Acupuncture for knee osteoarthritis – a randomised trial using a novel sham

Eric Manheimer*, Byungmook Lim*, Lixing Lao, Brian Berman
*equal contributors

This trial was originally published as: Berman B et al. Effectiveness of acupuncture as an adjunctive therapy in osteoarthritis of the knee. *Ann Intern Med* 2004;141:901-910. It is reproduced here with permission of the publishers.

**Abstract**

**Background** Evidence on the efficacy of acupuncture for reducing the pain and dysfunction of osteoarthritis is equivocal.

**Objective** To determine whether acupuncture provides greater pain relief and improved function compared with sham acupuncture or education in patients with osteoarthritis of the knee.

**Design** Randomised, controlled trial.

**Setting** Two outpatient clinics (an integrative medicine facility and a rheumatology facility) located in academic teaching hospitals and one clinical trials facility.

**Patients** 570 patients with osteoarthritis of the knee (mean age [±SD], 65.5 ± 8.4 years).

**Intervention** 23 true acupuncture sessions over 26 weeks. Controls received 6 two-hour sessions over 12 weeks or 23 sham acupuncture sessions over 26 weeks.

**Measurements** Primary outcomes were changes in the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and function scores at 8 and 26 weeks. Secondary outcomes were patient global assessment, 6-minute walk distance, and physical health scores of the 36-Item Short-Form Health Survey (SF-36).

**Results** Participants in the true acupuncture group experienced greater improvement in WOMAC function scores than the sham acupuncture group at 8 weeks (mean difference, -2.9 [95% CI, -5.0 to -0.8]; P=0.01) but not in WOMAC pain score (mean difference, -0.5 [CI, -1.2 to 0.2]; P=0.18) or the patient global assessment (mean difference, 0.16 [CI, -0.02 to 0.34]; P>0.2). At 26 weeks, the true acupuncture group experienced significantly greater improvement than the sham group in the WOMAC function score (mean difference, -2.5 [CI, -4.7 to -0.4]; P=0.01), WOMAC pain score (mean difference, -0.87 [CI, -1.58 to -0.16]; P=0.003), and patient global assessment (mean difference, 0.26 [CI, 0.07 to 0.45]; P=0.02).

**Limitations** At 26 weeks, 43% of the participants in the education group and 25% in each of the true and sham acupuncture groups were not available for analysis.

**Conclusions** Acupuncture seems to provide improvement in function and pain relief as an adjunctive therapy for osteoarthritis of the knee when compared with credible sham acupuncture and education control groups.

**Keywords**

Acupuncture, randomised controlled trial, osteoarthritis, placebos.
Clinical trials

Introduction

Osteoarthritis is the most common form of arthritis and is a major cause of morbidity, limitation of activity and health care use, especially among the elderly. Pain and functional limitation are the primary clinical manifestations of osteoarthritis of the knee. Current recommendations for managing osteoarthritis, including guidelines published by the American College of Rheumatology and European League of Associations of Rheumatology, focus on relieving pain and stiffness and maintaining or improving physical function as important goals of therapy. No curative therapies exist for osteoarthritis; thus, both pharmacologic and non-pharmacologic management focus on controlling pain and reducing functional limitation. However, pharmacologic management of osteoarthritis is often ineffective, and agents such as NSAIDs may cause unwanted and dangerous side effects.

Complementary and alternative medicine is another approach to treating osteoarthritis, particularly in Asian societies. Many US patients with osteoarthritis also use complementary and alternative medical therapies. A systematic review of acupuncture and knee osteoarthritis identified seven small randomised, controlled trials published in English. Within the methodological limitations of the studies, the evidence suggested that acupuncture seemed to alleviate knee pain and function compared with 'sham' acupuncture controls.

Before conducting our large scale trial, we completed both an uncontrolled pilot trial and a randomised, single blind, waiting list controlled trial of the effect of acupuncture on osteoarthritis of the knee. Both trials showed statistically significant improvements in self reported pain and physical function, compared with baseline (for the uncontrolled trial) and compared with the control group (for the waiting list controlled trial). However, for the controlled trial, which was of a longer duration, the effects diminished within 18 weeks of the final acupuncture treatment.

Methods

Patient recruitment

Patients were recruited, primarily through print and radio advertisements, from March 2000 through to December 2003, at three different sites: one university Integrative Medicine clinic, one private research firm, and one hospital clinic.

We determined the sample size (n=570) by a power analysis based on our randomised pilot trial, adjusted by the estimated decrease in effect size resulting from the inclusion of a sham acupuncture group designed to control for placebo effects.

Patients met the following inclusion criteria: age 50 years or older, a diagnosis of osteoarthritis of the knee, radiographic evidence of at least one osteophyte at the tibiofemoral joint (Kellgren-Lawrence grade ≥2), moderate or greater clinically significant knee pain on most days during the past month, and willingness to be randomly assigned. Exclusion criteria were the presence of serious medical conditions that precluded participation in study, bleeding disorders that might contraindicate acupuncture, intra-articular corticosteroid or hyaluronate injections (as well as any knee surgery or concomitant use of topical capsaicin cream) during the past six months, previous experience with acupuncture, or any planned events (including total knee replacement) that would interfere with participation in the study during the following 26 weeks.

After a brief telephone screening, patients were scheduled to visit one of the three participating sites to sign an informed consent statement and undergo a brief rheumatologic examination (including radiographic examination of affected knees) by a physician or a nurse practitioner. Patients were randomly assigned to one of three groups by a computer generated process using randomly selected blocks of three, six, and nine. We assured allocation concealment by using codes generated by a central statistical team and sent to the site coordinators in numbered, sealed, opaque envelopes. The research assistants who collected assessments from participants, the participants in the acupuncture groups, and the statistician were blinded to group assignment. Assessments were conducted at baseline and 4, 8, 14, and 26 weeks after randomisation.
Study interventions

We developed and modified the acupuncture treatment and sham control protocols from previously reported and validated procedures.\textsuperscript{15-18} During the trial, seven acupuncturists were used. In general, acupuncturists were assigned to the same participants throughout the 26 week treatment schedule. All acupuncturists were state licensed and had at least two years of clinical experience. The study’s principal acupuncturist trained and supervised the acupuncturists in performing true or sham procedures and avoiding interactions that could inadvertently communicate group assignment.

True acupuncture

The true acupuncture (experimental) group underwent 26 weeks of gradually tapering treatment according to the following schedule: eight weeks of two treatments per week followed by two weeks of one treatment per week, four weeks of one treatment every other week, and 12 weeks of one treatment per month. We based the acupuncture point selections on Traditional Chinese Medicine meridian theory to treat knee joint pain. These points consisted of five local points (GB34, SP9, ST36, ST35, Extra point Xiyan), and four distal points (BL60, GB39, SP6, KI3) on meridians that traverse the knee area.\textsuperscript{19;20} The same points were treated for each affected leg. If both knees were affected, nine needles were inserted in each leg. The acupuncturists inserted 32 gauge (0.25mm diameter) acupuncture needles to a depth of 0.3 to 1.0 inch (10 to 25mm), depending on point location. All participants in the treatment group achieved the \textit{de qi} sensation, a local sensation of heaviness, numbness, soreness, or paraesthesia that accompanies the insertion and manipulation of needles, at these nine points. Acupuncturists applied electrical stimulation at knee points Xiyan for 20 minutes.\textsuperscript{11}

To ensure that the procedures in the treatment and control groups were as similar as possible, we tapped two guiding tubes at two points (not classical acupuncture points) in the abdominal area, approximately 30mm lateral to and slightly above the umbilicus bilaterally, and immediately affixed a pair of needles to the surface of the same points, without needle insertion, with adhesive tape.

Sham control

For the sham treatment, we modified a combined insertion and no insertion procedure from our previously validated placebo acupuncture method.\textsuperscript{17;18} Acupuncturists inserted two needles into the points in the abdominal area, which were identical to points used in the true acupuncture group, and then immediately applied two pieces of adhesive tape next to the needles. In addition, they tapped a plastic needle guiding tube on nine true points in the leg to produce some discernible sensation and then immediately affixed a needle with a piece of adhesive tape, without needle insertion, for a total of 20 minutes.\textsuperscript{18} The sham acupuncture procedure was given on the same schedule as the experimental group and used the same active needle placements, except actual insertion did not occur at these nine points. A mock transcutaneous electrical stimulation unit (which emitted a sound and possessed a blinking light) was attached to the sham needles at the knee. To facilitate blinding, we used screens in both treatment and sham groups that were placed below the abdomen to prevent participants from actually observing the true or sham procedures at the knee area but to allow them to observe the procedure being performed in the abdomen area.

Education control

The education attention control consisted of six two hour group teaching sessions based on the Arthritis Self Management Program.\textsuperscript{21} In addition, we periodically mailed educational materials to the education group in an attempt to equalise the amount of experimental contact in all groups.

Outcome measures

The primary outcome variables for the study, specified \textit{a priori}, were the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain and function scores.\textsuperscript{22} Secondary outcomes were the 36-Item Short-Form Health Survey (SF-36) physical component score,\textsuperscript{23} the patient global assessment,\textsuperscript{24} and the six minute walk time. In addition to baseline, we assessed participants at 4, 8, 14, and 26 weeks on the two WOMAC scales, patient global assessment measures and patients’ self reports of adverse events, and at 8 and 26 weeks on the SF-36 and six minute walk outcomes.
Statistical analysis

Initial analyses tabulated demographic and baseline characteristics of the study participants by randomisation group, and we used chi square tests and one way analyses of variance to compare the three groups on these characteristics.

Longitudinal analyses examined mean change from baseline at 4, 8, 14, and 26 weeks by using a mixed model approach as implemented by the MIXED procedure in SAS, version 8.2 (SAS Institute, Inc Cary, North Carolina). Clinical site was controlled for by its inclusion as a random effect, although the site effect was small. We included the baseline value of the outcome variable as a covariate in all analyses of change from baseline to assess whether change differed by baseline level.

Results

The Figure shows the progress of participants through the trial. The proportion of dropouts between true and sham acupuncture groups did not differ. However, substantially fewer participants were available for assessment at 26 weeks in the education group (43% dropout) than in the true acupuncture group (25% dropout) or the sham control group (26% dropout). Table 1 shows that there were no pre-treatment differences among the three experimental groups.

Table 2 presents the mean changes from baseline for the five study outcomes. A trend for the three groups as a whole was statistically significant (that is, the time main effect) to improve over time on all of the outcomes except the six-minute walk. Of greater interest, however, are the differences in improvement in study outcomes observed between true versus sham acupuncture groups at the different points in time.

Pain

While pain among participants who were receiving true acupuncture decreased more than in the sham group at all of the post baseline assessments, this difference was not statistically significant at week eight. By week 14, the mean WOMAC pain score had decreased by 3.6 units in the acupuncture group (a 40% decrease from baseline) compared with 2.7 units in the sham group (P=0.02). These differences remained at week 26 (P=0.003).

Figure 1 Participant flow chart.
Function
The true acupuncture group’s improvement in function from baseline was significantly greater than that of the sham control group at weeks 8 (P=0.01), 14 (P=0.04), and 26 (P=0.009). A change of more than 12 units by 14 weeks is an almost 40% improvement from baseline.

Secondary outcomes
For most of the secondary outcomes, there was no statistically significant difference between the true and sham acupuncture groups (Table 2).

Masking effectiveness
To evaluate the masking effectiveness of the sham acupuncture procedure, we asked participants in the acupuncture and sham groups to report which treatment they believed they were receiving at both 4 and 26 weeks: ‘true acupuncture’, ‘sham acupuncture’, or ‘uncertain’. Most participants in both groups believed that they were receiving true acupuncture at both times, suggesting that the sham acupuncture procedure was a relatively credible blinding strategy. At four weeks, 67% in the true acupuncture group and 58% in the sham group believed that they were receiving true acupuncture at both times, suggesting that the sham acupuncture procedure was a relatively credible blinding strategy.
believed that they were receiving true acupuncture (P=0.06), and 25% and 33% were unsure, respectively. By the end of the trial, more individuals in the true group (75%) than in the sham group (58%) held this belief (P=0.003), and 23% and 32% were unsure, respectively. As indicated in Table 3, however, participant guesses on which treatment they were receiving were related to changes in WOMAC pain and function scores at both 8 and 26 weeks.

### Table 2: Mean Change from Baseline in Participant Outcomes*

<table>
<thead>
<tr>
<th>Week</th>
<th>Outcomes</th>
<th>True acupuncture</th>
<th>P value‡</th>
<th>Sham acupuncture</th>
<th>P value§</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Participants, n</td>
<td>173</td>
<td>163</td>
<td>124</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WOMAC pain score†</td>
<td>–2.22 ± 0.24</td>
<td>&gt;0.2</td>
<td>–1.98 ± 0.25</td>
<td>&lt;0.001</td>
<td>–0.84 ± 0.26</td>
</tr>
<tr>
<td></td>
<td>WOMAC function score†</td>
<td>–7.56 ± 0.78</td>
<td>0.15</td>
<td>–5.90 ± 0.66</td>
<td>0.05</td>
<td>–4.65 ± 0.81</td>
</tr>
<tr>
<td></td>
<td>Patient Global Assessment</td>
<td>0.13 ± 0.07</td>
<td>&gt;0.2</td>
<td>0.10 ± 0.07</td>
<td>&gt;0.2</td>
<td>0.07 ± 0.09</td>
</tr>
<tr>
<td>8</td>
<td>Participants, n</td>
<td>169</td>
<td>161</td>
<td>125</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WOMAC pain score†</td>
<td>–3.15 ± 0.29</td>
<td>0.18</td>
<td>–2.66 ± 0.26</td>
<td>&lt;0.001</td>
<td>–1.25 ± 0.30</td>
</tr>
<tr>
<td></td>
<td>WOMAC function score†</td>
<td>–10.77 ± 0.90</td>
<td>0.01</td>
<td>–7.84 ± 0.76</td>
<td>&lt;0.001</td>
<td>–5.30 ± 0.95</td>
</tr>
<tr>
<td></td>
<td>Patient Global Assessment</td>
<td>0.30 ± 0.07</td>
<td>&gt;0.2</td>
<td>0.14 ± 0.08</td>
<td>0.09</td>
<td>0.04 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>SF-36 Physical Health</td>
<td>9.2 ± 1.4</td>
<td>&gt;0.2</td>
<td>7.6 ± 1.2</td>
<td>0.02</td>
<td>4.3 ± 1.3</td>
</tr>
<tr>
<td></td>
<td>6-Minute Walk, ft</td>
<td>64.1 ± 18.0</td>
<td>&gt;0.2</td>
<td>67.7 ± 18.6</td>
<td>0.02</td>
<td>–1.0 ± 30.8</td>
</tr>
<tr>
<td>14</td>
<td>Participants, n</td>
<td>158</td>
<td>157</td>
<td>113</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WOMAC pain score†</td>
<td>–3.63 ± 0.31</td>
<td>&lt;0.02</td>
<td>–2.68 ± 0.33</td>
<td>0.001</td>
<td>–1.54 ± 0.35</td>
</tr>
<tr>
<td></td>
<td>WOMAC function score†</td>
<td>–12.18 ± 0.96</td>
<td>0.04</td>
<td>–9.40 ± 0.94</td>
<td>&lt;0.001</td>
<td>–5.62 ± 1.05</td>
</tr>
<tr>
<td></td>
<td>Patient Global Assessment</td>
<td>0.36 ± 0.08</td>
<td>&gt;0.2</td>
<td>0.26 ± 0.08</td>
<td>0.03</td>
<td>0.15 ± 0.09</td>
</tr>
<tr>
<td>26</td>
<td>Participants, n</td>
<td>142</td>
<td>141</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WOMAC pain score†</td>
<td>–3.79 ± 0.33</td>
<td>&lt;0.01</td>
<td>–2.92 ± 0.30</td>
<td>&lt;0.01</td>
<td>–1.69 ± 0.33</td>
</tr>
<tr>
<td></td>
<td>WOMAC function score†</td>
<td>–12.42 ± 1.12</td>
<td>&lt;0.01</td>
<td>–9.88 ± 0.93</td>
<td>0.01</td>
<td>–7.17 ± 1.07</td>
</tr>
<tr>
<td></td>
<td>Patient Global Assessment</td>
<td>0.45 ± 0.08</td>
<td>0.02</td>
<td>0.19 ± 0.09</td>
<td>&gt;0.2</td>
<td>0.22 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>SF-36 Physical Health</td>
<td>10.7 ± 1.6</td>
<td>0.21</td>
<td>8.2 ± 1.5</td>
<td>0.01</td>
<td>4.0 ± 1.5</td>
</tr>
<tr>
<td></td>
<td>6-Minute Walk, ft</td>
<td>74.2 ± 20.2</td>
<td>&gt;0.2</td>
<td>105.0 ± 21.4</td>
<td>&lt;0.01</td>
<td>–3.6 ± 40.8</td>
</tr>
</tbody>
</table>

*Test results and P values from mixed-model analysis of change from baseline, Toeplitz (banded) covariance matrix to control for within-participant correlation, clinical site as a random effect. Models include baseline value of the outcome variable, highly statistically significant in all cases. Values presented with a plus/minus sign are means ± SE. SF-36 = 36-Item Short-Form Health Survey; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

†Pain and function were the primary trial end points

‡P values compare true acupuncture with sham acupuncture groups

§P values compare sham acupuncture with education groups

|| Numbers of participants assessed for the SF-36 Physical Health and 6-Minute Walk outcomes measures were slightly different from the numbers assessed for the WOMAC outcome measures (see original publication, Table 2)

### Table 3: Relationship between Group Guesses at 26 Weeks and Western Ontario and McMaster Universities Osteoarthritis Index Pain and Function End Points*

<table>
<thead>
<tr>
<th>End point</th>
<th>Uncertain (n=49)</th>
<th>Sham acupuncture (n=45)</th>
<th>True acupuncture (n=187)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 wk</td>
<td>–2.8 (-3.9 to -1.8)</td>
<td>–1.0 (-1.8 to -0.2)</td>
<td>–3.8 (-4.2 to -3.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>26 wk</td>
<td>–2.9 (-4.0 to -1.9)</td>
<td>–1.2 (-2.0 to -0.3)</td>
<td>–4.0 (-4.6 to -3.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 wk</td>
<td>–7.6 (-10.7 to -4.5)</td>
<td>–3.1 (-5.6 to -0.7)</td>
<td>–11.8 (-13.3 to -10.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>26 wk</td>
<td>–8.9 (-12.3 to -5.5)</td>
<td>–4.7 (-7.1 to -2.3)</td>
<td>–13.4 (-15.2 to -11.6)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Values in parentheses are 95% CIs
Safety
Twenty six adverse events were reported for the 570 participants: 14 (7%) in the true acupuncture group, 5 (3%) in the sham control, and 7 (4%) in the education control. Of the 14 adverse events observed in the true acupuncture group, none was interpreted as treatment related, and the differences among groups did not reach statistical significance.

Concurrent treatments
Participants continued to receive medical care during the study from their primary care physicians and were allowed to receive their usual medications. We tracked these medications usages at each assessment after baseline, and we found no statistically significant differences between the proportions of participants in the true and sham groups.

Discussion
The results of our study extend those of our previous trial and demonstrate that true traditional Chinese acupuncture is safe and effective for reducing pain and improving physical function in patients with symptomatic knee osteoarthritis who have moderate or greater pain despite background therapy with analgesic or anti-inflammatory therapy.

We used a credible sham acupuncture group to control for the potential placebo effect in our trial. While the participants in the true acupuncture group were more likely to correctly guess their treatment, this masking procedure was reasonably successful in blinding participants in the sham control group since most participants believed that they were receiving true acupuncture (or were unsure) throughout the study. However, participants’ differential awareness of group membership may have contributed to the positive results found. Yet, these between group masking differences may have instead reflected the differential pain and function improvements due to the treatments themselves.

In any event, our trial ensured reasonable concealment of group allocation, provided evidence for the utility of the Center for Integrative Medicine sham procedure for use in acupuncture trials; contained adequate power; and adhered to the OARSI’s recommendation that symptom relief be assessed for 24 weeks or more. We interpret the superiority of true compared with sham acupuncture in improvements in pain and function as evidence of the treatment’s efficacy, especially given the masking success achieved for the sham procedure.

Although considerable attrition occurred over the six month study for all groups, this problem was not differential for the true and sham groups. We interpret this finding, coupled with the results of the analytic steps we took to study the differences between participants who dropped out of the trial versus those who completed it, as indicative that attrition did not confound the observed true versus sham differences. There is, however, no way to be absolutely certain that this problem did not affect our 26 week results.

Our study is one of the largest randomised, placebo controlled acupuncture trials to date and it involved a more intensive acupuncture regimen (23 sessions) for a longer period (26 weeks) than any other trial. While the duration of our acupuncture treatment may seem long, we used only eight weeks of focused treatment, followed by a tapered schedule for maintenance purposes. This is not an uncommon practice in China.

If these results are valid, our trial has two important clinical implications. First, the absence of any observed treatment related side effects attributable to acupuncture needling contrasts to current pharmacologic therapies for osteoarthritis that have side effects that may rival in severity the arthritis symptoms themselves. Second, observed acupuncture effects were achieved in addition to those of other viable osteoarthritis treatments, such as non-pharmacological therapies and NSAIDs or COX-2 selective inhibitors, since study participants were free to pursue any therapy they or their physician desired. Thus, acupuncture may have an important role in adjunctive therapy as part of a multidisciplinary integrative approach to treating symptoms related to knee osteoarthritis.

Acknowledgements
Funding
For preparing this monograph, EM and BB were funded by Grant Number R24 AT001293 from the National Center for Complementary and Alternative Medicine (NCCAM) of the US National Institutes of Health, and LL AND BB were funded by NCCAM Grant Numbers P01 AT002605-01A1 and U01 AT00171. BL was funded by Grant Number KRF-2005-214-E00056 from the Korea Research
Clinical trials

Foundation. The funding sources had no role in the writing of this monograph or in the decision to submit the paper for publication. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NCCAM or the Korea Research Foundation.

Reference list


Acupuncture for knee osteoarthritis – a randomised trial using a novel sham

Eric Manheimer, Byungmook Lim, Lixing Lao and Brian Berman

*Acupunct Med* 2006 24: 7-14
doi: 10.1136/aim.24.Suppl.7

Updated information and services can be found at:
http://aim.bmj.com/content/24/supplement/7

These include:

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://www.bmj.com/company/products-services/rights-and-licensing/

To order reprints go to:
http://journals.bmj.com/content/subscribers

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/