Effectiveness of acupuncture in the treatment of pain from osteoarthritis of the knee

A. Ferrández Infante, L. García Olmos, A. González Gamarra, M.J. Meis Meis and B.M. Sánchez Rodríguez

Aim. To determine the effectiveness of acupuncture in controlling pain from arthritis of the knee.

Design. Systematic review.

Data sources. MedLine, the Cochrane Library.

Study selection. Of the 9 studies located, only 4 met the inclusion criteria. All were controlled, randomized clinical trials that studied the effect of acupuncture only in the knee joint.

Data extraction. Primary outcome variables were intensity of pain, overall measure (general improvement, proportion of patients who recovered, subjective improvement in symptoms) and functional status. As secondary outcome measures we used objective physiological measures (range of knee movement, muscle strength, time needed to walk a certain distance, time needed to climb a certain number of stairs), general health status, and other information such as medication needed and side effects.

Results. There was moderately strong evidence that acupuncture was more effective in treating knee joint pain than no treatment. The difference can be explained by its marked placebo effect.

Conclusions. There is currently insufficient evidence to recommend acupuncture as a treatment for pain from osteoarthritis of the knee. Additional, better designed studies are needed to determine the actual role of acupuncture in this disease.

Key words: Arthritis. Knee. Acupuncture. Pain.
Introduction

The most frequent location of osteoarthritis, a highly prevalent disease, is the knee joint. The clinical features of arthritis of the knee joint give rise to a functional disability of the lower limb, which significantly impairs the patient’s quality of life. One of the main goals of arthritis treatment is pain relief. However, conventional treatment with analgesics and antiinflammatories is often ineffective, and is frequently linked with the appearance of serious side effects. Acupuncture, the main branch of traditional Chinese medicine, is a therapy based on principles related with ancient Chinese philosophy, and conceives disease as an imbalance between the forces of yin and yang. Although the exact mechanisms responsible for the therapeutic action of acupuncture remain unknown, many studies have clearly shown that this therapy induces physiological changes able to contribute to pain relief. Although the effectiveness of acupuncture in several diseases has been the subject of systematic reviews, no such studies have analyzed its effectiveness in controlling the pain of knee joint arthritis. This systematic review was undertaken to clarify whether acupuncture is an effective treatment for pain related with osteoarthritis of the knee.

Methods

Study selection
The studies we examined compared the following:

   a) Acupuncture plus standard treatment or medication versus standard treatment or medication alone.
   b) Acupuncture alone compared to a group that received no treatment.
2. Acupuncture compared to placebo or sham acupuncture.
   a) Acupuncture compared to placebo acupuncture (placing the needle on the surface of the skin without penetration).
   b) Acupuncture compared to sham acupuncture (placing the needles on sham points near the true acupuncture point).

The only types of study included in the present review were randomized, controlled clinical trials (RCTs). All other types of study were excluded.

We chose RCTs that studied individuals in whom osteoarthritis of the knee had been diagnosed on the basis of clinical and radiological features, regardless of the duration of the disease. Randomized clinical trials that investigated the effects of acupuncture on arthritis in locations other than the knee were excluded.

Articles that reported on acupuncture techniques based on insertion of the needles were included regardless of whether the needles were inserted at the classical points along the meridians used in classical Chinese medicine or at other points outside the meridians (used in modern acupuncture). We did not take into account the technique used to stimulate the needles (manual, electrical, etc.), although we excluded from the present review those studies in which needles were not actually inserted, eg, acupressure and laser therapies.

We included only those RCTs that used at least one of the primary outcome measures considered here: intensity of pain (visual analog scale, VAS) a global measure (global improvement, proportion of patients who recovered, subjective recall of symptoms) and functional status (WOMAC index).

As secondary outcome measures we used objective physiological measures (range of knee joint movement, muscle strength, time to walk a certain distance, time to climb a certain number of steps), general health status (SF-36 Health Survey, Nottingham...
Effectiveness of acupuncture in the treatment of pain from osteoarthritis of the knee

TABLE 1  Results of methodological quality evaluation of the trials included in the review

<table>
<thead>
<tr>
<th>RCT/criteria</th>
<th>B1</th>
<th>B2</th>
<th>C</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>L</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christensen et al</td>
<td>C</td>
<td>C</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+*</td>
<td>–</td>
<td>–*</td>
<td>+</td>
<td>–*</td>
</tr>
<tr>
<td>Takeda et al</td>
<td>C</td>
<td>C</td>
<td>C*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Berman et al</td>
<td>+</td>
<td>C</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Yurtkuran et al</td>
<td>C</td>
<td>C</td>
<td>+*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

B1: was the method of randomization described, and was it appropriate?; B2: was treatment allocation appropriately concealed?; C: were the groups initially similar in terms of key prognostic indicators?; E: were the care-provider blinded?; F: were co-interventions that might explain the results controlled for?; G: could the rate of compliance have biased the results?; H: were the patients blinded?; I: were the observers blinded?; L: could the withdrawal and drop-out rates have biased the results?; N: was the duration of follow-up similar in both groups?10

*Criteria on which the evaluators disagreed. C indicates unclear

Results

The MedLine search identified seven potentially eligible studies,11-17 two of which were found to be the same report in English11 and Danish.17 Two other studies were excluded because they were in German15 and Russian,16 as none of the reviewers was able to read these languages. We concluded that the abstract of the German study did not provide enough data to decide whether it met our selection criteria for inclusion. The abstract of the Russian study indicated that this trial did not investigate arthritis only in

Health Profile, disease impact profile), and other information such as medication needed and side effects.

Location and selection of studies
Randomized clinical trials that fulfilled the inclusion criteria were identified as described below:

– Computerized search of the MedLine database (1966-2000) and the Complementary Medicine field in the Cochrane Library (2000, volume 1) with the key words randomized controlled trial, controlled clinical trial, double-blind method, single-blind method, clinical trial, human, arthritis, osteoarthritis, genarthritis, knee, acupuncture, pain.

– Analysis of the references cited in the studies identified in the preceding step.

One reviewer (L.G.O.) developed the search strategies for MedLine and the Cochrane Library, and recorded to a PC-readable file the author, title, keywords and abstract of all articles identified. Later, two other reviewers (A.F.I. and B.G.R.) analyzed this information to identify trials that potentially met our inclusion criteria. The full text of the articles that reported these trials was obtained, and the same reviewers independently checked them against our selection criteria. The consensus method was used to solve disagreements regarding inclusion in the final sample of RCTs, and a third reviewer (A.G.G.) was consulted when no consensus was reached.

Assessment of methodological quality
The methodological quality of each RCT was evaluated independently by two reviewers who were blind as to the authors, affiliations and journal. Information on the authors, affiliations and journal was masked by a person uninvolved in the actual review of the texts. Consensus was used to solve disagreements, and when disagreement persisted a third reviewer was consulted.

The methodological quality of each RCT was evaluated with the checklist recommended in the guidelines for systematic reviews of the Cochrane Collaboration Back Review Group. We used only the ten features that referred to internal validity (criteria B1, B2, C, E, F, G, H, I, L, and n)10 (table 1).

Data extraction
Two reviewers extracted data for the primary and secondary outcome measures. We also recorded data on the characteristics of the study population (duration of pain from arthritis, age, sex) and on the acupuncture technique and reference interventions (type of intervention, frequency, intensity, duration and location).

Data analysis
For all studies we evaluated homogeneity, description of the methods, and presentation of the results. Because the studies differed in duration of the disease, type of acupuncture used, and outcome indicators, statistical analysis with pooled data was considered inappropriate, and we opted to undertake a qualitative review aimed at assigning levels of evidence to the effectiveness of acupuncture on the basis of methodological quality and the results of the different studies.

Level 1: strong evidence, provided by consistent findings across several high-quality RCTs.
Level 2: moderate evidence, provided by consistent findings in one high-quality RCT and one or more low-quality RCT.
Level 3: limited evidence, provided by consistent findings in one or more low-quality RCTs.
Level 4: no evidence, if no relevant RCTs were found or if the results across studies were contradictory.

An RCT was considered of high quality when more than five of the 10 validation criteria were scored as positive. The literature was considered contradictory when the finding for a specific outcome indicator was considered clearly positive in some studies but clearly negative in others, in at least one-third of the studies. The findings of each study were evaluated by blinded reviewers as positive, neutral, negative or unclear (if the results were poorly presented or if a significant error was detected in the study design).

Results

The MedLine search identified seven potentially eligible studies,11-17 two of which were found to be the same report in English11 and Danish.17 Two other studies were excluded because they were in German15 and Russian,16 as none of the reviewers was able to read these languages. We concluded that the abstract of the German study did not provide enough data to decide whether it met our selection criteria for inclusion. The abstract of the Russian study indicated that this trial did not investigate arthritis only in
the knee, and we therefore judged that it almost certainly did not meet our inclusion criteria. After studying the four remaining articles,11-14 both reviewers agreed that they were all eligible for inclusion.

The search in the Complementary Medicine field of the Cochrane Library identified two additional studies that might have been suitable for inclusion.18,19 Perusal of the full text led us to exclude both studies because they investigated arthritis in several other locations besides the knee.

No further studies suitable for inclusion were identified from the reference lists of the articles we examined. The reviewers disagreed on 12 of the 40 scores (30%). After the consensus meeting, 9 scores remained unresolved. A third reviewer decided the final result on the basis of the comments provided by both of the dissenting reviewers.

Only one study was found to be of high quality, with a positive score on more than 50% of the checklist items. Table 1 summarizes the results of the methodological quality evaluation of each RCT. The criteria that received the worst scores were B1, B2, e and G, and we therefore conclude that the randomization methods were not adequately described, the care providers were not blinded, and insufficient information was provided on compliance with treatment.

Two studies compared acupuncture and standard care alone.11,13 One study compared acupuncture and a placebo,14 and one other RCT compared the effects of real and sham acupuncture.12 Only one of the RCTs specified the setting of the study (hospital).14 Accreditation or certification of the acupuncturist was reported in only one RCT.12

Only one of four RCTs failed to note the duration of the disease.12 Sample sizes were small, ranging from 32 to 100 patients, and basic information on the participants’ age and sex was provided in all four studies. Each study used different acupuncture points, and treatment duration and frequency also differed between studies. In two studies the needles were stimulated until the Teh Chi sensation was obtained,11,13 and in one study the entire length of the needle was inserted if the Teh Chi effect was not attained.12

Table 2 summarizes the characteristics of each study included in the systematic review. After the best-evidence synthesis, the results for each type of comparison were as described below:

**Acupuncture versus no treatment**

We identified two studies that compared acupuncture plus standard treatment versus standard treatment alone.11,13 Both studies found that combined treatment controlled pain significantly better than did conventional treatment. Only one RCT was of high quality.13 We concluded that moderate evidence exists for the effectiveness of treatment with acupuncture versus no treatment.

**Acupuncture versus placebo or sham acupuncture**

We identified one study that compared acupuncture versus placebo,14 and one other that compared the effects of real and sham acupuncture.12 The former study was of poor quality, and showed that acupuncture improved symptoms significantly more than did placebo. The latter study was also of poor quality, and the results showed no statistically significant difference bet-

### Table 2

**Characteristics of the trials included in the review**

<table>
<thead>
<tr>
<th></th>
<th>Christensen et al</th>
<th>Takeda et al</th>
<th>Berman et al</th>
<th>Yurtkuran et al</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>32</td>
<td>40</td>
<td>58</td>
<td>100</td>
</tr>
<tr>
<td>Sex</td>
<td>11M/21W</td>
<td>20M/20W</td>
<td>23M/35W</td>
<td>9M/91W</td>
</tr>
<tr>
<td>Age (range, mean), years</td>
<td>48-75 (69.2)</td>
<td>61.6 49-86 (65.6)</td>
<td>45-70 (58.1)</td>
<td></td>
</tr>
<tr>
<td>Setting</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Hospital</td>
</tr>
<tr>
<td>Accreditation</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Frequency of treatment, sessions/week</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Duration of treatment, weeks</td>
<td>17 (short-term)</td>
<td>3</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>Teh chi effect required</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Technique</td>
<td>Acupuncture</td>
<td>Acupuncture</td>
<td>Acupuncture, electroacupuncture</td>
<td>Electroacupuncture</td>
</tr>
<tr>
<td>Side effects reported</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Control</td>
<td>Conventional treatment</td>
<td>Sham acupuncture</td>
<td>Conventional treatment</td>
<td>Placebo</td>
</tr>
<tr>
<td>Points used</td>
<td>Xiyan, S34, S35, S36, Sp10, LI4(^b)</td>
<td>Xiyan, EP<em>31</em>, S35, Sp9, GB34(^b)</td>
<td>Xiyan, S35, S36, K3, B60, Sp6, Sp9, GB14, GB29(^b)</td>
<td>S34, S35, GB34, Sp9(^b)</td>
</tr>
</tbody>
</table>

\(^a\)This study did not report the age range.

\(^b\)S indicates stomach; Sp, spleen; LI, large intestine; EP, extra point; GB, gallbladder; K, kidney; B, bladder.
ween real and sham acupuncture, although the results were better with actual acupuncture. From this study we concluded that there was no evidence that real acupuncture was more effective than placebo or sham acupuncture.

**Acupuncture versus conventional treatment**

We found no RCTs that compared acupuncture and conventional treatment, and have no information that would allow us to draw any conclusions about these studies.
This review examined four RCTs. The number of internal validity criteria that were scored as positive in our evaluation of methodological quality was low. Only one study received a positive score on more than 50% of the criteria, and could be considered of high quality. Although we would have liked to include as many RCTs as possible in our study, we may have omitted some relevant studies. Our lack of knowledge of languages other than English and Spanish prevented us from analyzing two studies that might have been included. Moreover, the number of journals published in languages other than English and covered in MedLine and the Cochrane Library databases is small, a factor that might have led to the exclusion of other important RCTs. This may have biased our review, as there is evidence that the omission of studies in languages other than English can significantly alter the results of systematic reviews. In future studies we intend to undertake a review of all RCTs located regardless of the language of publication. We also intend to consult other important databases such as Embase, in order to analyze as many RCTs as possible.

The methodological quality of the RCTs included in the present review was rather poor. We defined methodological quality on the basis of internal validity criteria that concerned characteristics of the study possibly related with biases of selection, performance, attrition and detection. The inclusion of poor-quality RCTs in our review might be assumed to have biased our results. However, in our analysis of the best evidence, strong evidence could be provided only by high-quality RCTs, a factor that rules out, to a large extent, this possible source of bias. Although the levels of evidence we established were arbitrary, we do not believe that a different rating system would have led to different findings.

The decision to use best evidence analysis rather than meta-analysis was based on the heterogeneity of the RCTs in clinical characteristics: the studies differed widely in duration of the disease, type of acupuncture, duration of follow-up and outcome measures. Because of the lack of consensus on the best way to evaluate the strength of evidence, the levels of evidence we set here were arbitrary. Although we feel it is unlikely, a different rating system of the levels of evidence might have produced different results.

The insertion points used, frequency of treatment and duration of treatment were described in all RCTs reviewed here; however, we found that treatment protocols differed markedly. None of the studies described adaptation of the treatment to a particular individual’s symptoms – one of the fundamental principles of traditional Chinese medicine. In two of the four RCTs, appearance of the Teh Chi effect was not required in the group of patients treated with acupuncture. Although it has not been proved and no consensus has been reached on this issue, many acupuncturists claim that the appearance of Teh Chi (feeling of tenderness, tingling, cramping, prickliness, warmth, irritation or numbness at the site of needle insertion and its surrounding area), with or without manipulation of the needles, is necessary to achieve optimum therapeutic effects. Only one of the four studies noted whether the acupuncturist was certified or licensed, and none of the studies noted how many years’ experience the practitioners had. Undoubtedly, differences in training and experience may have lowered the quality of acupuncture treatment, and may therefore have had an unfavorable affect on the results.

One of the advantages of acupuncture is the low incidence of side effects in comparison to other medical interventions. Three of the four studies in the present review reported on side effects: two studies found no side effects, whereas the third study noted that the side effects were negligible. The results of the present systematic review show that there is moderate evidence for the effectiveness of acupuncture when used in addition to conventional treatment for pain from osteoarthritis of the knee. This result may be explainable by the marked placebo effect of acupuncture. None of the RCTs included in the present review was designed in a manner that would provide clear-cut evidence of the best evidence, strong evidence could be provided only by high-quality RCTs, a factor that rules out, to a large extent, this possible source of bias. Although the levels of evidence we established were arbitrary, we do not believe that a different rating system would have led to different findings.

The decision to use best evidence analysis rather than meta-analysis was based on the heterogeneity of the RCTs in clinical characteristics: the studies differed widely in duration of the disease, type of acupuncture, duration of follow-up and outcome measures. Because of the lack of consensus on the best way to evaluate the strength of evidence, the levels of evidence we set here were arbitrary. Although we feel it is unlikely, a different rating system of the levels of evidence might have produced different results.

The insertion points used, frequency of treatment and duration of treatment were described in all RCTs reviewed here; however, we found that treatment protocols differed markedly. None of the studies described adaptation of the treatment to a particular individual’s symptoms – one of the fundamental principles of traditional Chinese medicine. In two of the four RCTs, appearance of the Teh Chi effect was not required in the group of patients treated with acupuncture. Although it has not been proved and no consensus has been reached on this issue, many acupuncturists claim that the appearance of Teh Chi (feeling of tenderness, tingling, cramping, prickliness, warmth, irritation or numbness at the site of needle insertion and its surrounding area), with or without manipulation of the needles, is necessary to achieve optimum therapeutic effects. Only one of the four studies noted whether the acupuncturist was certified or licensed, and none of the studies noted how many years’ experience the practitioners had. Undoubtedly, differences in training and experience may have lowered the quality of acupuncture treatment, and may therefore have had an unfavorable affect on the results.
against the potent placebo effect of acupuncture in the control group. Moreover, most of the studies had serious methodological shortcomings. Future RCTs should be designed in such a way as to: a) compare actual acupuncture and sham acupuncture or some other technique that simulates acupuncture; b) control more effectively for the *Teh Chi* effect; c) use appropriate sample sizes; d) use appropriately blinded random allocation; e) evaluate withdrawals and drop-outs appropriately with intention-to-treat analysis; f) use valid, individualized treatments, and g) include short-term and long-term follow-up.

References

For many of our patients, especially those with chronic illnesses, conventional treatments do not work satisfactorily, cause side effects, and do not always alleviate their suffering. Unfortunately osteoarthritis, like many other medical problems, remains chronic and incurable disease. It is well known that many patients with rheumatic disease—regardless of their level of education, background or socioeconomic level—have at some time tried alternative treatments. Because they feel defenseless against a slowly progressive and incapacitating disease, they seek treatments that are «natural» and magical. And because of the steady increase in the use of complementary and alternative medicine, there has also been increasing interest, on the part of physicians who treat rheumatic diseases, in learning about complementary therapies.1 The American College of Rheumatology (ACR) has recognized this current interest, and understands the motives and frustrations of patients who express interest in these approaches to treatment. To ignore these therapies would be a sign of disregard—an imprudent position for a professional organization. The ACR supports the integration of all treatments of proven efficacy and safety, and recommends that such treatments be reviewed scientifically.

Of the variety of therapies available, acupuncture is the most widely used technique.2 Within the current context that favors evidence-based medicine, we cannot ignore acupuncture. Accordingly, Ferrández et al. have undertaken a systematic review of the existing evidence for the effectiveness of acupuncture in the treatment of pain from osteoarthritis of the knee. They concluded that few well-designed studies have looked at the effectiveness of acupuncture for this problem, and that there is not enough evidence to recommend this therapy. An earlier systematic review reached the same conclusion.3

According to the National Consensus Conference on Acupuncture of the National Institutes of Health (NIH)4 in the USA, acupuncture «is a component of a health care system that has been used in China for at least 2000 years. The general theory of acupuncture is based on the premise that there are energy flows that move through the body and are essential for health». Western acupuncture is currently very different from that used in traditional Chinese medicine. Several studies have suggested that the analgesia produced by acupuncture is triggered by stimulation of small-caliber, high threshold nerves in the muscle tissue, from whence impulses are sent to the spinal cord and specific neurological centers, which release neurotransmitters that in turn block pain signals.5 Physicians who treat patients with chronic illnesses such as osteoarthritis of the knee should be aware that their patients are potential candidates for acupuncture treatment for their pain. They should therefore be prepared to inform patients about published data, and about the possible risks and benefits. They should also warn patients that acupuncture can be painful, and may be contraindicated in some cases. Acupuncture is an invasive medical procedure, and complications from inappropriate treatment have been reported, hence the importance of adequate training standards as well as full knowledge of anatomy and appropriate hygienic measures.5

There is no scientific evidence from large-scale randomized clinical trials that acupuncture is effective in rheumatic processes. Occasionally, the sample size is too small to allow the results to reach statistical significance. The role of acupuncture in the treatment of musculoskeletal pain should be more carefully evaluated, although studies that compare real acupuncture with sham acupuncture appear to be difficult to perform because of the possible placebo
effect, problems with interventions in the control group, and double blinding to avoid biasing the results. White et al. have prepared a consensus document for improving the design of studies that investigate the efficacy of acupuncture, and offer a number of recommendations to improve the quality of these studies. There appear to exist many treatment variables (such as point selection and form of stimulation) that should be controlled for when clinical trials are designed. For efficacy studies that aim to determine whether acupuncture is superior to a placebo, the authors of the review in this issue of ATENCIÓN PRIMARIA recommend sham procedures as the best method of control, as false negative results may be obtained if control procedures are inadequate.

Randomized, controlled studies of adequate methodological quality are important to evaluate acupuncture as a treatment alternative, as all treatments should be based on tests of validity with the same scientific methods. The only type of medicine is that which is based on science and evidence, and it is evidence-based medicine which we should aspire to practice.

References