

Osteoarthritis and Cartilage



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



M.S. Corbett †*, S.J.C. Rice †, V. Madurasinghe †^a, R. Slack †, D.A. Fayter †, M. Harden †, A.J. Sutton †, H. MacPherson †, N.F. Woolacott †

† University of York, UK

‡ University of Leicester, UK

ARTICLE INFO

Article history:

Received 12 November 2012

Accepted 13 May 2013

Keywords:

Osteoarthritis

Knee

Pain

Physical treatments

Network meta-analysis

SUMMARY

Objective: To compare the effectiveness of acupuncture with other relevant physical treatments for alleviating pain due to knee osteoarthritis.

Design: Systematic review with network meta-analysis, to allow comparison of treatments within a coherent framework. Comprehensive searches were undertaken up to January 2013 to identify randomised controlled trials in patients with osteoarthritis of the knee, which reported pain.

Results: Of 156 eligible studies, 114 trials (covering 22 treatments and 9,709 patients) provided data suitable for analysis. Most trials studied short-term effects and many were classed as being of poor quality with high risk of bias, commonly associated with lack of blinding (which was sometimes impossible to achieve). End of treatment results showed that eight interventions: interferential therapy, acupuncture, TENS, pulsed electrical stimulation, balneotherapy, aerobic exercise, sham acupuncture, and muscle-strengthening exercise produced a statistically significant reduction in pain when compared with standard care. In a sensitivity analysis of satisfactory and good quality studies, most studies were of acupuncture (11 trials) or muscle-strengthening exercise (9 trials); both interventions were statistically significantly better than standard care, with acupuncture being statistically significantly better than muscle-strengthening exercise (standardised mean difference: 0.49, 95% credible interval 0.00–0.98).

Conclusions: As a summary of the current available research, the network meta-analysis results indicate that acupuncture can be considered as one of the more effective physical treatments for alleviating osteoarthritis knee pain in the short-term. However, much of the evidence in this area of research is of poor quality, meaning there is uncertainty about the efficacy of many physical treatments.

© 2013 Osteoarthritis Research Society International. Published by Elsevier Ltd.

Open access under [CC BY-NC-SA license](#).

Introduction

The objective of treating osteoarthritis of the knee is usually the alleviation of pain and improving quality of life. Failure to control pain may result in reduced mobility and reduced participation in daily activities, which may further exacerbate symptoms. The regular use of pharmacological agents for pain may be associated with significant side effects (such as gastrointestinal bleeding)¹, and

many patients want non-pharmacological treatments for pain relief^{2,3}. Effective alternatives to pharmacological pain relief are therefore desirable.

Five guidelines (ACR⁴, AAOS⁵, OARSI⁶, EULAR⁷, and NICE⁸) have evaluated treatment effects on key outcomes of knee osteoarthritis (including pain, function, and disability). All recommend treatment with muscle-strengthening and aerobic exercise, education, weight loss (if required), and, where necessary, paracetamol and/or topical NSAIDs; when these are ineffective, a choice of one or more options from a range of pharmacological and non-pharmacological treatments is sometimes recommended, including transcutaneous electrical nerve stimulation (TENS), thermal (heat/cooling) treatments, insoles, and braces. The OARSI guideline recommended using acupuncture, AAOS found the acupuncture evidence to be inconclusive, and the ACR conditionally recommended acupuncture only for patients with moderate-to-severe pain who are unable or unwilling to undergo total knee arthroplasty. EULAR and NICE did

* Address correspondence and reprint requests to: M.S. Corbett, Centre for Reviews and Dissemination, University of York, Heslington, York, YO10 5DD, UK. Tel: 44-01904-321072, fax: 44-01904-32104.

E-mail address: mark.corbett@york.ac.uk (M.S. Corbett).

^a Now at: Perinatal Institute, Birmingham, UK.

not recommend use of acupuncture; one of the reasons for the commissioning of this review – as part of a programme of projects on acupuncture and chronic pain, funded by the National Institute for Health Research (NIHR) under its Programme Grant for Applied Research Programme – 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. The rationale for this systematic review was to compare acupuncture with available alternative physical treatments that might be prescribed by a GP, or used by a physiotherapist, since uncertainty exists regarding which treatments are best.

Although numerous reviews have evaluated individual types of physical treatment, few randomised trials have directly compared these treatments. One way to overcome this limitation is to use network meta-analysis, which allows assessment of relative efficacy when direct treatment comparisons are insufficient or unavailable. In the context of the present review it should enable all relevant physical treatments to be compared with each other. The purpose of this systematic review, therefore, was to conduct a comprehensive synthesis using network meta-analysis methods in order to compare the effectiveness of acupuncture with other relevant physical treatments for alleviating pain due to osteoarthritis of the knee.

Methods

A systematic review was conducted following the general principles outlined in the Centre for Reviews and Dissemination (CRD) Guidance⁹ and the PRISMA statement¹⁰. This paper reports an update of a systematic review and network meta-analysis conducted in 2011, which is available on the CRD website¹¹.

Literature search

A range of resources was searched for published and unpublished studies, grey literature, and on-going research (see eMethods 1). We searched 17 electronic databases from inception to January 2013, without language restrictions. A combination of relevant free text terms, synonyms and subject headings relating to osteoarthritis of the knee and named physical treatments were included in the strategy. Bibliographies of relevant reviews and guidelines were also checked, and Internet searches were made of websites relating to osteoarthritis.

Study selection and intervention definitions

Two reviewers independently screened all abstracts and full papers, with disagreements resolved by discussion, or a third reviewer. We included randomised controlled trials (RCTs) assessing pain (as a primary or secondary outcome) in adults with knee osteoarthritis (with a population mean age of ≥ 55 years). Eligible treatments were 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, TENS, and weight loss. The following were excluded: predominantly home-based and unsupervised exercise interventions, surgical interventions, pharmaceutical interventions, interventions which combined two or more physical treatments, and studies comparing only different regimens/durations/modalities of the same intervention. Populations with varus/valgus malalignment were excluded as were studies which did not report data in a format suitable for network meta-analysis (see Outcomes section).

We classified adjunctive components of the experimental interventions into five categories, based on what was reported in the trials: 'treatment as usual', 'treatment as usual' plus specified home exercise or education, 'treatment as usual' plus specified (trial-specific) analgesics, no medication, and no medication plus specified home exercise or education. Eligible comparators included any form of standard/usual care or waiting list control (which could incorporate analgesics, education, and exercise advice) all of which we called 'standard care'. Placebo interventions, no intervention, and sham acupuncture were also eligible. Sham acupuncture was treated as a separate comparator because of evidence suggesting it is more active than an inert 'placebo'^{12,13}. All pain scales were eligible.

Assessment of trial quality and data extraction

Trial quality was assessed using an adaptation of a checklist (14 questions) from a previous review by CRD¹⁴. Using an algorithm, studies were then graded as excellent, good, satisfactory or poor, and also given an assessment based on the Cochrane risk of bias tool¹⁵ [see eTables I(a and b)]. Data extraction and quality assessments were performed by one reviewer and independently checked by a second. Disagreements were resolved by discussion or a third reviewer.

Outcomes and data transformations

WOMAC pain (using a VAS or Likert scale) was the preferred pain measure. When studies did not measure WOMAC pain, another pain scale was included in the analysis with prioritisation of scales made on a clinical, or prevalence, basis (further details in the 2011 report)¹¹. Hedges-g standardised mean differences (SMDs) were calculated for the meta-analyses (studies reporting medians could not be analysed). Results for different doses/regimens of the same type of treatment within a study were pooled. In an initial analysis only final values were used. However, we included more studies by calculating final values for trials reporting change from baseline data, provided trial baseline data together with variance estimates (e.g., standard deviations) were also reported. In order to present more clinically meaningful results, we present both SMDs, and SMDs converted to the WOMAC pain VAS 0-100 scale.

Synthesis

A network meta-analysis draws on both direct evidence (treatments compared in the same trial) and indirect evidence (different treatments studied in separate trials, but compared when they use a common comparator), with the benefit of randomisation in each study retained. For indirect and direct evidence to be consistent, population and intervention characteristics must be similar across comparisons^{16–21}. Inconsistency between direct and indirect evidence was assessed using the node-splitting method^{17,22}. The SMD was assumed to be normally distributed and a random effects network meta-analysis model was selected since clinical and methodological heterogeneity within treatments appeared likely²³. Analyses were conducted using WinBUGS software (version 1.4). Further method detail can be found in eMethods 2.

We conducted analyses with interventions categorised both with, and without, any adjunct treatments. Furthermore, in order to attempt to assess both the immediacy and durability of effects, we planned analyses for three time points: end of treatment (our primary time point) as defined in the studies; 3 months from the start of treatment (the time point closest to 3 months from the start

of treatment, excluding outcomes recorded at less than 4 weeks from the start of treatment); and three months after the end of treatment (the time point closest to 3 months, but between 8 and 16 weeks, from end of treatment). However, due to a lack of medium- and long-term data, we report here results for the end of treatment time point only.

To evaluate the impact of study quality on the results, two sets of analyses were performed: one including all studies regardless of quality ('any-quality'), and a primary sensitivity analysis including studies of satisfactory, or better, quality ('better-quality'). Studies with atypical populations, interventions, or results were excluded in a second sensitivity analysis. When possible, examination of funnel plots was used to assess for publication bias.

Results

156 original trials (of 22 distinct interventions and comparators) met the inclusion criteria. Four of 10 foreign language papers which appeared eligible based on their English abstracts could not be translated, so had to be excluded from our analyses^{24–27}. One

retracted study was removed from all analyses²⁸. Twenty-two new studies were identified from the 2013 update searches. A study selection flow diagram is presented in eFig. 1.

Study characteristics

An overview of all eligible studies – regardless of whether they reported data suitable for network meta-analysis – is presented in Table I. The range of mean treatment durations (and timing of end of treatment assessment) varied widely from just a single session (TENS) to 69.3 weeks (weight loss interventions), although a majority of interventions were administered over a 2–6 week period. Most studies were classified as having recruited a general knee osteoarthritis population, 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.

Around three-quarters of the studies were classed as being of poor quality (110 of 152). The remainder were 'satisfactory' (33 studies) or 'good' (9 studies), together classed as 'better-quality'. In the network meta-analyses only 12 trials were considered to be at

Table I
Summary characteristics of trials included in the systematic review

Intervention	No. of trials eligible for the review (number of patients*)	Type of population recruited (number of studies)	Range of mean ages (years)	Range of % female	Comparators (number of treatment arms†)
Acupuncture	25 (2794)	General (23), both knees affected (1), awaiting surgery (1)	58–85	50–96	Sham acupuncture (15), standard care (13), TENS (3), muscle strengthening exercise (1), ice/cooling (1)
Balneotherapy	14 (1008)	General (12), both knees affected (2)	‡54–70	47–100	Placebo (8), standard care (6), heat treatment (1)
Braces	1 (24)	General (1)	59.5	63	Insoles (1)
Aerobic exercise	13 (1136)	General (9), both knees affected (2), overweight or obese (2)	‡54–75	50–100	Standard care (13), muscle strengthening exercise (2), weight loss (1)
Muscle-strengthening exercise	34 (3013)	General (26), both knees affected (5), awaiting surgery (2)	‡53–77	31–100	Standard care (22), placebo (4), no treatment (2), aerobic exercise (2), heat treatment (1), TENS (1), acupuncture (1), PES (1), manual therapy (1), NMES (2)
Heat treatment	7 (412)	General (7)	61–74	63–100	Placebo (4), standard care (1), TENS (1), muscle strengthening exercise (1), balneotherapy (1), ice/cooling (1)
Ice/cooling treatment	4 (211)	General (4)	56–61	48–91	TENS (2), acupuncture (1), standard care (1), heat treatment (1), placebo (1), no treatment (1)
Insoles	6 (893)	General (6)	58–68	54–100	Placebo (5), braces (1)
Interferential therapy	5 (240)	General (5)	59–67	67–80	Placebo (3), TENS (1), no treatment (1)
Laser/light therapy	9 (379)	General (6), both knees affected (3)	58–74	68–90	Placebo (8), standard care (1)
Manual therapy	6 (486)	General (6)	56–68	63–78	Standard care (4), placebo (2), muscle strengthening exercise (1)
NMES	3 (78)	General (3)	60–71	42–79	Standard care (2), muscle strengthening exercise (2)
PES	8 (392)	General (8)	55–70	46–100	Placebo (7), standard care (1), muscle strengthening exercise (1), no treatment (1)
PEMF	6 (521)	General (6)	60–69	28–80	Placebo (6)
Static magnets	3 (131)	General (3)	63–65	60–79	Placebo (3)
Tai Chi	4 (307)	General (4)	65–70	75–93	Standard care (4)
TENS	18 (805)	General (17), awaiting surgery (1)	56–85	48–97	Placebo (12), standard care (3), acupuncture (3), ice/cooling (2), heat treatment (1), interferential (1), no treatment, muscle strengthening exercise (1)
Weight loss (dieting)	5 (870)	Overweight or obese (5)	61–70	26–89	Standard care (5), 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 interventions. 'General' is used as a broad term for studies which did not recruit very specific population types.

Table IIa
Results of network meta-analyses for comparisons with standard care: trials of any-quality

Intervention (abbreviations in brackets relate to Fig. 1a–d)	No. of trials (No. of patients)	SMD (95% Cr I)	Difference expressed on a WOMAC VAS 0–100 pain scale (95% Cr I)*
Standard Care (comparator)		–	–
Interferential therapy (INT)	3 (98)	–1.63 (–2.39 to –0.87)	–26.90 (–39.39 to –14.40)
Acupuncture (ACU)	24 (1219)	–0.89 (–1.18 to –0.59)	–14.69 (–19.52 to –9.80)
TENS (TENS)	12 (285)	–0.65 (–1.06 to –0.25)	–10.77 (–17.50 to –4.05)
PES (PES)	6 (180)	–0.65 (–1.19 to –0.10)	–10.65 (–19.59 to –1.66)
Balneotherapy (BAL)	9 (275)	–0.60 (–1.04 to –0.15)	–9.87 (–17.15 to –2.48)
Aerobic exercise (AE EX)	11 (428)	–0.55 (–0.89 to –0.21)	–9.02 (–14.68 to –3.51)
Tai Chi (TAI)	4 (159)	–0.51 (–1.03 to 0.01)	–8.39 (–16.98 to 0.13)
Static magnets (MAG)	2 (41)	–0.50 (–1.34 to 0.33)	–8.27 (–22.08 to 5.43)
Sham Acupuncture (SH ACU)	14 (892)	–0.47 (–0.84 to –0.09)	–7.76 (–13.89 to –1.52)
Manual therapy (MAN)	4 (166)	–0.44 (–0.96 to 0.09)	–7.21 (–15.90 to 1.49)
Muscle-strengthening exercise (MU EX)	28 (1254)	–0.40 (–0.61 to –0.19)	–6.54 (–9.99 to –3.11)
Ice/cooling treatment (ICE)	3 (51)	–0.35 (–1.03 to 0.33)	–5.81 (–16.94 to 5.44)
Heat treatment (HEA)	5 (123)	–0.31 (–0.86 to 0.24)	–5.14 (–14.20 to 3.98)
Laser therapy (LAS)	5 (155)	–0.27 (–0.86 to 0.32)	–4.53 (–14.19 to 5.20)
Weight loss (WEI)	5 (436)	–0.26 (–0.67 to 0.15)	–4.25 (–10.97 to 2.43)
PEMF (PEMF)	5 (238)	–0.15 (–0.71 to 0.42)	–2.43 (–11.76 to 6.90)
Placebo (PLA)	42 (1077)	–0.07 (–0.42 to 0.29)	–1.15 (–6.98 to 4.70)
Braces (BRA)	1 (12)	0.00 (–1.39 to 1.39)	0.07 (–22.84 to 22.94)
Insoles (INS)	3 (197)	0.10 (–0.65 to 0.85)	1.64 (–10.71 to 13.97)
NMES (NMES)	2 (28)	0.22 (–0.62 to 1.05)	3.58 (–10.26 to 17.33)
No intervention (NO INT)	5 (87)	0.44 (–0.15 to 1.04)	7.25 (–2.51 to 17.12)
Data points: 131 Residual deviance: 152			
Between-study standard deviation: 0.45 (95% CI: 0.36–0.55)			

* see Table IIb footnote.

low risk of bias. Most trials were hampered by a lack of adequate blinding, and small sample sizes (which limited the effectiveness of randomisation, resulting in baseline imbalances). Full study quality assessment results are presented in eTable I(a and b). Study quality did vary by intervention, making the evidence base more robust in some areas than in others [see Table II(a)]. No evidence was found for publication bias (only assessable for muscle-strengthening exercise). Individual study characteristics and a reference list of all studies included in the systematic review can be found in eTable II.

Network meta-analysis

Overall, 114 trials (9,709 patients) reported data suitable for the end of treatment analyses. In addition to the 22 new studies identified from the update searches, nine studies – excluded from the original review analyses – were included in this updated analysis by calculating final values using change from baseline data. Our original analyses (based on searches up to 2010) provided no indication of a treatment effect difference related to the majority of

Table IIb
Results of network meta-analyses for comparisons with acupuncture: trials of any-quality

Intervention [abbreviations in brackets relate to Fig. 1(a–d)]	No. of trials (No. of patients)	SMD (95% Cr I)	Difference expressed on a WOMAC VAS 0–100 pain scale (95% Cr I)*
Acupuncture (comparator)		–	–
Interferential therapy (INT)	3 (98)	–0.74 (–1.54 to 0.05)	–12.21 (–25.33 to 0.84)
TENS (TENS)	12 (285)	0.24 (–0.22 to 0.70)	3.92 (–3.70 to 11.50)
PES (PES)	6 (180)	0.25 (–0.35 to 0.84)	4.04 (–5.78 to 13.87)
Balneotherapy (BAL)	9 (275)	0.29 (–0.22 to 0.81)	4.82 (–3.60 to 13.28)
Aerobic exercise (AE EX)	11 (428)	0.34 (–0.11 to 0.79)	5.67 (–1.84 to 13.00)
Tai Chi (TAI)	4 (159)	0.38 (–0.22 to 0.98)	6.30 (–3.58 to 16.12)
Static magnets (MAG)	2 (41)	0.39 (–0.48 to 1.25)	6.41 (–7.86 to 20.61)
Sham Acupuncture (SH ACU)	14 (892)	0.42 (0.15 to 0.70)	6.93 (2.50 to 11.46)
Manual therapy (MAN)	4 (166)	0.45 (–0.14 to 1.05)	7.47 (–2.30 to 17.23)
Muscle-strengthening exercise (MU EX)	28 (1254)	0.49 (0.15 to 0.84)	8.14 (2.41 to –13.83)
Ice/cooling treatment (ICE)	3 (51)	0.54 (–0.16 to 1.25)	8.88 (–2.70 to 20.61)
Heat treatment (HEA)	5 (123)	0.58 (–0.02 to 1.18)	9.55 (–0.30 to 19.44)
Laser therapy (LAS)	5 (155)	0.62 (–0.02 to 1.25)	10.16 (–0.28 to 20.61)
Weight loss (WEI)	5 (436)	0.63 (0.13 to 1.14)	10.44 (2.13 to 18.72)
PEMF (PEMF)	5 (238)	0.74 (0.13 to 1.36)	12.26 (2.22 to 22.36)
Placebo (PLA)	42 (1077)	0.82 (0.40 to 1.25)	13.53 (6.58 to 20.53)
Standard Care (SC)	53 (2308)	0.89 (0.59 to 1.18)	14.69 (9.80 to 19.52)
Braces (BRA)	1 (12)	0.89 (–0.51 to 2.31)	14.76 (–8.49 to 38.01)
Insoles (INS)	3 (197)	0.99 (0.21 to 1.78)	16.33 (3.41 to 29.30)
NMES (NMES)	2 (28)	1.11 (0.22 to 1.98)	18.27 (3.57 to 32.72)
No intervention (NO INT)	5 (87)	1.33 (0.69 to 1.97)	21.95 (11.30 to 32.52)
Data points: 131 Residual deviance: 152.			
Between-study standard deviation: 0.45 (95% CI: 0.36–0.55).			

* see Table IIb footnote.

adjunctive components of the experimental interventions (see eFig. 2). The exception was that standard care incorporating active analgesia was more effective than standard care with ‘treatment as usual’ (with or without home exercise/education). However, analgesic adjuncts were used in only eight trials. Furthermore, most studies were classified as using the ‘treatment as usual’ adjunct, where little adjunct detail was defined. We therefore focussed on comparing the interventions categorised without adjuncts.

Tables II(a and b) and Fig. 1(a and b) (caterpillar plots) present the primary results, with interventions ordered by treatment effect. The network is illustrated in eFig. 3. When compared with standard care, eight physical treatments had a mean effect suggesting benefit, namely interferential therapy, acupuncture, TENS, pulsed electrical stimulation, balneotherapy, aerobic exercise, sham acupuncture, and muscle-strengthening exercise [Fig. 1(a), Table II(a)]. When acupuncture (rather than standard care) was the comparator, acupuncture was significantly better at reducing pain than sham acupuncture, muscle-strengthening exercise, weight loss, PEMF, placebo, insoles, NMES, and no intervention [Fig. 1(b), Table II(b)].

Across all comparisons, inconsistency at a *P*-value less than 0.05 was only identified for the two comparisons involving PES.

The primary sensitivity analysis of only better-quality studies involved 35 trials, nine types of intervention and 3,499 patients. A small study of muscle-strengthening exercise vs PES was excluded as it was identified as causing inconsistency in the main analysis. The network is illustrated in Fig. 2. The reduction in the number of studies per comparison, as well as loops in the network, increased uncertainty around the true between-study variance. Some interventions were represented by few studies, although there were 11 acupuncture studies and nine muscle-strengthening exercise studies. There was a statistically significant reduction in pain compared with standard care for acupuncture, balneotherapy, sham acupuncture, and muscle-strengthening exercise [Fig. 1(c), Table II(c)]. Acupuncture was statistically significantly better at a 95% level of credibility than sham acupuncture, muscle-strengthening exercise, weight loss, aerobic exercise, and no intervention when the analysis of better-quality studies was presented as a comparison with acupuncture [Fig. 1(d), Table II(d)]. We found that acupuncture

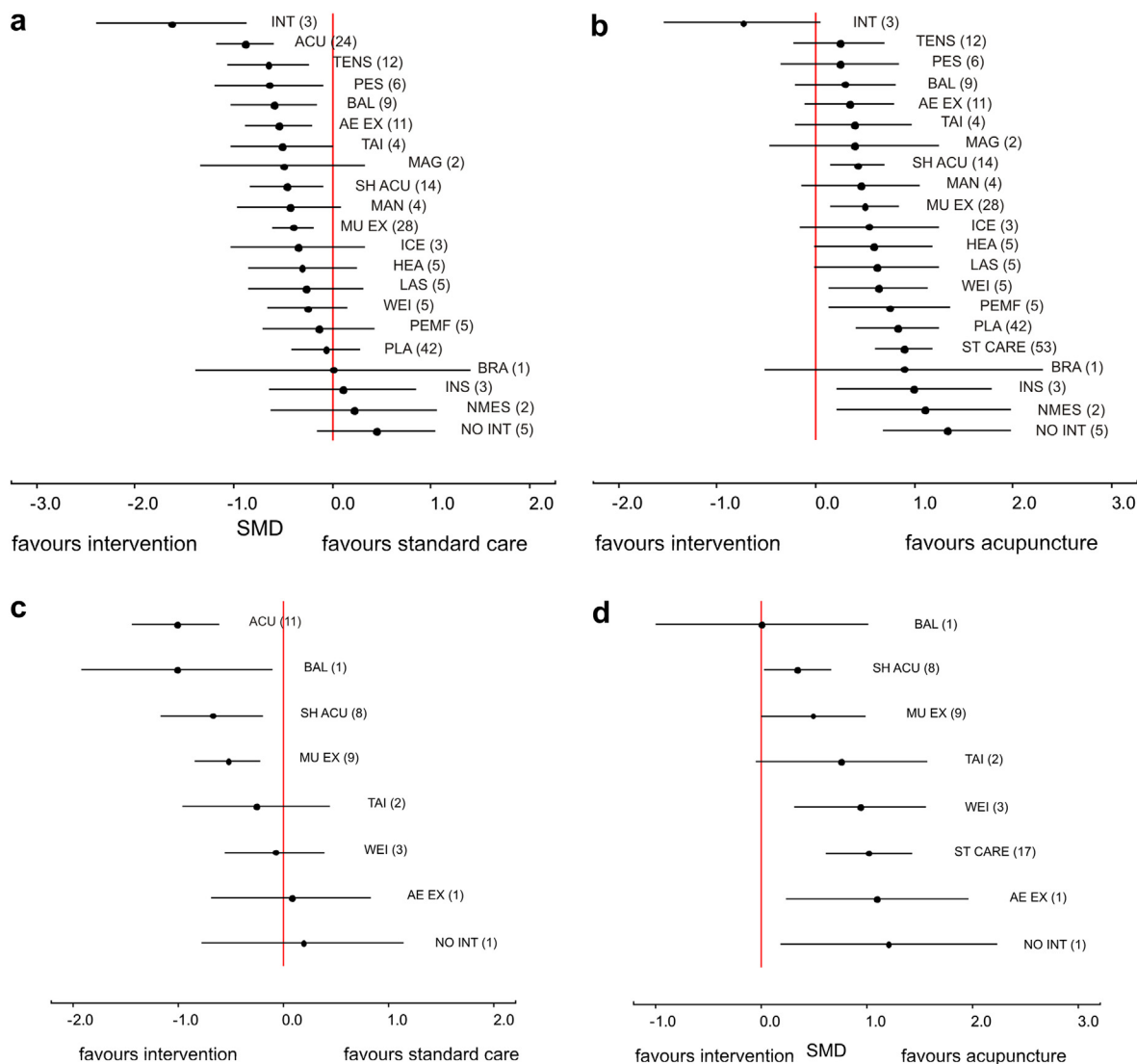


Fig. 1. a: SMDs of each treatment compared to standard care for the analysis including studies of any-quality. b: SMDs of each treatment compared to acupuncture for the analysis including studies of any-quality. c: SMDs of each treatment compared to standard care for the analysis including better-quality studies. d: SMDs of each treatment compared to acupuncture for the analysis including better-quality studies.

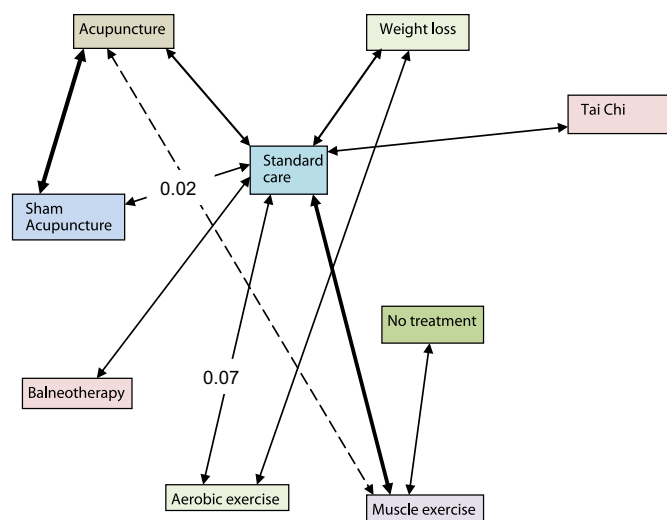


Fig. 2. Network diagram for the end of treatment analysis of better-quality trials. The number of trials and patients included in the analysis are stated in Tables II(c–d). Each solid arrow indicates that there is a data point for that comparison entered into the analysis. The thickness reflects the number of trials. The dotted line reflects an extra comparison in a multi-arm trial. The numbers are a measure of inconsistency: 0 is no inconsistency; 1 is complete inconsistency.

and balneotherapy were the two interventions with the highest rank, a probability statistic calculated from the treatment effect distributions (Table III), although there is uncertainty around these rankings as reflected in the overlapping credible intervals with sham acupuncture, muscle-strengthening exercise and Tai Chi.

Several trials were excluded in a secondary sensitivity analysis based on population or intervention differences, or on extreme data: the results were not sensitive to these changes, although the model fit improved. (see eResults 1).

For the analysis of better-quality studies, no network link could be made with the placebo-controlled studies. We therefore conducted a separate network meta-analysis for these studies. The results, and network, are presented in eResults 1. Both interferential therapy and heat treatment were statistically significantly more effective than placebo, but laser therapy, PES, and insoles were not.

Discussion

The comprehensive network meta-analysis reported here, in which physical treatments for osteoarthritis of the knee were compared with each other within a coherent framework, provides the first estimate of the relative effect of these treatments, which is essential for decision makers. A network meta-analysis provides a

basis of synthesising all the available evidence in a consistent framework, obviating the need to make decisions by subjective inferences from disparate data. Numerous systematic reviews, some summarised in a review of reviews,²⁹ have evaluated the interventions (or classes of interventions) included in this review. However, our analysis represents the use of the most practical methods currently available to compare a large number of different types of treatment, i.e., enabling us to compare the physical treatments (including acupuncture) with each other.

Of the 22 interventions evaluated, eight – interferential therapy, acupuncture, TENS, pulsed electrical stimulation, balneotherapy, aerobic exercise, sham acupuncture, and muscle-strengthening exercise – produced a statistically significant reduction in pain, compared with standard care. Of these, only acupuncture and muscle-strengthening exercise were represented by more than three trials in the sensitivity analysis of better-quality studies, with acupuncture (11 trials) being statistically significantly better than muscle-strengthening exercise (9 trials). Acupuncture, and balneotherapy (1 trial) were the interventions with the highest rank, although there is some uncertainty around these. For the better-quality placebo-controlled studies, interferential therapy (1 trial) showed a strong effect when compared to placebo.

Like a standard meta-analysis, a network meta-analysis requires an assumption of exchangeability between the trials. We sought to minimise concerns which might arise from within- or between-intervention heterogeneity by using an age restriction as part of our inclusion criteria, and by excluding interventions consisting of more than one physical treatment. The patient characteristics appeared broadly comparable across interventions. Some clinical heterogeneity is inevitable in a wide-ranging study such as this, but baseline pain did not appear to vary systematically between interventions, as far as it was possible to tell – given the wide variation of scales used. We conducted sensitivity analyses excluding trials causing heterogeneity. Our analyses used a random effects model to incorporate heterogeneity, and we undertook an evaluation of levels of inconsistency and model fit. Despite this, it is possible there are unknown confounding factors affecting the results of indirect comparisons, although in our results heterogeneity is accounted for in the credible intervals. The majority of trials which used placebo interventions studied electrical or electromagnetic interventions; it is not unreasonable to assume the placebo effects were similar (since the interventions were similar). A further strength of our review is that trials covering a diverse range of interventions were all assessed using the same quality assessment tools; this enabled fair comparisons to be made by evaluating the reliability of the evidence base for each intervention.

However, although we conducted a sensitivity analysis of the better-quality studies, this resulted in fewer trials per comparison,

Table IIc
Results of network meta-analyses for comparisons with standard care: trials of better-quality

Intervention [abbreviations in brackets relate to Fig. 1(a–d)]	No. of trials*	SMD (95% Cr I)	Difference expressed on a WOMAC VAS 0–100 pain scale (95% Cr I)*
Standard care (comparator)		–	–
Acupuncture (ACU)	11 (878)	–1.01 (–1.43 to 0.61)	–16.70 (–23.61 to –10.07)
Balneotherapy (BAL)	1 (40)	–1.01 (–1.92 to –0.11)	–16.65 (–31.73 to –1.74)
Sham Acupuncture (SH ACU)	8 (685)	–0.68 (–1.17 to –0.19)	–11.14 (–19.29 to –3.16)
Muscle-strengthening exercise (MU EX)	9 (450)	–0.52 (–0.84 to –0.22)	–8.62 (–13.92 to –3.58)
Tai Chi (TAI)	2 (51)	–0.26 (–0.96 to 0.44)	–4.29 (–15.87 to 7.23)
Weight loss (WEI)	3 (357)	–0.08 (–0.55 to 0.39)	–1.34 (–9.10 to 6.41)
Aerobic exercise (AE EX)	1 (80)	0.07 (–0.69 to 0.84)	1.23 (–11.30 to 13.78)
No intervention (NO INT)	1 (30)	0.19 (–0.77 to 1.14)	3.11 (–12.72 to 18.77)
Data points: 31 Residual deviance: 31.4			
Between-study standard deviation: 0.39 (95% CI: 0.24–0.58)			

* see Table II(d) footnote.

Table II
Results of network meta-analyses for comparisons with acupuncture: trials of better-quality

Intervention [abbreviations in brackets relate to Fig. 1(a–d)]	No. of Trials*	SMD (95% Cr I)	Difference expressed on a WOMAC VAS 0–100 pain scale (95% Cr I)*
Acupuncture (comparator)		–	–
Balneotherapy (BAL)	1 (40)	0.00 (–0.99 to 1.01)	0.05 (–16.36 to 16.62)
Sham Acupuncture (SH ACU)	8 (685)	0.34 (0.03 to 0.66)	5.57 (0.42 to 10.86)
Muscle-strengthening exercise (MU EX)	9 (450)	0.49 (0.00 to 0.98)	8.08 (0.02 to 16.21)
Tai Chi (TAI)	2 (51)	0.75 (–0.05 to 1.57)	12.42 (–0.81 to 25.84)
Weight loss (WEI)	3 (357)	0.93 (0.31 to 1.57)	15.36 (5.18 to 25.81)
Standard care (ST CARE)	17 (928)	1.01 (0.61 to 1.43)	16.70 (10.07 to 23.61)
Aerobic exercise (AE EX)	1 (80)	1.09 (0.23 to 1.96)	17.94 (3.82 to 32.27)
No intervention (NO INT)	1 (30)	1.20 (0.18 to 2.23)	19.80 (2.94 to 36.81)
Data points: 31 Residual deviance: 31.4.			
Between-study standard deviation: 0.39 (95% CI: 0.24–0.58)			

* To help evaluate these conversions, one study has reported the 'minimal clinically important change' (MCIC) as –15 mm (on a VAS 0–100 scale, and derived from a prior Delphi exercise³⁵), and the 'minimal perceptible clinical improvement' (MPCI) (the smallest change detectable by the patient) as –9.7 mm (on a WOMAC VAS 0–100 scale)³⁴. Another study estimated the 'minimal clinically important improvement' (MCII), although only for pain on movement, as being –19.9 mm on a VAS 0–100 scale; this figure varied by baseline pain score, with patients with less pain having a smaller MCII (10.8 mm) and patients with severe pain having a larger MCII (36.6 mm)³⁶.

and fewer network loops, meaning there is greater uncertainty about the true heterogeneity and about the differences between the direct and indirect evidence. Fewer loops in relation to the size of the network means there is less data to quantify inconsistency and so it is possible that uncertainty associated with inconsistency is not captured in the results. Further limitations are that we could not include all studies in our analyses due to the variable reporting of pain results, and the end of treatment data available was mostly short-term: of the trials which did investigate medium- or long-term effectiveness only a few could provide the data required by our analyses. However, given that the treatments under consideration are not intended as being cures, and that any treatment effect is expected to attenuate over time, a comparison of their maximum effect is not without merit.

It is important that our results are evaluated in context. Methodological limitations exist which are often inherent and unavoidable in clinical trials of physical treatments. Additionally, we found flaws which trialists could have avoided by using better methodology and reporting practices. Most of the studies in our review were rated as being of poor quality, and even many of the better-quality studies were pragmatic trials, where blinding of patients was not possible, i.e., most studies are likely to have been subject to some form of bias. For the trials where patients were not blinded, and treatments were compared with standard care, the overall treatment effect is likely to incorporate non-specific (placebo) effects. We assumed that such non-specific effects were similar across all interventions, but variation may in fact be present.

In light of our results, consideration of what might be the true (or specific) effect of acupuncture is warranted. A Cochrane review reported a statistically significant, clinically relevant, short-term improvement in pain, similar to our findings (acupuncture vs waiting list control, SMD –0.96, 95% CI: –1.19 to –0.72)³⁰. The

comparison of acupuncture with sham acupuncture also showed a similar effect to ours, and was described as being clinically irrelevant (SMD –0.35, 95% CI: –0.55 to –0.15). However, the largest study in this Cochrane analysis, indicating no significant difference, had, for many participants, the primary pain assessment 7 weeks after the end of treatment, and was one of two trials which used an intensive sham needling technique, which may have had physiologic effects. Also, our analysis included a recent large trial (discussed below) which used what appeared to be a very active sham. It is therefore possible that the pooled results from both reviews underestimate the short-term effect of acupuncture. It is also worth noting that the effect size of acupuncture vs sham is of the same order as that seen for NSAIDs vs placebo (SMD 0.32, 95% CI 0.24–0.39), which has also been described as being too small to be clinically significant³¹. An analysis of individual patient data on patients with knee osteoarthritis was recently reported for acupuncture studies (in which the allocation concealment methods had to be unambiguously adequate)³². These results also indicated acupuncture to be more effective than sham acupuncture, and also found a smaller effect size than when acupuncture was compared with no acupuncture (usual care) controls. Non-specific effects therefore seem to play an important role in the pain-alleviating effects of acupuncture. However, for our comparisons, the lack of blinding in trials of the other interventions (where blinding was not possible) in our network of better-quality studies would also be likely to result in non-specific effects contributing to results; it is reasonable to assume that fair comparisons between treatments have therefore been made.

Studies have presented evidence suggesting that sham acupuncture is associated with larger treatment effects than pharmacological and other physical placebos^{12,13}. However, one of two opposing factors – inadequacy of patient blinding by using unsuitable shams, or the use of physiologically active shams – may impact on the effect of sham acupuncture in a given trial; the former may result in an overestimation of the true effect of acupuncture, while the latter may result in an underestimation. In our review important details about sham acupuncture (e.g., depth of insertion) were sometimes poorly reported, or were not reported, so the possibility of further clinical heterogeneity remains. One study in particular had a very active sham, the depth of needle insertion was similar to depths used for (active) acupuncture in some of the other trials; different needle placement formed a large component of the sham. This large study, which found no difference in pain between acupuncture and sham, partly explains the relatively large effect estimate seen for sham acupuncture in our analyses (when compared with standard care)³³.

Table III
Ranking of interventions (using only better-quality studies)

Intervention	No. of trials	Median rank	95% Cr I
Acupuncture	11	2	1–3
Balneotherapy	1	2	1–6
Sham Acupuncture	8	3	2–6
Muscle-strengthening exercise	9	4	2–6
Tai Chi	2	5	2–9
Weight loss	3	6	4–9
Standard care	17	7	6–9
Aerobic exercise	1	8	3–9
No intervention	1	8	3–9

Several quantifications of the clinical relevance of improvements in knee pain scores exist [see Table II(d) footnote]. In this context, our results (derived from the better-quality trials) indicate that acupuncture produces both a ‘minimal perceptible clinical improvement’ (MPCI)³⁴ and quite possibly a ‘minimal clinically important change’^{34,35}, but may only yield a ‘minimal clinically important improvement’³⁶ for patients with low levels of pain. For muscle-strengthening exercise (with evidence from nine trials) a MPCI remains a possibility. Overall, our results suggest that few physical treatments are likely to have a clinically-relevant pain-relieving effect. Other factors to consider when interpreting effectiveness results are safety, the rapidity of onset – and durability – of treatment benefit, and the convenience, cost, and likelihood of patient adherence to treatment³⁷; these factors would clearly differ across the diverse range of interventions we studied, or when comparing them with pharmacological treatments.

Our analyses of the better-quality studies suggest that acupuncture should be considered as one of the physical treatment options for relieving pain due to osteoarthritis of the knee in the short-term. They indicate that balneotherapy, interferential therapy, and heat treatment may also be effective, but the results for all three interventions were informed by single small studies, so a cautious interpretation is warranted. It is worth noting that some of our results on effectiveness do not concur with existing guidance on physical treatments, specifically: EULAR (for insoles, braces, and weight loss), NICE (for TENS, insoles, braces, weight loss, manual therapy, and heat or cooling treatment), ACR (for weight loss, insoles, thermal agents, and Tai Chi), AAOS (for weight loss), and OARSI (for insoles, braces, heat or cooling treatment, TENS, and weight loss). Our analyses found little evidence (of significant differences from standard care, let alone clinically-relevant differences) to support such guidance with respect to treating pain, other than for TENS, where the evidence was of poor quality and likely to be unreliable. It should be remembered though that our review was focused on pain outcomes, rather than on function, disability, or cost-effectiveness.

Larger RCTs, with risk of bias reduced to a minimum and 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. The optimum timing and parameters of treatment for both acupuncture and muscle-strengthening exercise also need to be more clearly defined by future studies.

The evidence available for our network meta-analyses, in which physical interventions for osteoarthritis of the knee were compared with each other within a coherent framework, suggests that overall effectiveness is limited but that acupuncture can be considered as one of the more effective physical treatments for alleviating pain in the short-term. However, despite the large evidence-base found, the methodological limitations associated with many of the trials, indicate that high quality trials of many of the physical treatments are still required.

Contributors

HM and NW conceived the study along with Lesley Stewart and Mark Sculpher. NW, MC, VM, AS and HM developed the protocol. MH performed the searches. SR performed the analyses in collaboration with VM and AS. MC wrote the first draft of the manuscript. MC, NW, VM and RS were responsible for the acquisition of data. All authors critically revised the manuscript for important intellectual content and approved the final version of the manuscript. HM obtained public funding. NW is guarantor.

Conflicts of interest

The authors declare that they have no competing interests.

Funding

This article presents independent research funded by the NIHR under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0707-10186) titled, “Acupuncture for chronic pain and depression in primary care”. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in study design, data collection, data synthesis, data interpretation, or writing the report.

Acknowledgements

Thanks to Philip Conaghan, Mark Roman, Peter Hall, Mark Sculpher, Lesley Stewart, Andrea Manca, Cynthia Iglesias, Tony Danso-Appiah and Ann Hopton for their help at various stages of the review, particularly during protocol development.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.joca.2013.05.007>.

References

1. 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.
2. 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.
3. Arthritis Care. Osteoarthritis Nation: The Most Comprehensive UK Report of People with Osteoarthritis. London: Arthritis Care; 2004.
4. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, *et al.* American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res* 2012;64:465–74.
5. American Academy of Orthopaedic Surgeons. Treatment of Osteoarthritis of the Knee (Non-arthroplasty). Rosemont, IL: American Academy of Orthopaedic Surgeons (AAOS). Available from: <http://www.aaos.org/research/guidelines/OAKguideline.pdf>; 2008.
6. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, *et al.* OARSI recommendations for the management of hip and knee osteoarthritis, part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthritis Cartilage* 2008; 16:137–62.
7. 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 (ESCSIT). *Ann Rheum Dis* 2003;62:1145–55.
8. National Collaborating Centre for Chronic Conditions. Osteoarthritis: National Clinical Guideline for Care and Management in Adults. London: Royal College of Physicians; 2008.
9. Centre for Reviews and Dissemination. Systematic Reviews: CRD’s Guidance for Undertaking Reviews in Health Care. York: Centre for Reviews and Dissemination; 2009.

10. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
11. Corbett M, Rice S, Slack R, Harden M, Madurasinghe V, Sutton A, et al. Acupuncture and Other Physical Treatments for the Relief of Chronic Pain Due to Osteoarthritis of the Knee: A Systematic Review and Network Meta-analysis. CRD Report 40. Centre for Reviews and Dissemination [cited 2012 April 30]. Available from: http://www.york.ac.uk/inst/crd/CRD_Reports/crdreport40.pdf; 2011.
12. Lund I, Naslund J, Lundeberg T. Minimal acupuncture is not a valid placebo control in randomised controlled trials of acupuncture: a physiologist's perspective. *Chin Med* 2009;4:1.
13. Linde K, Niemann K, Meissner K. Are sham acupuncture interventions more effective than (other) placebos? A re-analysis of data from the Cochrane review on placebo effects. *Forsch Komplementmed* 2010;17:259–64.
14. Rodgers M, McKenna C, Palmer S, Chambers D, Van Hout S, Golder S, et al. Curative catheter ablation in atrial fibrillation and typical atrial flutter: systematic review and economic evaluation. *Health Technol Assess* 2008;12:1–198.
15. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
16. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.
17. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932–44.
18. 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.
19. Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *Am J Epidemiol* 2009;169:1158–65.
20. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897–900.
21. Salanti G, Higgins JPT, Ades A, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279–301.
22. Guobing LU, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006;101:447–59.
23. Ades AE, Welton N, Lu G. Mixed treatment comparisons. Community Based Medicine, University of Bristol, 2007 [cited 2011 March 9]. Available from: <https://www.bris.ac.uk/cobm/research/mpes/mtc.html>.
24. Salli A, Ugurlu H, Emlik D. Comparison of the effectiveness of concentric, combined concentric-eccentric and isometric exercises on symptoms and functional capacity in patients with knee osteoarthritis. *Turk Fiz Tip Rehab Derg* 2006;52:61–7.
25. Durmus D, Alayli G, Canturk F. Effects of biofeedback assisted isometric exercise and electrical stimulation on pain, anxiety and depression scores in knee osteoarthritis. *Turk Fiz Tip Rehab Derg* 2005;51:142–5.
26. Jezek J, Laznický J, Kyjovský A. Pulsatile magnetic field within the framework of comprehensive spa treatment in patients with clinically manifest gonarthrosis. *Fysiatr Revmatol Vestn* 1990;68:203–8.
27. Jensen H, Harreby M, Kjer J. Is infrared laser effective in painful arthrosis of the knee? *Ugeskr Laeger* 1987;149:3104–6.
28. 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;16:64–7. Retraction in: Schumacher HR. *J Clin Rheumatol* 2010;16(7):357.
29. 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.
30. Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S, et al. Acupuncture for peripheral joint osteoarthritis. *Cochrane Database Syst Rev* 2010, <http://dx.doi.org/10.1002/14651858.CD001977.pub2>. Issue 1. Art. No.:CD001977.
31. Bjordal JM, Ljunggren AE, Klovning A, Slørdal L. Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: meta-analysis of randomised placebo controlled trials. *BMJ* 2004;329:1317.
32. Vickers AJ, Cronin AM, Maschino AC, Lewith G, MacPherson H, Foster NE, et al. Acupuncture for chronic pain: individual patient data meta-analysis. *Arch Intern Med* 2012;172:1444–53.
33. Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Marcus DM, et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. *Arthritis Care Res* 2010;62:1229–36.
34. 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.
35. Bellamy N, Carrette 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.
36. 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.
37. 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.