention set, higher-quality studies

76	45
54	93
34	82
78	90
52	33
40	91
68	96
63	44
43	

End of treatment, grouped intervention set, any-quality studies

140	54	42	63	31	93	33	62	106
120	86	41	137	94	57	104	96	116
131	159	56	138	45	71	77	32	146
95	34	78	152	51	82	75	67	36
128	99	85	144	151	60	91	129	147
76	101	52	55	35	79	26a	149	74
142	113	81	87	38	28	26b	148	
37	49	40	59	97	27	66	109	
160	107	68	43	124	90	61	139	
110	46	48	126	125	72	155	44	

End of treatment, grouped intervention set, higher-quality studies

76	78	43	33
110	52	45	91
54	40	93	155
34	68	57	96
113	63	82	109
46	152	90	44

End of treatment, therapy-only intervention set, any-quality studies

140	110	46	48	126	125	72	155	44
120	54	42	63	31	93	33	62	106
131	86	41	137	94	57	104	96	116
95	159	56	138	45	71	77	32	146
128	34	78	152	51	82	75	67	36
132	99	85	144	151	60	91	129	147
76	101	52	55	35	79	26a	149	74
142	113	81	87	38	28	26b	148	
37	49	40	59	97	27	66	109	
160	107	68	43	124	90	61	139	

End of treatment, therapy-only intervention set, higher-quality studies

76	68	90
54	63	33
34	43	91
46	45	96
78	93	44
52	57	
40	82	

Three months from the start of treatment, therapy-plus-adjunct intervention set, any-quality studies

120	101	68	51	27	155	106
95	107	48	35	90	105	146
128	42	137	38	72	62	154
132	41	144	97	104	96	74
76	56	87	93	77	32	
142	78	59	71	75	67	
110	85	43	82	91	149	
54	52	69	60	26a	148	
86	81	94	79	26b	139	
34	40	45	28	66	44	

Three months from the start of treatment, therapy-plus-adjunct intervention set, higher-quality studies

76	68	90
54	43	91
34	69	96
78	45	44
52	93	
40	82	

Three months from the start of treatment, grouped intervention set, any-quality studies

140	99	52	59	57	77	32	36
120	101	81	43	71	75	67	147
95	113	40	69	82	91	149	74
128	107	68	94	60	26a	148	
76	46	48	45	79	26b	109	
142	42	137	51	28	66	139	
110	41	152	35	27	155	44	
54	56	144	38	90	105	106	
86	78	55	97	72	62	146	
34	85	87	93	104	96	154	

Three months from the start of treatment, grouped intervention set, higher-quality studies

76	78	69	155
110	52	45	96
54	40	93	109
34	68	57	44
113	152	82	154
46	43	91	90

Three months from the start of treatment, therapy-only intervention set, any-quality studies

140	34	85	87	93	104	96	154
120	99	52	59	57	77	32	36
95	101	81	43	71	75	67	147
128	113	40	69	82	91	149	74
132	107	68	94	60	26a	148	
76	46	48	45	79	26b	109	
142	42	137	51	28	66	139	
110	41	152	35	27	155	44	
54	56	144	38	90	105	106	
86	78	55	97	72	62	146	

Three months from the start of treatment, therapy-only intervention set, higher-quality studies

76	40	57
54	68	82
34	43	91
99	69	96
78	45	44
52	93	

10.4.4.2 *Overall WOMAC score analyses*

End of treatment, therapy-plus-adjunct intervention set, any-quality studies

34	35
42	38
41	90
44	75
94	91
45	

End of treatment, therapy-plus-adjunct interventions set, higher-quality studies

34	90
44	91
45	

End of treatment, grouped intervention set, any-quality studies

42	95	137
41	128	94
44	110	90
45	34	75
35	99	91
38	46	

End of treatment, grouped intervention set, higher-quality studies

44	46
45	90
34	91

End of treatment, therapy-only intervention set, any-quality studies

42	95	137
41	128	94
44	110	90
45	34	75
35	99	91
38	46	

End of treatment, therapy-only intervention set, higher-quality studies

44	46
45	90
34	91

Three months from the start of treatment, therapy-plus-adjunct intervention set, any-quality studies

34	45
42	35
41	38
44	90
69	75
94	91

Three months from the start of treatment, therapy-plus-adjunct intervention set, higher-quality studies

34	45
44	91
69	

Three months from the start of treatment, grouped intervention set, any-quality studies

42	95	94
41	128	154
44	110	90
69	34	75
45	99	91
35	46	
38	137	

Three months from the start of treatment, grouped intervention set, higher-quality studies

44	34
69	46
45	91

Three months from the start of treatment, therapy-only intervention set, any-quality studies

42	95	94
41	128	154
44	110	90
69	34	75
45	99	91
35	46	
38	137	

Three months from the start of treatment, therapy-only intervention set, higher-quality studies

44	34
69	46
45	91

10.5 Network meta-analysis WinBUGS code

```
#Random effects model for multi-arm trials (any number of arms)

model{
for(i in 1:N){

      prec[i]<-1/var[i]
      diff[i]~dnorm(delta[i],prec[i])              # model
      delta[i] ~ dnorm(md[i],taud[i])              # trial-specific distributions
            md[i] <-  d[t[i]] - d[b[i]]  + sw[i]      # mean of distributions
             taud[i] <- tau *2*(nd[i])/(nd[i]+1)       #precision of distributions

dev[i]<-(diff[i]-delta[i])*(diff[i]-delta[i])/var[i]
 }

 #adjustment, multi-arm RCTs

      sw[1] <-0
      sw[2]<-((delta[1]  - d[t[1]] + d[b[1]])*equals(s[1],s[2]))/nd[2]


      for(i in 3:N){

            sw[i] <- ((delta[i-1]  - d[t[i-1]] + d[b[i-1]])*equals(s[i-1],s[i])+(delta[i-2]  - d[t[i-2]] + d[b[i-2]])*equals(s[i-2],s[i]))/nd[i]


            }

sumdev<-sum(dev[])                              # residual deviance

d[1]<-0
for (k in 2:NT){d[k] ~ dnorm(0,.0001) }                  #  vague priors for basic parameters

sd~dunif(0,2)                        #  vague prior for random effects standard deviation
tau<-1/pow(sd,2)
vr<1/tau                                  #  calculates the between study variance
      }
```

list(N=95, NT=36)

diff[]	var[]	t[]	b[]	nd[]	s[]
0.107898285	0.023267397	26	4	1	1
-0.149161681	0.034305608	11	2	1	2
-0.114197207	0.071545009	31	3	1	3
-0.242640196	0.108982733	32	5	1	4
.....					

```
#Random effects model adjusted to evaluate consistency

model{

#w is the treatment comparison being tested for inconsistency between the direct and indirect evidence

w<-4

for(i in 1:N){

        prec[i]<-1/var[i]
        diff[i]~dnorm(diff1[i],prec[i])                    # model

        #splits the direct and indirect evidence

        diff1[i]<-delta[i]*equals(con[i],w)+phi[i]*(1-equals(con[i],w))

        phi[i] ~ dnorm(md[i],taud[i])             # trial-specific distributions

     md[i] <-  d[t[i]] - d[b[i]]  + sw[i]                  # mean of distributions
                taud[i] <- tau *2*(nd[i])/(nd[i]+1)       #precision of distributions

 delta[i] ~ dnorm(md1[i],taud[i])             # trial-specific distributions


      md1[i] <-  d1[t[i]] - d1[b[i]]  + sw1[i]                  # mean of distributions
```

```
dev[i]<-(diff[i]-delta[i])*(diff[i]-delta[i])/var[i]
  }

#adjustment, multi-arm RCTs
#this retains the correlation between multiple arms

      sw[1] <-0
      sw[2]<-((phi[1]  - d[t[1]] + d[b[1]])*(1-equals(con[1],w))*equals(s[1],s[2]))/nd[2]+((delta[1]  - d1[t[1]] +
d1[b[1]])*equals(con[1],w)*equals(s[1],s[2]))/nd[2]

      sw1[1] <-0
      sw1[2]<-((delta[1]  - d1[t[1]] + d1[b[1]])*equals(con[1],w)*equals(s[1],s[2]))/nd[2]+((phi[1]  - d[t[1]] + d[b[1]])*(1-
equals(con[1],w))*equals(s[1],s[2]))/nd[2]


      for(i in 3:N){

            sw[i] <- ((phi[i-1]  - d[t[i-1]] + d[b[i-1]])*(1-equals(con[i-1],w))*equals(s[i-1],s[i])+(phi[i-2]  - d[t[i-2]] + d[b[i-2]])*(1-equals(con[i-
2],w))*equals(s[i-2],s[i]))/nd[i]+((delta[i-1]  - d1[t[i-1]] + d1[b[i-1]])*equals(con[i-1],w)*equals(s[i-1],s[i])+(delta[i-2]  - d1[t[i-2]] + d1[b[i-
2]])*equals(con[i-2],w)*equals(s[i-2],s[i]))/nd[i]


            sw1[i] <- ((delta[i-1]  - d1[t[i-1]] + d1[b[i-1]])*equals(con[i-1],w)*equals(s[i-1],s[i])+(delta[i-2]  - d1[t[i-2]] + d1[b[i-2]])*equals(con[i-
2],w)*equals(s[i-2],s[i]))/nd[i]+((phi[i-1]  - d[t[i-1]] + d[b[i-1]])*(1-equals(con[i-1],w))*equals(s[i-1],s[i])+(phi[i-2]  - d[t[i-2]] + d[b[i-2]])*(1-
equals(con[i-2],w))*equals(s[i-2],s[i]))/nd[i]

            }

sumdev<-sum(dev[])                                    # residual deviance

d[1]<-0
for (k in 2:NT){d[k] ~ dnorm(0,.0001) }                  #  vague priors for basic parameters

d1[1]<-0
for (k in 2:NT){d1[k] ~ dnorm(0,.0001) }


sd~dunif(0,2)                              #  vague prior for random effects standard deviation
```

```
tau<-1/pow(sd,2)

#this calculates the direct and indirect estimates for the same comparison

P[1]<-d[32]-d[5]
P[2]<-d1[32]-d1[5]
for(m in 1:2){
rk[m]<-rank(P[],m)
best[m]<-equals(rk[m],1)

      }


      }

list(N=95, NT=36)

diff[] var[]      t[]       b[]       nd[]      s[]       con[]

0.107898285       0.023267397   26        4         1         1         20
-0.149161681      0.034305608   11        2         1         2         20
-0.114197207      0.071545009   31        3         1         3         20
-0.242640196      0.108982733   32        5         1         4         4
0.061092357       0.100046653   17        32        1         5         20
.................
```

10.6 Data Extraction Tables

Full data extraction tables are available at http://www.york.ac.uk/inst/crd/Documents/AllInterventionsDataExtraction.pdf

10.7 PRISMA Checklist

Section/topic	Checklist item		Page number
Title	1	Identify the report as a systematic review, meta-analysis, or both.	i
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1-6
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	7-8
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	9
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	10
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	11
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	12-13
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	147-149
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	13
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	13
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	14
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	14
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	14-15

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	14-16
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	13-14
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	17-20
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	91
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	22-90 and 178
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).	22-90
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.	95-131
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	22-90 and 95-131
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	22-90
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression) (see Item 16).	95-131
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).	134-135
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).	135-137
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	138
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Iii

10.8 List of excluded studies

Abasolo L, Carmona L, Hernandez-Garcia C, Lajas C, Loza E, Blanco M, et al. Musculoskeletal work disability for clinicians: time course and effectiveness of a specialized intervention program by diagnosis. *Arthritis Rheum* 2007;57:335-42.

Adedoyin RA, Olaogun MOB, Oyeyemi AL. Transcutaneous electrical nerve stimulation and interferential current combined with exercise for the treatment of knee osteoarthritis: a randomised controlled trial. *Hong Kong Physiotherapy Journal* 2005;23:13-19.

Ahsin S. Electro-acupuncture shows promise for knee arthritis, but we are not yet convinced. *Focus on Alternative and Complementary Therapies* 2010;15:32-33.

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;147:60-6.

Alcock K, Barrett A, Rothwell G, Ainsworth G. Study to evaluate the effects of an exercise and education based osteoarthritis knee programme on patient's pain, function and ability to self manage one year post discharge. In: *The Chartered Society of Physiotherapy Annual Congress and Exhibition: Defining Practice*; 2003 17-19 October; Birmingham, England. 2003.

Allegrante JP, Kovar PA, MacKenzie CR, Peterson MG, Gutin B. USA impact of a supervised walking program and education on functional status: results from a controlled trial in patients with osteoarthritis of the knee. *Patient Educ Couns* 1991;18:283-4.

Ammer K, Petschnig R. [Comparative study between acupuncture and physical therapy in with gonarthrosis]. *Wien Med Wochenschr* 1988;138:566-9.

Aoki O, Tsumura N, Kimura A, Okuyama S, Takikawa S, Hirata S. Home stretching exercise is effective for improving knee range of motion and gait in patients with knee osteoarthritis. *Journal of Physical Therapy Science* 2009;21:113-9.

Ay S, Evcik D. The effects of pulsed electromagnetic fields in the treatment of knee osteoarthritis: a randomized, placebo-controlled trial. *Rheumatol Int* 2009;29:663-6.

Bagheri J, Esfahanian F, Fooladgar E, Oraizi E. Effects of laser acupuncture on osteoarthritis of knee joint. *Osteoporos Int* 2003;14:S19-S20.

Band PA. Comparing different therapeutic classes for the treatment of osteoarthritis of the knee: data from the OMERACT-OARSI responder criteria analysis. *Osteoarthritis Cartilage* 2005;13:449.

Bansil CK, Joshi JB. Effectiveness of shortwave diathermy and ultrasound in the treatment of osteo-arthritis of the knee joint. *Med J Zambia* 1975;9:138-39.

Barton GR, Sach TH, Jenkinson C, Doherty M, Avery AJ, Muir KR. Lifestyle interventions for knee pain in overweight and obese adults aged > or = 45: economic evaluation of randomised controlled trial. *BMJ* 2009;339:b2273.

Battisti E, Piazza E, Rigato M, Nuti R, Bianciardi L, Scribano A, et al. Efficacy and safety of a musically modulated electromagnetic field (TAMMEF) in patients affected by knee osteoarthritis. *Clin Exp Rheumatol* 2004;22:568-72.

Bautch JC, Malone DG, Vailas AC. Effects of exercise on knee joints with osteoarthritis: a pilot study of biologic markers. *Arthritis Care Res* 1997;10:48-55.

Bennell KL, Hinman RS, Metcalf BR, Buchbinder R, McConnell J, McColl G, et al. Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. *Ann Rheum Dis* 2005;64:906-12.

Berman BM, Lao L, Greene M, Anderson RW, Wong RH, Langenberg P, et al. Efficacy of traditional Chinese acupuncture in the treatment of symptomatic knee osteoarthritis: a pilot study. *Osteoarthritis Cartilage* 1995;3:139-42.

Birmingham TB, Kramer JF, Kirkley A, Inglis JT, Spaulding SJ, Vandervoort AA. Knee bracing for medial compartment osteoarthritis: effects on proprioception and postural control. *Rheumatology* 2001;40:285-89.

Bliddal H, Astrup A, Christensen R. Sustained weight loss reduces pain in obese patients with osteoarthritis in the knee: a randomized trial. In: *Proceedings of the 11th World Congress on Pain*; 2005; Sydney, Australia. 2005. p. 363.

Bruyere O. Both weight-bearing and non-weight-bearing exercise improved function in patients with knee osteoarthritis. *Evid Based Med* 2009;14:178.

Budiman-Mak E, Collins E, Langbein EW, Foley S, Oconnell S, Bammert C, et al. Home-based exercise and weight control in overweight elderly individuals with knee osteoarthritis. *Arthritis Rheum* 2005;52:S504.

Bulthuis Y, Drossaers-Bakker KW, Taal E, Rasker J, Oostveen J, van't Pad Bosch P, et al. Arthritis

patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. *Rheumatology* 2007;46:1712-7.
Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A, et al. Self management of arthritis in primary care: randomised controlled trial. *BMJ* 2006;333:879-83.
Campbell R, Evans M, Tucker M, Quilty B, Dieppe P, Donovan JL. Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patients with osteoarthritis of the knee. *J Epidemiol Community Health* 2001;55:132-8.
Cao L-H, Wang S-C, Zhang Q-J, Li Z-D, Ma Y-H, Su J-C, et al. [Evaluation of the clinical effect of knee osteoarthritis treated with moxibustion according to WHOQOL-BREF]. *Zhongguo Gu Shang* 2009;22:813-5.
Cetin N, Aytar A, Atalay A, Akman MN. Comparing hot pack, short-wave diathermy, ultrasound, and TENS on isokinetic strength, pain, and functional status of women with osteoarthritic knees: a single-blind, randomized, controlled trial. *Am J Phys Med Rehabil* 2008;87:443-51.
Chahade W, Samara AM, Silva NA, Seda H, Radominski SC, Keiserman M, et al. Eficácia sintomática dos insaponificáveis de abacate e soja (IAS)* no tratamento de osteoartrose (OA) de quadril e joelho. *Rev Bras Med* 2004;61:711-18.
Chaipinyo K, Karoonsupcharoen O. No difference between home-based strength training and home-based balance training on pain in patients with knee osteoarthritis: a randomised trial. *Aust J Physiother* 2009;55:25-30.
Chamberlain MA, Care G, Harfield B. Physiotherapy in osteoarthrosis of the knees. A controlled trial of hospital versus home exercises. *Int Rehabil Med* 1982;4:101-6.
Cheing GLY, Hui-Chan CWY. Would the addition of TENS to exercise training produce better physical performance outcomes in people with knee osteoarthritis than either intervention alone? *Clin Rehabil* 2004;18:487-97.
Chen KW, Perlman A, Liao JG, Lam A, Staller J, Sigal LH. Effects of external qigong therapy on osteoarthritis of the knee. A randomized controlled trial. *Clin Rheumatol* 2008;27:1497-505.
Chen Y. Treatment of genual osteoarthritis by massotherapy. *J Tradit Chin Med* 2000;20:191-4.
Chiosoig FN, Hendriks O, Malone J. A pilot study of the therapeutic effects of bipolar and quadripolar interferential therapy, using bilateral osteoarthritis as a model. *Physiotherapy Ireland* 1994;15:3-7.
Christensen BV, Iuhl IU, Vilbek H, Bulow HH, Dreijer NC, Rasmussen HF. Acupuncture treatment of severe knee osteoarthrosis: a long-term study. *Acta Anaesthesiol Scand* 1992;36:519-25.
Christensen BV, Juhl IU, Wilbek H, Bulow HH, Dreijer NC, Rasmussen HF. [Acupuncture treatment of knee arthrosis. A long-term study]. *Ugeskr Laeger* 1993;155:4007-11.
Christie A, Moe RH. Aerobic walking and strengthening exercises have similar effectiveness for knee osteoarthritis. *Aust J Physiother* 2005;51:193.
Chua SD, Jr., Messier SP, Legault C, Lenz ME, Thonar EJMA, Loeser RF. Effect of an exercise and dietary intervention on serum biomarkers in overweight and obese adults with osteoarthritis of the knee. *Osteoarthritis Cartilage* 2008;16:1047-53.
Cochrane T, Davey RC, Matthes Edwards SM. Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis. *Health Technol Assess* 2005;9:1-114.
Coleman S, Briffa N, Burton H, Carroll G, Inderjeeth C, Cook N, et al. Effects of self-management, education, and specific exercises, delivered by health professionals, using behaviour modification in patients with osteoarthritis of the knee. *Ann Rheum Dis* 2005;64:67-68.
Coleman S, Briffa NK, Carroll G, Inderjeeth C, Cook N, McQuade J. Effects of self-management, education and specific exercises, delivered by health professionals, in patients with osteoarthritis of the knee. *BMC Musculoskelet Disord* 2008;9:133.
Coupe VMH, Veenhof C, van Tulder MW, Dekker J, Bijlsma JWJ, Van den Ende CHM. The cost effectiveness of behavioural graded activity in patients with osteoarthritis of hip and/or knee. *Ann Rheum Dis* 2007;66:215-21.
Cummings M. Research reviews Acupuncture has a role in treating OA knee (n=393). *Acupunct Med* 2002;20:41-2.
Cusack T, McAteer MF, Daly LE, McCarthy CJ. Knee osteoarthritis: a randomized controlled trial comparing hydrotherapy and continuous short-wave diathermy. *Arthritis Rheum* 2005;52:S506.
Dennis DA, Komistek RD, Nadaud MC, Mahfouz M. Evaluation of off-loading braces for treatment of unicompartmental knee arthrosis. *J Arthroplasty* 2006;21(Suppl 1):2-8.
Deyle GD, Allison SC, Matekel RL, Ryder MG, Stang JM, Gohdes DD, et al. Physical therapy treatment effectiveness for osteoarthritis of the knee: a randomized comparison of supervised clinical exercise and manual therapy procedures versus a home exercise program. *Phys Ther*

2005;85:1301-17.
Deyle GD, Henderson NE, Matekel RL, Helewa A. Manual physical therapy and exercise improved function in osteoarthritis of the knee. *Evid Based Med* 2000;5:145.
Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee: a randomized, controlled trial. *Ann Intern Med* 2000;132:173-81.
Dincer U, Cakar E, Ozdemir B, Kiralp MZ, Dursun H. [Comparison of effects of combined physical therapy program and exercise on corrupted balance functions in patient with knee bilateral osteoarthritis]. *Romatizma* 2008;23:9-13.
Ding M-H, Zhang H, Li Y. [A randomized controlled study on warming needle moxibustion for treatment of knee osteoarthritis]. *Zhongguo Zhen Jiu* 2009;29:603-7.
Diracoglu D, Aydin R, Baskent A, Celik A. Effects of kinesthesia and balance exercises in knee osteoarthritis. *J Clin Rheumatol* 2005;11:303-10.
Doi T, Akai M, Fujino K, Iwaya T, Kurosawa H, Hayashi K, et al. Effect of home exercise of quadriceps on knee osteoarthritis compared with nonsteroidal antiinflammatory drugs: a randomized controlled trial. *Am J Phys Med Rehabil* 2008;87:258-69.
Domaille MJ, Mascarenhas R, Dayal N, Kirwan JR. Evaluation of an evidence-based physiotherapy programme for the management of patients with osteoarthritis of the knee. *Rheumatology* 2005;44(Suppl 1):i147.
Domínguez Carrillo LG. [Comparación de la electroestimulación transcutánea vs láser de bajo nivel en la analgesia de pacientes con gonartrosis]. *Cirugia y cirujanos* 2001;69:13-17.
Dominick KL, Golightly YM, Bosworth HB. Pilot study of home-based strength training for knee osteoarthritis. *Arthritis Rheum* 2004;50:S469-S70.
Endres HG, Victor N, Haake M, Witte S, Streitberger K, Zenz M. [Acupuncture for the treatment of chronic knee and back pain]. *Dtsch Arztebl* 2007;104:A123-30.
Endres HG, Zenz M, Schaub C, Molsberger A, Haake M, Streitberger K, et al. [German Acupuncture Trials (gerac) address problems of methodology associated with acupuncture studies] [erratum appears in Schmerz. 2005 Aug;19(4):333]. *Schmerz* 2005;19:201-13.
Erbahceci F, Dursun E. The effects of exercises in women with knee osteoarthritis. *Ann Rheum Dis* 2007;66:639-40.
Erhart JC, Mundermann A, Elspas B, Giori NJ, Andriacchi TP. A variable-stiffness shoe lowers the knee adduction moment in subjects with symptoms of medial compartment knee osteoarthritis. *J Biomech* 2008;41:2720-5.
Ettinger W, Afable R. Physical disability from knee osteoarthritis: the role of exercise as an intervention. *Med Sci Sports Exerc* 1994;26:1435-40.
Eungpinichpong W. The efficacy of physical exercise programmes for patients with osteoarthritis of the knee as determined by clinical and gait parameters. *New Zealand Journal Physiotherapy* 1998;26:5.
Evcik D, Kavuncu V, Yeter A, Yigit I. The efficacy of balneotherapy and mud-pack therapy in patients with knee osteoarthritis. *Joint Bone Spine* 2007;74:60-5.
Evcik D, Sonel B. Effectiveness of a home-based exercise therapy and walking program on osteoarthritis of the knee. *Rheumatol Int* 2002;22:103-6.
Eyigor S. A comparison of muscle training methods in patients with knee osteoarthritis. *Clin Rheumatol* 2004;23:109-15.
Falconer J, Hayes KW, Chang RW. Effect of ultrasound on mobility in osteoarthritis of the knee. A randomized clinical trial. *Arthritis Care Res* 1992;5:29-35.
Fargas-Babjak A, Rooney P, Gerecz E. Randomized trial of Codetron for pain control in osteoarthritis of the hip/knee. *Clin J Pain* 1989;5:137-41.
Farina G, Arrigo C, Bocchino L, Morgante L, Morgante S, Miceli G, et al. Efficacy of thermal mud bath in patients with knee osteoarthrosis, with evaluation of concentration of blood amino-acids. *Ann Rheum Dis* 2003;62:275-75.
Farr JN, Going SB, McKnight PE, Kasle S, Cussler EC, Cornett M. Progressive resistance training improves overall physical activity levels in patients with early osteoarthritis of the knee: a randomized controlled trial. *Phys Ther* 2010;90:356-66.
Ferrell BA, Josephson KR, Pollan AM, Loy S, Ferrell BR. A randomized trial of walking versus physical methods for chronic pain management. *Aging Clin Exp Res* 1997;9:99-105.
Fioravanti A, Bisogno S, Nerucci F, Cicero MR, Locunsolo S, Marcolongo R. [Assessment of the effectiveness and tolerability of mud-packs therapy with fluorurate radioactivity water on gonarthrosis. A comparative study versus short wave therapy]. *Minerva Med* 2000;91:291-8.
Fisher N, Pendergast D, Gresham G, Calkins E. Muscle rehabilitation: its effect on muscular and

functional performance of patients with knee osteoarthritis. *Arch Phys Med Rehabil* 1991;72:367-74.
Fisher NM. Effects of three exercise rehabilitation programs on the relationship between muscular strength and efficiency in knee osteoarthritis. *Arthritis Rheum* 2005;52:1130.
Fisher NM, Gresham G, Pendergast DR. Effects of a quantitative progressive rehabilitation program applied unilaterally to the osteoarthritic knee. *Arch Phys Med Rehabil* 1993;74:1319-26.
Fisher NM, Gresham GE, Abrams M, Hicks J, Horrigan D, Pendergast DR. Quantitative effects of physical therapy on muscular and functional performance in subjects with osteoarthritis of the knees. *Arch Phys Med Rehabil* 1993;74:840-7.
Fisher NM, Gresham GE, Pendergast DR. Quantitative progressive exercise rehabilitation for osteoarthritis of the knee. *Phys Med Rehabil Clin N Am* 1994;5:785-802.
Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis: a randomised controlled trial comparing a gym based and a hydrotherapy based strengthening programme. *Ann Rheum Dis* 2003;62:1162-7.
Forogh B, Karami M, Emadifar R, Saeedi H. Comparing the efficacy of a lateral-wedge insole with subtalar supporting and an in-shoe lateral-wedge in patients with varus deformity osteoarthritis of the knee. In: *4th World Congress of the International Society of Physical and Rehabilitation Medicine*; 2007; Seoul, South Korea. International Society of Physical and Rehabilitation Medicine; 2007.
Foster NE, Deyle GD, Allison SC. Supervised clinic-based treatment enhances the effects of home-exercise programmes for knee osteoarthritis, but only in the short term. *Focus on Alternative and Complementary Therapies* 2006;11:133-5.
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;335:436.
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. *Rheumatology* 2008;47(Suppl 2):ii2-ii2.
Foster NE, Thomas E, Hill JC, Hay EM. The relationship between patient and practitioner expectations and preferences and clinical outcomes in a trial of exercise and acupuncture for knee osteoarthritis. *Eur J Pain* 2010;14:402-09.
Foy CG, Penninx B, Shumaker S, Pahor M. Long-term exercise therapy resolves the influence of ethnicity upon baseline health status among older adults with knee osteoarthritis. *J Am Geriatr Soc* 2004;52:P433.
Fraenkel L, Fried T. Patients with knee osteoarthritis (OA) prefer exercise over NSAIDs. *Arthritis Rheum* 2006;54:S667-S68.
Fransen M. Home exercise produces small reductions in pain and functional limitation in patients with osteoarthritic knees. *Aust J Physiother* 1999;45:322.
Fransen M, Crosbie J, Edmonds J. Effectiveness of physiotherapy for patients with OA knee. *Aust N Z J Med* 2000;30:529.
Fransen M, Crosbie J, Edmonds J. Physical therapy is effective for patients with osteoarthritis of the knee: a randomized controlled clinical trial. *J Rheumatol* 2001;28:156-64.
Fransen M, Margiotta E, Heussier J, Edmonds J. Group exercise for subjects with osteoarthritis of the knee. *Aust J Physiother* 1995;41:255-60.
Fransen M, Nairn L, Winstanley J, Lam P, Edmonds J. Physical activity for osteoarthritis management: a randomized controlled clinical trial evaluating hydrotherapy or Tai Chi classes. *Arthritis Rheum* 2007;57:407-14.
Furmanov S, Nelson D, Li Y, Shakoor N. The influence of exercise oil pain, muscle strength and proprioception in knee osteoarthritis(OA): does disease severity influence outcome? *Arthritis Rheum* 2004;50:4089-89.
Gaines JM. *The effects of neuromuscular electrical stimulation on chronic knee pain and functional performance in older adults with osteoarthritis of the knee* [PhD]. Baltimore, Maryland: John Hopkins University; 2002.
Gaines JM, Metter EJ, Talbot LA. The effect of neuromuscular electrical stimulation on arthritis knee pain in older adults with osteoarthritis of the knee. *Appl Nurs Res* 2004;17:201-6.
Garche U, Kruger H. Einsatz isokinetischer Systeme bei Gonarthrosen wahrend stationarer Rehabilitation: Teil II - Isokinetischer Training. *Pravention Rehabilitation* 1997;9:161-5.
Gasparyan LV, Solovieva TIe. Low level laser therapy of myofascial pain syndromes of the patients with osteoarthritis of knee and hip joints. *SPIE proceedings series* 2001;4422:12-17.
Ghroubi S, Elleuch H, Kaffel N, Echikh T, Abid M, Elleuch MH. [Contribution of exercise and diet in the

management of knee osteoarthritis in the obese]. *Ann Readapt Med Phys* 2008;51:663-70.
Gill SD, McBurney H, Schulz DL. Land-based versus pool-based exercise for people awaiting joint replacement surgery of the hip or knee: results of a randomized controlled trial. *Arch Phys Med Rehabil* 2009;90:388-94.
Gilsenan CO, Keogan F, Blake C, O'Connell P. Efficacy of open and closed kinetic chain strengthening exercises in subjects with osteoarthritis of the knee: a prospective, randomised trial. *Phys Ther Rev* 2006;11:221-2.
Graham J, Fisher N. Effects of resistance training on muscle function in individuals with knee osteoarthritis. *Arthritis Rheum* 2003;48(Suppl):S445.
Gremion G, Gaillard D, Leyvraz P-F, Jolles BM. Effect of biomagnetic therapy versus physiotherapy for treatment of knee osteoarthritis: a randomized controlled trial. *J Rehabil Med* 2009;41:1090-5.
Guillemin F, Virion JM, Escudier P, De Talance N, Weryha G. Effect on osteoarthritis of spa therapy at Bourbonne-les-Bains. *Joint Bone Spine* 2001;68:499-503.
Gur A, Cevik R, Cosut A, Nas K, Sarac AJ. Comparison of effects of low power laser and physiotherapy in the management of painful knee osteoarthritis. *Osteoporos Int* 2003;14:S53-S53.
Haaz S, Wrobleski P, Ruffing C, Bathon JM, Fontaine KR, Bartlett SJ. Improving HRQOL through diet and exercise in obese persons with knee OA. *Arthritis Rheum* 2004;50:S452-S53.
Halbert J, Crotty M, Weller D, Ahern M, Silagy C. Primary care-based physical activity programs: effectiveness in sedentary older patients with osteoarthritis symptoms. *Arthritis Rheum* 2001;45:228-34.
Harlow T, Greaves C, White A, Brown L, Hart A, Ernst E. Randomised controlled trial of magnetic bracelets for relieving pain in osteoarthritis of the hip and knee. *BMJ* 2004;329:1450-4.
Hart LE. Combination of manual physical therapy and exercises for osteoarthritis of the knee. *Clin J Sport Med* 2000;10:305.
Hart LE. Home exercise for knee pain and knee osteoarthritis. *Clin J Sport Med* 2003;13:388-9.
Hartman CA, Manos TM, Winter C, Hartman DM, Li B, Smith JC. Effects of Tai Chi training on function and quality of life indicators in older adults with osteoarthritis. *J Am Geriatr Soc* 2000;48:1553-59.
Hederstrom M, Olsson E, Granath G, Moller E, Oscarsson M, Weijdenhielm L. Physiotherapy in medial gonarthrosis. *Acta Orthop Scand* 1989;60(Suppl 231):33.
Hegedus B, Viharos L, Gervain M, Galfi M. The effect of low-level laser in knee osteoarthritis: a double-blind, randomized, placebo-controlled trial. *Photomed Laser Surg* 2009;27:577-84.
Heuts PHTG, de Bie R, Drietelaar M, Aretz K, Hopman-Rock M, Bastiaenen CHG, et al. Self-management in osteoarthritis of hip or knee: a randomized clinical trial in a primary healthcare setting. *J Rheumatol* 2005;32:543-9.
Hicks-Little C, Hubbard T, Cordova M. The role of exercise in the treatment of knee osteoarthritis. *Athl Ther Today* 2008;13:7-10.
Hinman RS, Bowles KA, Bennell KL. Laterally wedged insoles in knee osteoarthritis: do biomechanical effects decline after one month of wear? *BMC Musculoskelet Disord* 2009;10:146.
Hinman RS, Bowles KA, Payne C, Bennell KL. Effect of length on laterally-wedged insoles in knee osteoarthritis. *Arthritis Rheum* 2008;59:144-7.
Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. *Phys Ther* 2007;87:32-43.
Hinman RS, Metcalf BR, Payne C, Bennell KL, Wrigley TW. Immediate biomechanical effects of laterally-wedged shoe insoles in medial knee osteoarthritis. *Arthritis Rheum* 2005;52:1970.
Hopman-Rock M, Westhoff MH. The effects of a health educational and exercise program for older adults with osteoarthritis for the hip or knee. *J Rheumatol* 2000;27:1947-54.
Horlick SG, Loomer RL. Valgus knee bracing for medial gonarthrosis. *Clin J Sport Med* 1993;3:251-55.
Horstmann T, Mayer F, Heitkamp HC, Merk J, Axmann D, Bork H, et al. [Isokinetic strength-training in patients with osteoarthritis of the knee]. *Z Rheumatol* 2000;59:93-100.
Huang MH, Chen CH, Chen TW, Weng MC, Wang WT, Wang YL. The effects of weight reduction on the rehabilitation of patients with knee osteoarthritis and obesity. *Arthritis Care Res* 2000;13:398-405.
Huang M-H, Lin Y-S, Yang R-C, Lee C-L. A comparison of various therapeutic exercises on the functional status of patients with knee osteoarthritis. *Semin Arthritis Rheum* 2003;32:398-406.
Huang M-H, Yang R-C, Lee C-L, Chen T-W, Wang M-C. Preliminary results of integrated therapy for

patients with knee osteoarthritis. *Arthritis Rheum* 2005;53:812-20.
Hughes SL, Seymour RB, Campbell RT, Huber G, Pollak N, Sharma L, et al. Long-term impact of fit and strong! on older adults with osteoarthritis. *Gerontologist* 2006;46:801-14.
Hurley MV, Scott DL. Improvements in quadriceps sensorimotor function and disability of patients with knee osteoarthritis following a clinically practicable exercise regime. *Br J Rheumatol* 1998;37:1181-87.
Irie H. [Effects of whole-body vibration exercise on osteoarthritis of the knee]. *Journal of the Japanese Orthopaedic Association* 2009;83:S32.
Jan MH, Lai JS. The effects of physiotherapy on osteoarthritic knees of females. *J Formos Med Assoc* 1991;90:1008-13.
Jan M-H, Chai H-M, Wang C-L, Lin Y-F, Tsai L-Y. Effects of repetitive shortwave diathermy for reducing synovitis in patients with knee osteoarthritis: an ultrasonographic study. *Phys Ther* 2006;86:236-44.
Jan M-H, Lin C-H, Lin Y-F, Lin J-J, Lin D-H. Effects of weight-bearing versus nonweight-bearing exercise on function, walking speed, and position sense in participants with knee osteoarthritis: a randomized controlled trial. *Arch Phys Med Rehabil* 2009;90:897-904.
Jan M-H, Tang P-F, Lin J-J, Tseng S-C, Lin Y-F, Lin D-H. Efficacy of a target-matching foot-stepping exercise on proprioception and function in patients with knee osteoarthritis. *J Orthop Sports Phys Ther* 2008;38:19-25.
Jawad AS, Goodwill CJ. TVS brace in patients with rheumatoid arthritis or osteoarthritis of the knee. *Br J Rheumatol* 1986;25:416-7.
Jensen H, Zesler R, Christensen T. Transcutaneous electrical nerve stimulation (TNS) for painful osteoarthrosis of the knee. *Int J Rehabil Res* 1991;14:356-8.
Jessep SA, Walsh NE, Ratcliffe J, Hurley MV. Long-term clinical benefits and costs of an integrated rehabilitation programme compared with outpatient physiotherapy for chronic knee pain. *Physiotherapy* 2009;95:94-102.
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;9:18-19.
Jiang A, Zhang L, Zhao C, Yang F. Clinical effect of acupuncture treatment in 109 cases of knee osteoarthritis. *J Tradit Chin Med* 2001;21:282-5.
Jokic A, Hansjurgens K, Sremcevic N, Hansjurgens A, Pekmezovic T, Vasic S, et al. Gonarthrosis: treatment by horizontal therapy. A comparison study. In: *4th World Congress of the International Society of Physical and Rehabilitation Medicine*; 2007; Seoul, South Korea. International Society of Physical and Rehabilitation Medicine; 2007.
Kachelman J, Bibeau W, Topp R, Swank NRA, Quesada P, Nyland J, et al. Effects of prehabilitation on strength and range of motion among end-stage knee osteoarthritis patients. *Med Sci Sports Exerc* 2007;39:S251.
Kao KC, Hong S, Taylor CE, Nouvong A, Masih S, Perell KL, et al. Laterally wedged shoe inserts as an intervention for medial knee osteoarthritis. *J Investig Med* 2006;54:302.
Kitay GS, Koren MJ, Helfet DL, Markenson JA. Efficacy of local mechanical vibrations, continuous passive motion and thermotherapy in the management of osteoarthritis of the knee: a prospective, randomized, controlled, double cross-over study. *Ann Rheum Dis* 2007;66:500-01.
Kitay GS, Koren MJ, Helfet DL, Parides MK, Markenson JA. Efficacy of combined local mechanical vibrations, continuous passive motion and thermotherapy in the management of osteoarthritis of the knee. *Osteoarthritis Cartilage* 2009;17:1269-74.
Klaber MJA, Frost H, Osborn A, Richardson PH. Controlled study to evaluate the effectiveness of pulsed short wave for pain relief in osteoarthritic hips & knees. In: *12th World Confederation of Physical Therapy Congress*; 1995 25-30 June; Washington DC. 1995. p. 1084.
Klaber Moffett JA, Richardson PH, Frost H, Osborn A. A placebo controlled double blind trial to evaluate the effectiveness of pulsed short wave therapy for osteoarthritic hip and knee pain. *Pain* 1996;67:121-7.
Kocaman O, Koyuncu H, Dinc A, Toros H, Karamehmetoglu SS. The comparison of the effects of electrical stimulation and exercise in the treatment of knee osteoarthritis. *Turk Fiz Tip Rehab Derg* 2008;54:54-58.
Korzh IV, Fedotova IF. The response of knee osteoarthritis patients to treatment with rofecoxib and physiotherapy. *Ann Rheum Dis* 2004;63:357-57.
Korzh IV, Fedotova IF. Comparison of effects of celecoxib and physiotherapy in patients with knee osteoarthritis. *Ann Rheum Dis* 2005;64:479-79.
Kostopoulos D. *Comparative effects of aquatic recreational and aquatic exercise programs on*

mobility, pain perception, and treatment satisfaction among elderly persons with osteoarthritis of the knee. New York: New York University; 2000.
Kozanoglu E, Basaran S, Guzel R, Guler-Uysal F. Short term efficacy of ibuprofen phonophoresis versus continuous ultrasound therapy in knee osteoarthritis. *Swiss Med Wkly* 2003;133:333-8.
Kreindler H, Lewis C, Rush S, Schaefer K. Effects of three exercise protocols on strength of persons with osteoarthritis of the knee. *Top Geriatr Rehabil* 1989;4:32-9.
Kruger H, Garche U. [The use of isokinetic systems at osteoarthritis during in-patient rehabilitation, part II - Isokinetic training]. *Pravention und Rehabilitation* 1997;9:161-65.
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. *Am J Chin Med* 2001;29:187-99.
Laufer Y, Zilberman R, Porat R, Nahir AM. Effect of pulsed short-wave diathermy on pain and function of subjects with osteoarthritis of the knee: a placebo-controlled double-blind clinical trial. *Clin Rehabil* 2005;19:255-63.
Laurie S. Acupuncture for the treatment of bilateral osteoarthritis of the knee. *J Acupunct Assoc Chartered Physiotherapists* 2005;2005:56-60.
Law PPW, Cheing GLY, Tsui AYY. Does transcutaneous electrical nerve stimulation improve the physical performance of people with knee osteoarthritis? *J Clin Rheumatol* 2004;10:295-99.
Lee H-Y. [Comparison of effects among Tai-Chi exercise, aquatic exercise, and a self-help program for patients with knee osteoarthritis]. *Daehan Ganho Haghoeji* 2006;36:571-80.
Lee H-Y, Lee KJ. [Effects of Tai Chi exercise in elderly with knee osteoarthritis]. *Daehan Ganho Haghoeji* 2008;38:11-8.
Li C-D, Huang X-Y, Yang X-G, Wang Q-F, Huang S-Q. [Observation on therapeutic effect of warming needle moxibustion on knee osteoarthritis of deficiency-cold type]. *Zhongguo Zhen Jiu* 2006;26:189-91.
Li J-W, Xiang S-Y, Ma Z-Y, Feng Y-B, Tong H-Y, Geng H-P, et al. [Clinical observation on cake-separated mild-warm moxibustion for treatment of knee osteoarthritis]. *Zhongguo Zhen Jiu* 2008;28:17-9.
Li Q, Zhu J. [Observation on therapeutic effect of moxibustion at Shenque (CV 8) combined with electroacupuncture on gonarthritis of type of accumulation of pathogenic cold due to yang deficiency]. *Zhongguo Zhen Jiu* 2008;28:565-8.
Liddle S. Integrated exercise and self-management programmes in osteoarthritis of the hip and knee. *Phys Ther Rev* 2007;12:285-6.
Lievens P, Voorde Jvd. The influence of cycloidal vibrations on the knee joint mobility of osteoarthritic patients. *Physiotherapy (London)* 1984;70:241-42.
Lim BW. A comparative study of open and closed kinetic chain exercise regimes in patients with knee osteoarthritis. *Physiotherapy Singapore* 2002;5:34-40.
Lin D-H, Lin Y-F, Chai H-M, Han Y-C, Jan M-H. Comparison of proprioceptive functions between computerized proprioception facilitation exercise and closed kinetic chain exercise in patients with knee osteoarthritis. *Clin Rheumatol* 2007;26:520-8.
Lin L-F, Liang Y-P. [Clinical observation on electric heat needle therapy for treatment of gonarthritis]. *Zhongguo Zhen Jiu* 2005;25:689-90.
Lin SYC, Davey RC, Cochrane T. Community rehabilitation for older adults with osteoarthritis of the lower limb: a controlled clinical trial. *Clin Rehabil* 2004;18:92-101.
Lin YC, Davey RC. Community-based water therapy for elderly with osteoarthritis (OA) of the knee or hip. In: Martos E, editor. *XXVII FIMS: World Congress of Sports Medicine*. Bologna: Medimond Inc.; 2002. p. 203-06.
Linde K, Streng A, Hoppe A, Jurgens S, Weidenhammer W, Melchart D. The programme for the evaluation of patient care with acupuncture (PEP-Ac) - A project sponsored by ten German social health insurance funds. *Acupunct Med* 2006;24:S25-S32.
Ling SM. Does acupuncture improve symptoms in patients with osteoarthritis who are awaiting knee replacement surgery? *Nat Clin Pract Rheumatol* 2008;4:286-7.
Lloyd CH, Royer TD, Barrios J, Davis IM. Medially wedged insoles reduce knee kain during functional activities in subjects with lateral knee osteoarthritis. *Med Sci Sports Exerc* 2007;39:S390.
Lone AR, Wafai ZA, Buth BA, Wani TA, Koul PA, Khan SH. Analgesic efficacy of transcutaneous electrical nerve stimulation compared with diclofenac sodium in osteo-arthritis of the knee. *Physiotherapy* 2003;89:478-85.
Loomer R, Horlick S. Valgus knee bracing for medial gonarthrosis. In: *Proceedings of 7th World Congress of Interantional Society for Prosthetics and Orthotics*; 1992; Chicago, Illinois. 1992. p. 218.

Lund H, Henriksen M, Bartels EM, Danneskiold-Samsoe B, Bliddal H. Can stimulating massage improve joint repositioning error in patients with knee osteoarthritis? *J Geriatr Phys Ther* 2009;32:111-6.
Maiko OI, Gogoleva EF. [Outpatient bioresonance treatment of gonarthrosis]. *Ter Arkh* 2000;72:50-3.
Marks R. The effect of isometric quadriceps strength training in mid-range for osteoarthritis of the knee. *Arthritis Care Res* 1993;6:52-6.
Marks R. The effect of isometric quadriceps strength training in mid-range for osteo-arthritis of the knee. *New Zealand Journal of Physiotherapy* 1993;21:16-8.
McAuley R, Ysla R. Soft laser: A treatment for osteoarthritis of the knee? *Arch Phys Med Rehabil* 1985;66:553-54.
McDonald HL. Patients who wore standard magnetic bracelets reported reduced pain from osteoarthritis of the hip or knee compared with patients wearing placebo bracelets. *Evid Based Nurs* 2005;8:89.
McKnight PE, Kasle S, Going S, Villanueva I, Cornett M, Farr J, et al. A comparison of strength training, self-management, and the combination for early osteoarthritis of the knee. *Arthritis Care Res* 2010;62:45-53.
Meng C-R, Fan L, Fu W-B, Li Y. Clinical research on abdominal acupuncture plus conventional acupuncture for knee osteoarthritis. *J Tradit Chin Med* 2009;29:249-52.
Messier SP. Knee strengthening exercises at home, but not dietary intervention, reduce knee pain and improve knee function in overweight and obese people over 2 years. *Evid Based Med* 2010;15:21-2.
Messier SP, Loeser RF, Mitchell MN, Valle G, Morgan TP, Rejeski WJ, et al. Exercise and weight loss in obese older adults with knee osteoarthritis: a preliminary study. *J Am Geriatr Soc* 2000;48:1062-72.
Messier SP, Royer TD, Craven TE, O'Toole ML, Burns R, Ettinger WH, Jr. Long-term exercise and its effect on balance in older, osteoarthritic adults: results from the Fitness, Arthritis, and Seniors Trial (FAST). *J Am Geriatr Soc* 2000;48:131-8.
Metin I, Hakguder HA. [Comparison of clinical efficacy of intermittent and continuous cold treatment in acute knee osteoarthritis]. *Turk Fiz Tip Rehab Derg* 2007;53:65-68.
Michailov St Dinitrov I, Papasova V, Kalaidshiev G, Iliev G. [Die Behandlung der Gonarthrose bei älteren Menschen mit dem neuen bulgarischen Apparat "Arfai"]. *Z Physiother Jg* 1981;33:21-27.
Mihailov M, Popa D. Efficacy of physical therapy upon pain and disability in knee osteoarthritis. *Osteoporos Int* 2007;18(Suppl 1):S172-S73.
Mihailov M, Popa D, Pascalau N. Is exercise effective treatment for osteoarthritis of the knee? *Osteoporos Int* 2009;20(Suppl 1):117.
Mikesky AE, Mazzuca SA, Brandt KD, Perkins SM, Lane KA. Effects of strength training on incidence and progression of knee osteoarthritis (OA). *Arthritis Rheum* 2005;52:1331.
Miller GD, Nicklas BJ, Loeser RF. Inflammatory biomarkers and physical function in older, obese adults with knee pain and self-reported osteoarthritis after intensive weight-loss therapy. *J Am Geriatr Soc* 2008;56:644-51.
Minor M. Exercise in the management of osteoarthritis of the knee and hip. *Arthritis Care Res* 1994;7:198-204.
Mohomed NN. Manual physical therapy and exercise improved function in osteoarthritis of the knee. *J Bone Joint Surg Am* 2000;82:1324.
Murphy SL, Strasburg DM, Lyden AK, Koliba JF, Smith DM, Dadabhoy DP, et al. The effect of activity strategy training on pain and physical activity in older adults with knee or hip osteoarthritis. *Arthritis Rheum* 2008;58:S244.
Nadaud MC, Komistek RD, Mahfouz MR, Dennis DA, Anderle MR. In vivo three-dimensional determination of the effectiveness of the osteoarthritic knee brace: a multiple brace analysis. *J Bone Joint Surg Am* 2005;87(Suppl 2):114-9.
Nejrup K, Olivarius NdF, Jacobsen JL, Siersma V. Randomised controlled trial of extraarticular gold bead implantation for treatment of knee osteoarthritis: a pilot study. *Clin Rheumatol* 2008;27:1363-9.
Nelson E, Cornett N, Ryan PC, Villanueva I, Going S, Flint-Wagner H, et al. Development of a physical training program for early knee osteoarthritis patients. *Arthritis Rheum* 2004;50:S451.
Nicklas BJ, Ambrosius W, Messier SP, Miller GD, Penninx BWJH, Loeser RF, et al. Diet-induced weight loss, exercise, and chronic inflammation in older, obese adults: a randomized controlled clinical trial. *Am J Clin Nutr* 2004;79:544-51.
Ochi M. The effects of foot orthoses and valgus bracing on the osteoarthritis knees. *Journal of the*

Japanese Orthopaedic Association 2009;83:S27.
Ohta H. Effect of muscle and stretching exercise in osteoarthritis of the knee. *Journal of the Japanese Orthopaedic Association* 2000;74:S216.
Oida Y, Morozumi K, Nakamura N, Kitabatake Y, Shiozawa S, Sato S, et al. [Effectiveness of a community health service program using exercise intervention for elderly people with osteoarthritis of the knees: a randomized controlled trial]. *Nippon Koshu Eisei Zasshi* 2008;55:228-37.
Oldham J, Howe T. The effectiveness of placebo muscle stimulation in quadriceps muscle rehabilitation: a preliminary evaluation. *Clin Eff Nurs* 1997;1:25-30.
Oldham JA, Howe TE, Patterson T, Smith GP, Tallis RC. Electrotherapeutic rehabilitation of the quadriceps in elderly osteoarthritic patients: a double blind assessment of patterned neuromuscular stimulation. *Clin Rehabil* 1995;9:10-20.
Ones K, Tetik S, Tetik C, Ones N. The effects of heat on osteoarthritis of the knee. *Pain Clinic* 2006;18:67-75.
O'Reilly SC, Muir KR, Doherty M. Effectiveness of home exercise on pain and disability from osteoarthritis of the knee: a randomised controlled trial. *Ann Rheum Dis* 1999;58:15-19.
Ozdincler AR, Yeldan I, Kinali P. The effects of closed kinetic chain exercise on pain and functional performance of patients with knee osteoarthritis. *Pain Clinic* 2005;17:107-15.
Park JS, Kwon SJ, Kwon YS. [The effects of Dogbi(ST35) & Sulan moxibustion on knee joint pain, range of motion and discomfort during ADL in the aged]. *Daehan Ganho Haghoeji* 2006;36:189-96.
Patel A, Buszewicz M, Beecham J, Griffin M, Rait G, Nazareth I, et al. Economic evaluation of arthritis self management in primary care. *BMJ* 2009;339:794.
Penninx BW, Messier SP, Rejeski WJ, Williamson JD, DiBari M, Cavazzini C, et al. Physical exercise and the prevention of disability in activities of daily living in older persons with osteoarthritis. *Arch Intern Med* 2001;161:2309-16.
Petrella RJ, Bartha C. Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial. *J Rheumatol* 2000;27:2215-21.
Petronijevic V, Lazovic M, Kocic M, Djokic B. Pulsed electromagnetic field and low-energy laser in the treatment of acute knee osteoarthritis. In: Peek WJ, Lankhorst GJ, editors. *1st World Congress of the International Society of Physical and Rehabilitation Medicine*; 2001 7-13 July; Amsterdam, The Netherlands. Medimond S R L; 2001. p. 287-91.
Piechocki M. Physiotherapy equal to placebo for treating knee OA. *Orthopedics Today* 2005;25:22.
Pisters MF, Veenhof C, De Bakker DH, Schellevis FG, Dekker J. Physical activity and exercise adherence in physical therapy exercise treatment in patients with osteoarthritis of hip or knee. *Arthritis Rheum* 2008;58:S607-S08.
Quilty B, Tucker M, Campbell R, Dieppe P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. *J Rheumatol* 2003;30:1311-7.
Quirk AS, Newman RJ, Newman KJ. An evaluation of interferential therapy, shortwave diathermy and exercise in the treatment of osteoarthrosis of the knee. *Physiotherapy* 1985;71:55-7.
Ravaud P, Flipo RM, Boutron I, Roy C, Mahmoudi A, Giraudeau B, et al. ARTIST (osteoarthritis intervention standardized) study of standardised consultation versus usual care for patients with osteoarthritis of the knee in primary care in France: pragmatic randomised controlled trial. *BMJ* 2009;338:b421.
Ravaud P, Giraudeau B, Logeart I, Larguier JS, Rolland D, Treves R, et al. Management of osteoarthritis (OA) with an unsupervised home based exercise programme and/or patient administered assessment tools: a cluster randomised controlled trial with a 2x2 factorial design. *Ann Rheum Dis* 2004;63:703-8.
Rawiworrakul T, Sirapo-Ngam Y, Davis AHT, Malathum P, Kulthanan T, Vorapongsathorn T. A community-based exercise program promotes self-efficacy for exercise among Thai women with osteoarthritis of the knee. *Thai Journal of Nursing Research* 2007;11:132-50.
Reid D, McNair PJ. The effects of a 6 week stretching intervention on range of motion, torque and stiffness in people with osteoarthritis of the knee. *New Zealand Journal of Physiotherapy* 2008;36:91.
Reinhold T, Witt CM, Jena S, Brinkhaus B, Willich SN. Quality of life and cost-effectiveness of acupuncture treatment in patients with osteoarthritis pain. *Eur J Health Econ* 2008;9:209-19.
Rioja-Toro J, Estevez P, Prada-Espinal D, Gonzaley-Rebollo Y. Los amplitude and extremely low frequency magnetic fields for the treatment of chronic knee pain. *Rehabilitacion* 2008;42:127-36.

Rohde J, Jaschke B. [The "joint school". Part 2: Investigations into the effect of bicycle exercise on joint pain in gonarthrosis]. *Manuelle Medizin* 2004;42:279-86.
Rudisile-Smith C, Loosli M, Brumitt J. Clinically applicable papers: land-based versus pool-based exercise for people awaiting joing replacement surgery of the hip or knee: results of a randomized controlled trial. *New Zealand Journal of Physiotherapy* 2009;37:160.
Safaie MM, Azami S, Mokmeli S, Grigorian G. Comparing the efficacy of low level laser therapy (LLLT) with physiotherapy in treatment of knee osteoarthritis (OA). *Photomed Laser Surg* 2006;25:332.
Salim M. Transcutaneous electrical nerve stimulation (TENS) in chronic pain. *Alternative Therapies in Clinical Practice* 1996;3:33-35.
Sasaki T, Yasuda K. Clinical evaluation of the treatment of osteoarthritic knees using a newly designed wedged insole. *Clin Orthop Relat Res* 1987:181-7.
Scharf H-P, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C, et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. *Ann Intern Med* 2006;145:12-20.
Schmidtke-Schrezenmeier G, Reck R, Gerster G. [Behandlung der nichtaktivierten Gonarthrose. Besserung durch ein Phytotherapeutikum]. *Therapiewoche* 1992;42:1322-25.
Seay J, Messier SP. Effects of exercise and diet on knee strength in older overweight and obese adults with knee osteoarthritis. *Med Sci Sports Exerc* 2004;36:S94.
Segal NA, Foster NA, Dhamani S, Ohashi K, Yack HJ. Effects of concurrent use of an ankle support with a laterally wedged insole for medial knee osteoarthritis. *PM R* 2009;1:214-22.
Seto H, Ikeda H, Hisaoka H, Kurosawa H. Effect of heat- and steam-generating sheet on daily activities of living in patients with osteoarthritis of the knee: randomized prospective study (erratum: Journal of Orthopaedic Science (2008) 13, 3, (187-191)). *J Orthop Sci* 2008;13:578.
Seto H, Ikeda H, Hisaoka H, Kurosawa H. Effect of heat- and steam-generating sheet on daily activities of living in patients with osteoarthritis of the knee: randomized prospective study. *J Orthop Sci* 2008;13:187-91.
Shakoor N, Nelson D, Li Y, Block J. The influence of exercise on pain, muscle strength, and proprioception in knee osteoarthritis (OA). *Osteoarthritis Cartilage* 2004;12:S136-S37.
Sherman AM, Shumaker SA, Rejeski W, Morgan T, Applegate WB, Ettinger W. Social support, social integration, and health-related quality of life over time: results from the Fitness and Arthritis in Seniors Trial (FAST). *Psychol Health* 2006;21:463-80.
Silva ALP, Imoto DM, Croci AT. Comparison of cryotherapy, exercise and short waves in knee ostioarthritis treatment. *Acta Ortopédica Brasileira* 2007;15:204-09.
Silva LE, Pessanha AC, Oliveira LM, Myamoto S, Valim V, Jones A, et al. Efficacy of water exercise in the treatment of patients with knee osteoarthritis: a randomized, single-blind, controlled clinical trial. *Ann Rheum Dis* 2005;64:556-56.
Silva LE, Valim V, Pessanha APC, Oliveira LM, Myamoto S, Jones A, et al. Hydrotherapy versus conventional land-based exercise for the management of patients with osteoarthritis of the knee: a randomized clinical trial. *Phys Ther* 2008;88:12-21.
Slivar SR, Zilic I, Subert N, Mrkonja L, Marosevic D. The influence of exercise on quadriceps strength among different decade in the elderly with knee osteoarthritis. In: *Proceedings of the 7th Mediterranean Congress of Physical and Rehabilitation Medicine*: Mediterranean Forum of Physical and Rehabilitation Medicine 2008. p. 163. Available from: http://www.mfprm.org/PRM_proceedings_7_444.pdf
Soeken K. Placebo effects may have confounded electro-acupuncture study results. *Focus on Alternative and Complementary Therapies* 2003;8:53-4.
Song R, Lee E-O, Lam P, Bae S-C. Effects of tai chi exercise on pain, balance, muscle strength, and perceived difficulties in physical functioning in older women with osteoarthritis: a randomized clinical trial. *J Rheumatol* 2003;30:2039-44.
Song R, Roberts BL, Lee E-O, Lam P, Bae S-C. A randomized study of the effects of t'ai chi on muscle strength, bone mineral density, and fear of falling in women with osteoarthritis. *J Altern Complement Med* 2010;16:227-33.
Stigsgaard L, Christensen R, Astrup A, Bliddal H. Application of an intensive dietary regime in rehabilitation of overweight patients with knee osteoarthritis: a randomised controlled weight loss/self-management trial. *Ann Rheum Dis* 2006;65(Suppl 2):655.
Su J-C, Cao L-H, Li Z-D, Wang S-C, Zhang Q-J, Ma Y-H, et al. [Controlled clinical trials of initial observation on therapeutic effects of moxibustion for osteoarthritis of the knee: multi-center clinical effect]. *Zhongguo Gu Shang* 2009;22:914-6.
Suraj, Kumar S. Effect of knee complex mobilization on pain and active range of motion arc in osteoarthritis knee joint. *Phys Ther Sport* 2006;7:176.

Svarcova J, Trnavsky K, Zvarova J. The influence of ultrasound, galvanic currents and shortwave diathermy on pain intensity in patients with osteoarthritis. *Scand J Rheumatol Suppl* 1987;67:83-5.
Svarcova J, Zvarova J, Kouba A, Trnavsky K. [Does physiotherapy affect the pain in activated arthrosis?]. *Z Physiother* 1988;40:333-36.
Szucs L, Ratko I, Lesko T, Szoor I, Genti G, Balint G. Double-blind trial on the effectiveness of the Puspokladany thermal water on arthrosis of the knee-joints. *J R Soc Health* 1989;109:7-9.
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. *Acupunct Med* 2009;27:3-8.
Takazawa Y. [A randomized trial comparing three different modes of exercises on osteoarthritis of the knee]. *Journal of the Japanese Orthopaedic Association* 2008;82:S18.
Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. *J Am Geriatr Soc* 2003;51:387-92.
Terek M, Radunovic G, Marcetic D, Gavrilov N. Long lasting improvement of functional status of patients with knee osteoarthritis after physical therapy at 6 months follow-up. *J Rehabil Med* 2008;40(Suppl 47):225.
The Norwegian Knowledge Centre for the Health Services. Effectiveness of physical therapy, restricted to electrotherapy and exercise, for osteoarthritis of the knee (Structured abstract). *[serial online]* 2004:104. [cited Available from: http://www.mrw.interscience.wiley.com/cochrane/clhta/articles/HTA-32005000204/frame.html
Thomas KS, Miller P, Doherty M, Muir KR, Jones AC, O'Reilly SC. Cost effectiveness of a two-year home exercise program for the treatment of knee pain. *Arthritis Rheum* 2005;53:388-94.
Thomas KS, Muir KR, Doherty M, Jones AC, O'Reilly SC, Bassey EJ. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. *BMJ* 2002;325:752.
Thym DB, Olmedija dAR, Dall'Aqua JM, Baldan C. Evaluation of pain relief using a combination of TENS and cryotherapy. *Krankengymnastik* 2005;57:894-901.
Tillu A, Roberts C, Tillu S. Unilateral versus bilateral acupuncture on knee function in advanced osteoarthritis of the knee - a prospective randomised trial. *Acupunct Med* 2001;19:15-8.
Tillu A, Tillu S, Vowler S. Effect of acupuncture on knee function in advanced osteoarthritis of the knee: a prospective, non-randomised controlled study. *Acupunct Med* 2002;20:19-21.
Toda Y. [The effect of weight loss therapy on the obese patients with osteoarthritis of the knee]. *Journal of the Japanese Orthopaedic Association* 1998;72:S502.
Toda Y. The effect of energy restriction, walking, and exercise on lower extremity lean body mass in obese women with osteoarthritis of the knee. *J Orthop Sci* 2001;6:148-54.
Toda Y. [The effect of insole with figure 8 elastic ankle fixation band for patients with osteoarthritis of the knee]. *Journal of the Japanese Orthopaedic Association* 2001;75:S350.
Toda Y. [A comparison of the efficacy of conservative therapies for obese patients with osteoarthritis of the knee]. *Ryumachi* 2002;42:795-800.
Toda Y. [Usefulness of an insole with subtalar strapping for analgesia in patients with osteoarthritis of the knee]. *Journal of the Japanese Orthopaedic Association* 2004;78:S343.
Toda Y. An optimal duration of insole with subtalar strapping wear each day for patients with varus deformity osteoarthritis of the knee. *Osteoarthritis Cartilage* 2004;12:S25.
Toda Y, Segal N, Kato A, Yamamoto S, Irie M. Effect of a novel insole on the subtalar joint of patients with medial compartment osteoarthritis of the knee. *J Rheumatol* 2001;28:2705-10.
Toda Y, Tsukimura N. A comparative study on the effect of the insole materials with subtalar strapping in patients with medial compartment osteoarthritis of the knee. *Mod Rheumatol* 2004;14:459-65.
Toda Y, Tsukimura N. A six month follow-up of a randomized trial to compare the efficacy of a lateral wedged insole with subtalar strapping and an in-shoe lateral wedged insole in patients with varus deformity osteoarthritis of the knee. *Arthritis Rheum* 2004;50:S343.
Toda Y, Tsukimura N. Randomised trial to compare the clinical effects of an insole with subtalar strapping and knee support with hinged struts for patients with varus deformity osteoarthritis of the knee. *Arthritis Rheum* 2004;50:S342-S43.
Toda Y, Tsukimura N. A six-month followup of a randomized trial comparing the efficacy of a lateral-wedge insole with subtalar strapping and an in-shoe lateral-wedge insole in patients with varus deformity osteoarthritis of the knee. *Arthritis Rheum* 2004;50:3129-36.
Toda Y, Tsukimura N. A 2-year follow-up of a study to compare the efficacy of lateral wedged insoles

with subtalar strapping and in-shoe lateral wedged insoles in patients with varus deformity osteoarthritis of the knee. *Osteoarthritis Cartilage* 2006;14:231-7.
Toda Y, Tsukimura N, Kato A. The effects of different elevations of laterally wedged insoles with subtalar strapping on medial compartment osteoarthritis of the knee. *Arch Phys Med Rehabil* 2004;85:673-7.
Toda Y, Tsukimura N, Segal N. An optimal duration of daily wear for an insole with subtalar strapping in patients with varus deformity osteoarthritis of the knee. *Osteoarthritis Cartilage* 2005;13:353-60.
Trock DH, Bollet AJ, Dyer RH, Jr., Fielding LP, Miner WK, Markoll R. A double-blind trial of the clinical effects of pulsed electromagnetic fields in osteoarthritis. *J Rheumatol* 1993;20:456-60.
Tsauo J-Y, Cheng P-F, Yang R-S. The effects of sensorimotor training on knee proprioception and function for patients with knee osteoarthritis: a preliminary report. *Clin Rehabil* 2008;22:448-57.
Tunay VB, Baltaci G. Optimal stimulation and treatment duration of neuromuscular electrical stimulation in the management of osteoarthritic knee pain. *Ann Rheum Dis* 2006;65:671-72.
Tunay VB, Baltaci G. Home-based exercise programme and clinical-based functional exercise programme for patients with knee osteoarhritis: a randomised controlled trial. *Ann Rheum Dis* 2007;66:638-39.
Tunay VB, Baltaci G, Yakut E, Vardar N. Time-benefit in the physiotherapy management of knee osteoarthritis: Clinical or group-based physiotherapy. *Ann Rheum Dis* 2006;65:672-72.
Tunay VB, Ozer D, Duzgun I, Baltaci G, Ergun N. Does neuromuscular electrical stimulation improve the performance-based physical functioning of people with knee osteoarthritis? *Ann Rheum Dis* 2007;66:515-15.
Tuzun EH, Otman S, Kirdi N. Comparison of different methods of pulsed shortwave diathermy in knee osteoarthritis. *Pain Clinic* 2003;15:421-27.
van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Bijlsma JW. Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up. *Ann Rheum Dis* 2001;60:1123-30.
van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Lemmens JA, et al. The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized clinical trial. *J Rheumatol* 1998;25:2432-9.
Veenhof C, Koke AJA, Dekker J, Oostendorp RA, Bijlsma JWJ, van Tulder MW, et al. Effectiveness of behavioral graded activity in patients with osteoarthritis of the hip and/or knee: a randomized clinical trial. *Arthritis Rheum* 2006;55:925-34.
Veenhof C, Van den Ende CHM, Dekker J, Kiike AJA, Oostendorp RA, Bijlsma JWJ. Which patients with osteoarthritis of hip and/or knee benefit most from behavioral graded activity? *Int J Behav Med* 2007;14:86-91.
Volklein R, Callies R. [Changes in pain by different types of diadynamic current in gonarthrosis and lumbar syndrome]. *Z Physiother* 1990;42:113-18.
Von Svarcova J, Zvarova J, Pichova A, Kouba A, Simacek K, Uhlemann C, et al. [Comparison of the analgesic effects of electroacupuncture and of galvanic current in patients with activated ostoarthrosis (a controlled clinical study)]. *Z Physiother* 1990;42:375-78.
Wang C, Schmid CH, Hibberd P, Kalish R, Roubenoff R, Rones R, et al. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. *Arthritis Rheum* 2008;58:S401.
Wang C, Schmid CH, Hibberd PL, Kalish R, Roubenoff R, Rones R, et al. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. *Arthritis Rheum* 2009;61:1545-53.
Wang C, Schmid CH, Hibberd PL, Kalish R, Roubenoff R, Rones R, et al. Tai Chi for treating knee osteoarthritis: designing a long-term follow up randomized controlled trial. *BMC Musculoskelet Disord* 2008;9:108.
Wang J-G, He L-J. [Observation on the therapeutic effect of warming needle moxibustion on knee osteoarthritis]. *Zhongguo Zhen Jiu* 2007;27:191-2.
Wang J-L, Chai C-H, Xu Y-M. [Clinical observations on the effect of Tuigua manipulation combined with quadriceps exercise for the treatment of degenerative gonarthritis]. *Zhongguo Gu Shang* 2008;21:887-9.
Wang T-J, Belza B, Elaine Thompson F, Whitney JD, Bennett K. Effects of aquatic exercise on flexibility, strength and aerobic fitness in adults with osteoarthritis of the hip or knee. *J Adv Nurs* 2007;57:141-52.
Wang TT. *Aquatic exercise improves flexibility, strength, and walk time in osteoarthritis*. University of Washington; 2004.
Weng M-C, Lee C-L, Chen C-H, Hsu J-J, Lee W-D, Huang M-H, et al. Effects of different stretching

techniques on the outcomes of isokinetic exercise in patients with knee osteoarthritis. *Gaoxiong Yi Xue Ke Xue Za Zhi* 2009;25:306-15.
Wetzels R, van Weel C, Grol R, Wensing M. Family practice nurses supporting self-management in older patients with mild osteoarthritis: a randomized trial. *BMC Fam Pract* 2008;9:7.
White A. Acupuncture is a very promising treatment for osteoarthritis of the knee. *Focus on Alternative and Complementary Therapies* 2001;6:251-2.
Whitehurst DG, Bryan S, Thomas E, Hay EM, Young J, Foster NE. The cost-utility of acupuncture as an adjunct to exercise-based physiotherapy for osteoarthritis of the knee. *Rheumatology* 2009;48:I106-I06.
Wolf T, Fussganger-May T, Weber M. A study to test the efficacy of laser-shower-treatment. *Schmerz Akupunkt* 2006;32:159-62.
Wu Y-y, Liao J-p. [Treatment of knee joint pain with superficial needling]. *Zhongguo Zhen Jiu* 2005;25:261-2.
Wyatt FB, Milam S, Manske RC, Deere R. The effects of aquatic and traditional exercise programs on persons with knee osteoarthritis. *J Strength Cond Res* 2001;15:337-40.
Wyatt MR, Yein K, Melton JTK, Williamson L. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise), and standard management for patients awaiting knee replacement. *Rheumatology* 2007;46:I15-I15.
Xu H, Wu H-G. [Clinical observation of electroacupuncture combined with low-dose diclofenac in treating osteoarthritis of the knee]. *Zhong Xi Yi Jie He Xue Bao* 2007;5:457-9.
Yagli NV, Yakut E, Akdogan A, Yakut Y. [The effect of quadriceps muscle strength on walking time and functional capacity of patients with osteoarthritis of the knee]. *Fizyoterapi Rehabilitasyon* 2007;18:153.
Yakut E, Yagli N, Akdogan A, Kiraz S. [The role of Pilates exercises in patients with osteoarthritis of the knee: a pilot study]. *Fizyoterapi Rehabilitasyon* 2006;17:51-60.
Yurtkuran M, Alp A, Konur S, Ozcakir S, Bingol U. Laser acupuncture in knee osteoarthritis: a double-blind, randomized controlled study. *Photomed Laser Surg* 2007;25:14-20.
Yuruk ZOB, Gunel MK. [Effectiveness of classic physiotherapy program on functioning, quality of life and depression in women with knee osteoarthritis] *Fizyoterapi Rehabilitasyon* 2008;19:164.
Zeng H-W, Nie B, Shi L-l. [Observation on therapeutic effect of blood-letting puncture combined with red-hot needle therapy on knee osteoarthritis]. *Zhongguo Zhen Jiu* 2008;28:493-5.
Zhou F. Warming acupuncture in treating arthralgia of knee. *International Journal of Clinical Acupuncture* 1997;8:425-6.
Zivkovic VD, Lazovic M, Kocic M, Kojovic Z. Different physiotherapy programs for patients with knee osteoarthritis. *Ann Rheum Dis* 2005;64:557-58.
Zwoelfer W, Grubhofer G, Cartellieri M, Spacek A. Acupuncture in gonarthrotic pain Bachmann's knee program. *Am J Chin Med* 1992;20:325-29.