Review

Acupuncture for peripheral joint osteoarthritis
A systematic review and meta-analysis

Y. D. Kwon¹,², M. H. Pittler¹ and E. Ernst¹

Objective. To evaluate the evidence for the effectiveness of acupuncture in peripheral joint osteoarthritis (OA).

Methods. Systematic searches were conducted on Medline, Embase, AMED, Cochrane Library, CINAHL, British Nursing Index, PsychINFO and CAMPAIN until July 2005. Hand-searches included conference proceedings and our own files. There were no restrictions regarding the language of publication. All randomized controlled trials (RCTs) of acupuncture for patients with peripheral joint OA were considered for inclusion. Trials assessing needle acupuncture with or without electrical stimulation were considered if sham- or placebo-controlled or controlled against a comparator intervention. Trials testing other forms of acupuncture were excluded. Methodological quality was assessed and, where possible, meta-analyses were performed.

Results. Thirty-one possibly relevant studies were identified and 18 RCTs were included. Ten trials tested manual acupuncture and eight trials tested electro-acupuncture. Overall, ten studies demonstrated greater pain reduction in acupuncture groups compared with controls. The meta-analysis of homogeneous data showed a significant effect of manual acupuncture compared with sham acupuncture (standardized mean difference 0.24, 95% confidence interval 0.01–0.47, \( P = 0.04, n = 329 \)), which is supported by data for knee OA. The extent of heterogeneity in trials of electro-acupuncture prevented a meaningful meta-analysis.

Conclusions. Sham-controlled RCTs suggest specific effects of acupuncture for pain control in patients with peripheral joint OA. Considering its favourable safety profile acupuncture seems an option worthy of consideration particularly for knee OA. Further studies are required particularly for manual or electro-acupuncture in hip OA.

KEY WORDS: Acupuncture, Osteoarthritis, Randomized controlled trial, Systematic review.

Introduction

Osteoarthritis (OA) is the most common form of arthritis, and the most common reason for total hip and total knee replacement [1]. The underlying disease processes of OA involve cartilage degeneration, proliferation and remodelling of subchondral bone structure. Weight-bearing peripheral and axial joints are most often affected [2]. OA is associated with symptoms of pain and functional disability. Physical disability arising from pain and loss of functional capacity reduces the quality of life and increases the risk of further morbidity and mortality [3]. Among adults aged ≥30 yrs, symptomatic knee OA occurs in ~6% and symptomatic hip OA in about 3% [1]. Before the age of 50 yrs, the prevalence of OA in most joints is higher in men than in women, whereas in later life women are more often affected than men in hands, feet and knees [4].

The treatment of OA is largely symptomatic and includes analgesics, NSAIDS, glucosamine, topical analgesics such as capsaicin cream as well as exercise, behavioural interventions and surgical treatment [5]. Most drug treatments are associated with well-documented risks such as gastrointestinal irritation and bleeding, renal and hepatic toxicity, as well as an increased risk of hypertension. Some of these adverse events are most prominent in the elderly—the very group most commonly affected by OA [6].

Non-pharmacological treatments such as acupuncture are therefore attractive. Acupuncture is often used for OA and chronic pain relief [2, 7]. In the US, over 2 million people use acupuncture annually [8]. To evaluate the evidence for the effectiveness of acupuncture in peripheral joint OA, we conducted a systematic review and meta-analysis of randomized controlled trials (RCTs).

Methods

Data sources

Searches were performed in July 2005 using Medline, Embase, AMED, CINAHL, British Nursing Index, PsychINFO, CAMPAIN and Cochrane Library. Search terms used were OA, degenerative arthritis, osteoarthrosis, joint pain, knee pain, hip pain, arthritis, acupuncture, ear acupuncture and electro-acupuncture. In addition, our own files were manually searched and authors were contacted. Original articles were obtained, and all reference lists were scanned for further relevant articles.

¹Complementary Medicine, Peninsula Medical School, Universities of Exeter and Plymouth, UK and ²Department of Rehabilitation and Acupuncture, School of Oriental Medicine, Wonkwang University, South Korea.

Submitted 12 January 2006; revised version accepted 12 May 2006.

Correspondence to: Max H. Pittler, MD PhD, Complementary Medicine, Peninsula Medical School, Universities of Exeter and Plymouth, 25 Victoria Park Road, Exeter EX2 4NT, UK. E-mail: max.pittler@pms.ac.uk
Study selection
All articles were included that reported an RCT in which patients with peripheral joint OA were treated with needle acupuncture with or without electrical stimulation. Trials testing other forms of acupuncture, such as laser acupuncture or acupuncture using moxibustion were excluded. Studies comparing two different forms of acupuncture and those in which no data or statistical comparisons were reported were also excluded. No language restrictions were applied.

Data extraction
Data were extracted independently by two of the authors (Y.D.K. and M.H.P) using a specifically designed data extraction form. For each study, trial design, randomization, blinding and handling of drop outs, inclusion and exclusion criteria, details of treatment and control procedures, main outcomes measure and main results were recorded. Differences during this process were settled by discussion.

Data synthesis
The mean change of pain scores assessed on 100 mm visual analogue scale (VAS) and on the Western Ontario MacMaster Osteoarthritis Index (WOMAC) pain scale compared with baseline were defined as primary outcome measures. They were used to assess the differences between the intervention groups and the control groups. Datapoints chosen were those at the end of the treatment period. Means and 95% confidence intervals (CIs) were calculated using the Cochrane Collaboration’s standard meta-analysis software (RevMan 4.27, Update Software Ltd., Oxford, UK). For some studies, the information was insufficient. In all of these cases, we contacted the original authors requesting further information. The variance of the change was imputed using a correlation factor of 0.4 suggested by the Cochrane Collaboration. The chi-square test for heterogeneity was used to assess whether the distribution of the results was compatible with the assumption that inter-trial differences were attributable to chance variation alone. Homogeneous datasets were meta-analysed using a fixed effects model. Sensitivity analyses were planned, but abandoned due to the small number of trials that provided adequate data. Data which were not suitable for meta-analysis were assessed and weighed in a subjective manner according to important design features relating to study quality (Jadad score), sample size, analyses (e.g. baseline/inter-group comparisons), characteristics of treatment and control groups (e.g. waiting list, sham/placebo control).

Quality assessment
Study quality was assessed using the Jadad score [9]. Taking into account the difficulties in blinding, we used a modification of this scale [10]. For a total of 5 points, we awarded 1 point each: if the study was described as randomized; for appropriate method; if subjects were blinded to intervention; if evaluator was blinded to intervention; for description of withdrawals and dropouts. Subject blinding was assumed where the control intervention was indistinguishable from acupuncture, even if the word ‘blinding’ did not occur in the report. Trials with 4 or 5 points were considered to be of high quality. For knee and hip OA, acupuncture treatment was considered adequate if a minimum of four points were needled for a minimum of 20 min, at least six treatments were given with a frequency of at least one session weekly, sufficient electrical stimulation for a minimal sensation was given or deqi (needle sensation) was achieved.

Study quality
Two sham-controlled studies reported assessor and subject blinding and scored the maximum on the modified Jadad scale [33, 35], whereas eight studies scored 4 points [26–28, 30, 31, 36, 38, 40]. Patient blinding was judged to have been achieved in six studies [25, 28, 30, 33, 35, 36]. The assessor was reported as blinded in 13 studies (Table 1) [25, 27, 28, 30, 31, 33–36, 38–40]. The treatment duration was 20–30 min in most trials [23–25, 27–39].

Outcomes
Overall, 10 studies suggested greater pain reduction in acupuncture groups compared with control groups [25, 27, 30–32, 35–38, 40]. Of the 10 high-quality studies scoring 4–5 points [26–28, 30, 31, 33, 35, 36, 38, 40], seven trials reported acupuncture to be superior to a range of control interventions [27, 30, 31, 35, 36, 38, 40]. The extent of heterogeneity in the dataset prevented a meaningful meta-analysis across all trials.
<table>
<thead>
<tr>
<th>First author (date)</th>
<th>Study design, quality score</th>
<th>Sample size, OA site</th>
<th>Experimental treatment</th>
<th>Control intervention</th>
<th>Main outcome measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christensen [27] (1992)</td>
<td>2 parallel groups, assessor blind 4</td>
<td>32, knee</td>
<td>A: formula AT, ST-34, 35, 36, SP-10, XL-2, LI-4, 2 sessions weekly for 3 weeks</td>
<td>B: waiting list</td>
<td>(1) Pain VAS</td>
<td>(1) Intergroup difference (&lt;0.01)</td>
</tr>
<tr>
<td>Witt [30] (2005)</td>
<td>3 parallel groups, patient blind 4</td>
<td>300, knee</td>
<td>A: Chinese AT, ST-34, 35, 36, SP-9, 10, UB-40, Kid-10, GB-33, 34, Li-8, Heding. Xiyian, and selective additional points (SP-4, 5, 6, ST-6, UB-20, 57, 58, 60,62, Kid-3), 1–2 sessions weekly for 8 weeks, 12 sessions in total</td>
<td>B: sham (superficial insertion on non-AT points) C: waiting list</td>
<td>(1) WOMAC (pain, stiffness and function)</td>
<td>(1) Differences compared with B, C (&lt;0.01)</td>
</tr>
<tr>
<td>Petrou [25] (1988)</td>
<td>2 parallel groups, assessor and patient blind 3</td>
<td>31, knee</td>
<td>A: formula AT, ST-35, 36, 43, EX-31, 32, UB-40, LI-4, 3 sessions weekly for 3 weeks, 8 sessions in total</td>
<td>B: sham (superficial insertion into near non-classical, non-tender points)</td>
<td>(1) Starting pain (4-point scale) (2) Night pain (4-point scale) (3) Walking pain, pain descending stairs and walking time</td>
<td>(1) Intergroup difference (&lt;0.001)</td>
</tr>
<tr>
<td>Takeda [28] (1994)</td>
<td>2 parallel groups, assessor and patient blind 4</td>
<td>40, knee</td>
<td>A: formula AT, ST-35, GB-34, SP-9, Extra-31, 32, 3 sessions weekly for 3 weeks</td>
<td>B: sham (superficial insertion about 1 inch from the AT points)</td>
<td>(1) McGill Pain Questionnaire (2) WOMAC (pain, stiffness and function) (3) Pain threshold (dolorimeter)</td>
<td>(1)–(3) No intergroup differences</td>
</tr>
<tr>
<td>Molsberger [30] (1994)</td>
<td>2 parallel groups, assessor and patient blind 4</td>
<td>97, knee</td>
<td>A: formula AT, ST-34, 35, 36, UB-9, 10, GB-34, Extra-31, 32, 2 sessions weekly for 5 weeks</td>
<td>B: sham (superficial insertion at non-AT points on homolateral leg)</td>
<td>(1) Pain VAS (2) Function (Lysholm score)</td>
<td>(1) Intergroup difference (&lt;0.01)</td>
</tr>
<tr>
<td>Fink [33] (2001)</td>
<td>2 parallel groups, assessor and patient blind 5</td>
<td>67, hip</td>
<td>A: formula AT, GB-30, 31, 34, BL-37, ST-40, BL-54, six ‘ah shi’ points, 10 sessions in total within 3 weeks</td>
<td>B: sham (insertion at same depth at least 5 cm away from AT point)</td>
<td>(1) Pain VAS (2) Functional impairment (hip function index)</td>
<td>(1) No intergroup difference</td>
</tr>
<tr>
<td>Berman [35] (2004)</td>
<td>3 parallel groups, assessor and patient blind 5</td>
<td>570, knee</td>
<td>A: Chinese EA, GB-34, SP-9, ST-35, 36, Xiyian, UB-60, GB-39, SP-6, Kid-3, 23 sessions in total within 26 weeks</td>
<td>B: sham (combined insertion on abdomen and noninsertion procedure on same AT points as A) C: education</td>
<td>(1) WOMAC (pain) (2) WOMAC (function) (3) Patient global assessment</td>
<td>(1)(2) Differences compared with B (&lt;0.01) and C</td>
</tr>
<tr>
<td>Vas [40] (2004)</td>
<td>2 parallel groups, assessor blind 4</td>
<td>97, knee</td>
<td>A: formula EA, GB-34, SP-9, EX-LE5, ST-36, Ki-3, SP-6, LI-4, ST-40, 1 session weekly for 12 weeks plus diclofenac</td>
<td>B: placebo EA (placebo needles, no perforation of the skin, electrical stimulation) plus diclofenac</td>
<td>(1) Pain VAS (2) WOMAC pain (3) WOMAC stiffness (4) WOMAC function</td>
<td>(1)–(4) Intergroup differences (&lt;0.001)</td>
</tr>
<tr>
<td>Dickens [26] (1989)</td>
<td>2 parallel groups, assessor blind 4</td>
<td>12, thumb</td>
<td>A: formula AT, point not described, 6 sessions within 2 weeks</td>
<td>B: mock TENS</td>
<td>(1) Pain VAS</td>
<td>(1) Changes compared with baseline for AT (P=0.02)</td>
</tr>
<tr>
<td>Yurtkuran [29] (1999)</td>
<td>4 parallel groups, assessor blind 3</td>
<td>100, knee</td>
<td>A: formula EA, SP-9, GB-34, ST-34, 35, 5 sessions weekly for 2 weeks</td>
<td>B: TENS C: ice massage D: placebo TENS (electrodes disconnected)</td>
<td>(1) Present pain intensity (5-point scale) (2) Stiffness (3) 50-foot walking time (4) Quadriceps muscle strength (5) Active knee flexion</td>
<td>(1)–(5) Changes compared with baseline for EA (P&lt;0.05) (1)–(5) No differences between treatment groups</td>
</tr>
<tr>
<td>Ng [34] (2003)</td>
<td>3 parallel groups, assessor blind 3</td>
<td>24, knee</td>
<td>A: formula EA, ST-35, EX-LE4, 8 sessions in total within 2 weeks</td>
<td>B: TENS C: general education on knee care only</td>
<td>(1) Numerical pain rating scale (2) Passive range of movement (3) The timed up and go test (TUGT)</td>
<td>(1) Change in A, B compared with baseline (2) No changes in either group (3) Change in A compared with baseline</td>
</tr>
<tr>
<td>Ammer [24] (1988)</td>
<td>2 parallel groups, open 1</td>
<td>28, knee</td>
<td>A: formula AT, UB-54, Liv-9, GB-34, 30, 32, ST-36, 2 sessions weekly for 4 weeks, 8 sessions in total</td>
<td>B: physical therapy</td>
<td>(1) Starting pain, walking pain (4-point scale) (2) Total pain (4-point scale)</td>
<td>(1) No intergroup differences</td>
</tr>
</tbody>
</table>

Continued
### TABLE 1. Continued

<table>
<thead>
<tr>
<th>First author (date)</th>
<th>Study design, quality score</th>
<th>Sample size, OA site</th>
<th>Experimental treatment</th>
<th>Control intervention</th>
<th>Main outcome measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stener-Victorin [39] (2004)</strong></td>
<td>3 parallel groups, open 3</td>
<td>45, hip</td>
<td>A: formula EA, BL-54, 36, GB-29, 30, 31, ST-31, GB-34, BL-60, 2 sessions weekly for 5 weeks, 10 sessions in total plus education</td>
<td>B: hydrotherapy plus education</td>
<td>C: education alone</td>
<td>(1) Pain VAS (related to motion and load) (2) Daytime pain VAS (3) Nighttime pain VAS (4) Lequesne's functional index (5) 50 ft walk time</td>
</tr>
<tr>
<td><strong>Haslam [32] (2001)</strong></td>
<td>2 parallel groups, open 3</td>
<td>32, hip</td>
<td>A: formula AT, GB-29, 30, 34, 43, ST-44, LI-4, four 'ah shi' points, 1 session weekly for 6 weeks, 6 sessions in total</td>
<td>B: advice plus a set of five exercises</td>
<td>(1) modified version of WOMAC (pain, stiffness and physical function)</td>
<td>(1) Intergroup difference (P=0.02)</td>
</tr>
<tr>
<td><strong>Jia [23] (2005)</strong></td>
<td>3 parallel groups, open 2</td>
<td>120, knee</td>
<td>A: Chinese AT, ST-34, 36, SP-9, 10, GB-34, medial Xiyan, lateral Xiyan, 'ah shi' points, 1 session daily, 45 sessions within 2 months</td>
<td>B: Chinese AT plus functional exercise</td>
<td>C: functional exercise (oral)</td>
<td>(1) effective rate (4-point scale) (2) recurrence rate</td>
</tr>
<tr>
<td><strong>Berman [31] (1999)</strong></td>
<td>Partial crossover, assessor blind 4</td>
<td>73, knee</td>
<td>A: Chinese EA, GB-34, SP-9, ST-35, 36, Xiyan, UB-60, GB-39, SP-6, Kid -3, 2 sessions weekly for 8 weeks, 16 sessions in total</td>
<td>B: conventional therapy alone (oral)</td>
<td>(1) WOMAC (total, pain and disability) (2) Lequesne indices</td>
<td>(1)(2) Intergroup differences (P&lt;0.001)</td>
</tr>
<tr>
<td><strong>Sangdee [37] (2002)</strong></td>
<td>4 parallel groups, assessor blind 3</td>
<td>193, knee</td>
<td>A: formula EA, ST-35, Liv-8, medial Xiyan, trigger point, 3 sessions weekly for 4 weeks, 12 sessions in total</td>
<td>B: formula EA plus diclofenac, C: diclofenac plus placebo EA (patch electrodes to the selected AT points) D: placebo EA plus placebo tablet</td>
<td>(1) Pain VAS (2) WOMAC pain (3) WOMAC stiffness (4) WOMAC disability (5) Lequesne's functional index (6) 50 ft walk time</td>
<td>(1) Intergroup difference (P&lt;0.05), A vs C and A vs D (2) Intergroup difference (P&lt;0.05), A vs D (3)(4)(6) No intergroup differences</td>
</tr>
<tr>
<td><strong>Tukmachi [38] (2004)</strong></td>
<td>3 parallel groups, assessor blind 4</td>
<td>30, knee</td>
<td>A: Chinese EA, LI-4, SP-9, 10, Xiyan, GB-34, ST-36, LR-3, BL-40, 57, 2 sessions weekly for 5 weeks, 10 sessions in total</td>
<td>B: Chinese EA plus symptomatic medication</td>
<td>C: symptomatic medication</td>
<td>(1) Pain VAS (2) WOMAC (pain) (3) WOMAC (stiffness)</td>
</tr>
</tbody>
</table>

Manual acupuncture was compared with waiting list control groups in two studies, which reported intergroup differences for pain on VAS and WOMAC scales (Table 1). When compared with superficial needling as sham control, intergroup differences on pain measures were reported in three of four trials (Table 1). Three trials provided sufficient data for meta-analysis, which showed a significant effect of manual acupuncture compared with sham for treating peripheral joint OA (standardized mean difference 0.24, 95% CI 0.01–0.47, P = 0.04, n = 329, Fig. 2). This was confirmed by the results for knee OA (Fig. 3). Compared with mock TENS, there were no intergroup differences [26] and compared with conventional therapy the results were mixed [23, 24, 32]. In one case of insufficient data reporting [30], additional information was requested from the original author but has not been received so far.

Electro-acupuncture was compared against sham or placebo acupuncture (Table 1) in three trials [35, 37, 40], which reported intergroup differences for pain on VAS and WOMAC scales. Electro-acupuncture was compared with conventional therapy in seven trials [29, 31, 34, 35, 37–39]. Intergroup differences in favour of electro-acupuncture as compared with oral medication such as diclofenac on pain VAS [37, 38], and WOMAC scales [31, 38]. Intergroup differences on pain were not reported compared with transcutaneous electrical nerve stimulation (TENS) [29, 34] and confirmed by the results for knee OA (Fig. 3). Compared with placebo TENS [26, 29]. Compared with conventional medication, electro-acupuncture was superior to NSAIDs on VAS [37, 38] and WOMAC [31, 38]. Further trials are required to confirm these data.

Discussion

A previous review indicated favourable evidence of acupuncture for pain control when compared with sham acupuncture, and concluded that the evidence suggests that acupuncture may play a role in the treatment of knee OA [42]. Since the publication of this article, a number of further RCTs were published. Whereas the previous review [42] could not perform meta-analyses due to the numerous types of control groups and insufficient reporting of data, the newly identified trials enabled us to conduct meta-analyses. Our meta-analysis of a homogeneous dataset suggests significant pain reduction for patients with peripheral joint OA, particularly knee OA (Figs. 2 and 3). A best evidence synthesis [43], which contains a degree of subjectivity, suggests that, overall, the data for manual acupuncture could be classified as a fairly strong evidence.

We opted to assess quality using a modified Jadad scale whereby a point is scored for subject and assessor blinding separately. Of the 18 RCTs, only five trials were both patient-blinded and assessor-blinded [25, 28, 30, 33, 35], whereas four studies [23, 24, 32, 39] did not make an attempt at either subject or assessor blinding. Trials with inadequate levels of blinding are likely to exaggerate the treatment effects and thus limit the reliability of the findings [44, 45]. Of course, many other systems to assess methodological quality exist; some of which are more elaborate. However, the Jadad score is an accepted, frequently used and easy to apply score. Jüni et al. [46] have shown that there are no strong advantages of these instruments over the Jadad score.

Acupuncture was compared with various types of control interventions. Depending on the nature of the control intervention, different conclusions can be drawn. Manual acupuncture seems to reduce pain compared with waiting-list controls. This may suggest effectiveness against the natural course of OA but does not allow for placebo effects. Trials of manual acupuncture were also conducted against sham acupuncture, and the meta-analysis of three RCTs, which provided sufficient data, suggests a positive effect. In addition, five trials [30, 35–37, 40] reported a degree of blinding, suggesting beneficial effects compared with sham or placebo control for VAS or WOMAC pain scores. This indicates analgesic effects of acupuncture beyond a placebo response. There were no intergroup differences compared with placebo TENS [26, 29]. Compared with conventional medication, electro-acupuncture was superior to NSAIDs on VAS [37, 38] and WOMAC [31, 38]. Further trials are required to confirm these data.

The placebo effect of acupuncture as a treatment for pain can be impressive. Several placebo or sham acupuncture methods have been tried. These range from puncturing the skin outside acupuncture points, eliciting a sensation on the skin without sensation, and nine of 14 studies on OA of the thumb joint [26].

FIG. 2. Meta-analysis of randomized controlled trials (specifics are mentioned at the top of each figure).

FIG. 3. Meta-analysis of randomized controlled trials (specifics are mentioned at the top of each figure).
puncturing the skin, superficially puncturing the skin without stimulation or in cases of electro-acupuncture, using electro-stimulators with disconnected cables. To reliably account for the placebo effect, it is crucial that the sham procedure is indistinguishable from the real treatment. Therefore, the success of the blinding procedures should be assessed.

This systematic review has several limitations. We cannot be absolutely certain that our searches located all relevant RCTs, which is a limitation that indeed applies to systematic reviews in general. Although strong efforts were made to retrieve all RCTs on the subject, it is conceivable that some were not found. In this study, a large number of different databases were searched and there were no restrictions in terms of publication language. Further limitations include the paucity and the often suboptimal quality of the primary data. It must be noted that design features such as placebo or blinding are difficult to incorporate for acupuncture trials and that research funds for acupuncture studies are in short supply. Even though the total number of 18 RCTs is encouraging, it is too small considering the heterogeneity of the overall dataset.

One argument for using acupuncture in the management of painful musculoskeletal conditions is that it is safer than standard drug treatments [10]. Serious adverse effects of acupuncture have been reported, although these may be rare [47]. Relative to those of standard drug treatments these may be infrequent or even negligible [2, 48]. Several prospective studies have shown that mild adverse effects after acupuncture occur in about 7% of all cases [49].

In conclusion, sham-controlled RCTs suggest specific effects of acupuncture for pain control in patients with peripheral joint OA. Considering its favourable safety profile, acupuncture seems an option worthy of consideration particularly for knee OA. Further studies are required particularly for manual or electro-acupuncture for hip OA.

<table>
<thead>
<tr>
<th>Key messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Acupuncture is often used for OA and chronic pain relief.</td>
</tr>
<tr>
<td>- The meta-analysis of three trials showed a significant effect of manual acupuncture compared with sham acupuncture.</td>
</tr>
<tr>
<td>- Beneficial effects were confirmed for knee OA.</td>
</tr>
</tbody>
</table>

Acknowledgement

Y.D.K. was supported by a Soongsan grant from Wonkwang University, South Korea in 2005.

The authors have declared no conflict of interest.

References

26. Dickens W, Lewith GT. A single-blind, controlled and randomised clinical trial to evaluate the effect of acupuncture in the


