Acupuncture is Effective for Chronic Knee Pain: A Reanalysis of the Australian Acupuncture Trial

Arthur Yin Fan, MD, PhD, LAc; Kehua Zhou, MD, DPT, PT, LAc; Sherman Gu, BMed, AdvDipMyother, MAppSc; Yong Ming Li, MD, PhD

ABSTRACT

Context • In the October 2014 issue of the Journal of the American Medical Association (JAMA), Hinman et al published the results of an Australian clinical trial on acupuncture in a paper entitled “Acupuncture for Chronic Knee Pain: A Randomized Clinical Trial” (JAMA report), in which they concluded that neither acupuncture nor laser acupuncture had any greater effects than sham laser acupuncture for pain or function for patients aged 50 y and older with moderate-to-severe knee pain. That study has been criticized extensively by international scholars for its validity because serious methodological flaws existed throughout the study’s design, implementation, and conclusions.

Objective • The current study intended to re-examine the prior study’s conclusions about the efficacy of acupuncture for chronic knee pain.

Design • The current research team performed a reanalysis of relevant data from the JAMA report.

Intervention • The original study included 4 groups: (1) an acupuncture group, which received needle acupuncture, inferred by the current authors to have been set up to be a positive control in the original study; (2) a laser acupuncture group, which received laser acupuncture; (3) a sham laser acupuncture group, which received sham laser acupuncture and acted as the negative controls for the laser acupuncture intervention; and (4) a control group, which received conventional care but no acupuncture or laser treatments. The study lasted 12 wk.

Outcome Measures • The measures included evaluations in the following areas: (1) poststudy modifications—an evaluation of the consistency of the JAMA report with the study’s intentions as identified for a grant that was originally approved and funded by the Australian National Health and Medical Research Council (NHMRC) in 2009, as indicated in the study’s trial registration, and as compared with the published protocols and to the study’s originally stated objectives; (2) high heterogeneity—an assessment of the heterogeneity among the 4 groups for the overall outcome related to pain; (3) ineffectiveness of laser acupuncture—an analysis of laser acupuncture’s efficacy for chronic knee pain as stated in the JAMA report, using effect size (ES); (4) effectiveness of acupuncture—a reanalysis of acupuncture’s efficacy for chronic knee pain in comparison with the original analysis in the JAMA report, using ES; and (5) acupuncture after data adjustment—a new analysis of acupuncture’s efficacy for chronic knee pain using data from the original study that was discussed in the JAMA report, using ES, with an estimation after data adjustment and elimination of the dilution effect of the Zelen design.

Results • Contrary to a general impression that acupuncture was the focus, laser acupuncture was the primary intervention tested in the actual study, “Laser Acupuncture in Patients With Chronic Knee Pain: A Randomized, Placebo Controlled Trial.” The study discussed in the JAMA report was neither a truly randomized, controlled trial (RCT) for acupuncture nor was it an appropriately designed, randomized study in general. High heterogeneity was found among its groups in the evaluation of overall pain in patients. Both the ES of 0.60 that had been set by Hinman et al for the minimal clinically important difference (MCID) and the resulting interpretation of results in the JAMA report were not appropriate. Using the original study’s criteria of efficacy, the reanalysis has confirmed that the laser acupuncture was not effective, whereas the acupuncture was found to be moderately effective for chronic knee pain (P < .05) for both overall pain and function at 12 wk, with an ES of 0.58, or after the adjustment of the data, with an ES of 0.67.

Conclusions • The JAMA study was neither a conventional RCT nor an appropriately randomized trial, and its results are probably invalid. The ES of 0.60 for the MCID that was used in the JAMA study and the resulting explanation were not appropriate. Even with an ES of 0.60 for the MCID, acupuncture remained effective after data adjustment. Consequently, compared with conventional care, acupuncture treatment was found to be moderately effective for chronic knee pain in patients aged 50 y and older. (Altern Ther Health Med. 2016;22(3):32-36.)
Gu, BMed, AdvDipMyother, MAppSc, is a clinical assistant professor in the Department of Health Care Studies and the lead physical therapist at the Wound Care Clinic at Daemen College in Amherst, New York. Sherman Gu, BMed, AdvDipMyother, MAppSc, is a registered Chinese medicine practitioner, and officer at the Federation of Chinese Medicine & Acupuncture Societies of Australia Ltd in Wantirna South, Victoria, Australia. Yong Ming Li, MD, PhD, is former President at the American Traditional Chinese Medicine Society in New York, New York.

Corresponding author: Arthur Yin Fan, MD, PhD, LAc  
E-mail address: arthurfan@chinesemedicinedoctor.us

In the October 2014 issue of the Journal of the American Medical Association (JAMA), Hinman et al published the results of an Australian clinical trial on acupuncture in a paper entitled “Acupuncture for Chronic Knee Pain: A Randomized Clinical Trial,” in which the researchers concluded that neither acupuncture nor laser acupuncture had any greater effects than sham laser acupuncture on pain or function for patients aged 50 years and older with moderate-to-severe knee pain. That study has been extensively criticized by international scholars for its validity because serious methodological flaws existed throughout the study’s design, implementation, and conclusions.2-12

Further, Hinman et al’s alteration of the original study’s objectives and protocols after completion of the trial and the fact that they failed to mention the changes in their final report has raised concerns about the integrity of the investigators.2-4,10-12 Those concerns may justify a finding that the results and conclusions reported by Hinman et al1 are probably not valid or appropriate.

Using relevant, available, supplementary data, together with the final published report of the Australian clinical trial on acupuncture in JAMA—referred to as the JAMA report—in the rest of the current article—the authors have reanalyzed the effects of acupuncture for chronic knee pain as evaluated in the original study.

POSTSTUDY MODIFICATIONS

Evidence from the trial report in JAMA1 and from information gathered through clinical-trial registration13 indicates that laser acupuncture rather than acupuncture was the primary intervention tested. Laser-acupuncture and acupuncture differ in many ways.2-4,9-12 Based on the traditional and official definition, in which the term acupuncture refers to needling therapy, an actual insertion of a solid needle into the body;14 the categorization of noninvasive, laser acupuncture as acupuncture is debatable. Thus, the mixed usage of the term acupuncture in referring to both laser and needle therapy by Hinman et al1 was unprofessional and misleading.2-9-12

Therefore, the current article uses the word acupuncture to refer strictly to needling therapy; to be as clear as possible. As per the JAMA report, 4 arms were designed for the study:1 (1) an acupuncture group, which received actual needle treatments; (2) a laser acupuncture group, which received laser acupuncture; (3) a sham laser acupuncture group, which received sham laser acupuncture and acted as the negative controls for the laser acupuncture intervention; and (4) a control group, which received conventional care but no acupuncture or laser treatments. Because the primary aim of the trial was to test laser acupuncture only,13,15 the reasonable explanation for including acupuncture is that it was used as a positive control.2,4,10-12

In addition to the study’s published protocol,15 which stated that laser acupuncture was the primary intervention tested, a synopsis of the original research proposal that had been posted by Monash University and the funding application for the research to the Australian National Health and Medical Research Council (NHMRC) in 2009, also indicated the same objective.16,17

Even more clearly stating that objective, the title of the research proposal in both cases was “Laser Acupuncture in Patients with Chronic Knee Pain: A Randomized, Placebo Controlled Trial,” and it was funded by NHMRC in 2009.16,17

Acupuncture had been reported to be an effective therapy for knee osteoarthritis in a much larger trial in 2004, and the researchers for the study on laser acupuncture had cited that reference in their research proposal. They also stated clearly in their original proposal that one of their aims was to compare laser acupuncture with acupuncture but did not claim an intention to compare the acupuncture group with the control group.6,17 The power of the statistical test and other resources in the original study were optimized only to evaluate the laser intervention, instead of 2 interventions, laser and acupuncture, as was indicated in the JAMA report.4,7-13,15

Because the JAMA report mixed the terms laser acupuncture and acupuncture (Hinman et al1 named it as needle acupuncture) with the general term acupuncture, the authors blurred the focus of the original study’s design. However, the title and contents of the published JAMA paper actually reported that acupuncture using needling rather than laser acupuncture was the primary intervention tested. In addition, the title of the JAMA report stated that the study was a randomized, clinical trial rather than a randomized, controlled trial (RCT). That change in title suggests that the investigators were aware that their post hoc modifications led to a nontypical RCT. A serious deficiency of the trial after the modification of objectives was a lack of a sham control for acupuncture.2-4,9-12 Modification of the research objectives after completing the study, without a reasonable explanation and with a failure to mention the changes in the final report, was unethical and a violation of professional standards.1-4,11-12

Fan—Acupuncture and Chronic Knee Pain
HIGH HETEROGENEITY

Apparently, the JAMA report was not a conventional RCT report because a Zelen design was used. The high heterogeneity of the 4 groups in the study further decreased its reliability and validity as a randomized trial. Using Review Manager 5.3 (a software downloaded from http://www.tech.cochrane.org), the current research team pooled the data for the overall pain score, one of the primary outcome measures discussed in the JAMA report for week 12, to test the overall effects of acupuncture versus the control and laser acupuncture versus sham laser acupuncture. The team also evaluated the potential heterogeneity among the findings, following the method recommended by the Cochrane Collaboration.18

The measurements indicated that the overall effects were favorable to the groups who were treated with acupuncture or laser treatment although the effects were marginal—mean difference (MD), -0.56; 95% confidence interval (CI), -1.12 to -0.00; P = .05; n = 256 (Figure 1).

A statistically significant difference existed between the acupuncture group and the control group—MD, -1.10; 95% CI, -1.88 to -0.32; n = 133. However, a substantial heterogeneity (I² = 73%) among the 4 groups significantly weakened the strength of that positive estimate. The sources of that heterogeneity might have included (1) selection bias, because the acupuncture group had the worst scores at baseline in all of the primary outcome measures and in almost all of the secondary outcome measures; (2) the integrity of the intervention (ie, acupuncture and laser acupuncture are not comparable as the former involves skin penetration with solid needles placed at multiple acupuncture points during a single visit, whereas the latter uses a laser device with a single output point that was used to stimulate selected points during the treatment); (3) performance bias (ie, systematic differences among the groups in blinding) as detailed in number 4; and (4) an inconsistency that may have come from an attrition bias and a data dilution bias that was associated with the Zelen design24 (ie, systematic differences among the groups in withdrawals [dropouts]). In the JAMA report, at week 12 of the trial, the dropout rate was (1) acupuncture group—8.57% (ie, [70-64]/70); (2) laser acupuncture group—8.45% (ie, [71-65]/71); (3) sham laser acupuncture group—17.14% (ie, [70-58]/70); and (4) control group—2.82% (ie, [71-69]/71).1

INEFFECTIVENESS OF LASER ACUPUNCTURE

As per the JAMA report by Hinman et al,1 compared with the control at week 12, laser acupuncture did not induce significant improvements in function as assessed by the Western Ontario and McMaster Universities Arthritis Index (WOMAC) and by 7 other secondary outcomes measures, with the exception of a slight improvement in the overall pain score, with a difference of -0.8 (P = .03). Using the calculation of effect size (ES) for mean differences between groups with unequal sample sizes within a pre-post design,19 the ES for pain relief was calculated as being 0.33 for the laser acupuncture group when compared with the control group. That ES was much smaller than the ES of 0.60 for the minimal clinically important difference (MCID) that had been set by the investigators in the original design of the study.1,5 Compared with the ES for the sham laser-acupuncture group, the ES19 for the laser acupuncture group was -0.05.

Thus, as stated in the JAMA report by Hinman et al,1 laser acupuncture was not effective for chronic knee pain. Consequently, negative results were found for the clinical trial, which had originally been designed for the assessment of laser acupuncture in patients with chronic knee pain.

EFFECTIVENESS OF ACUPUNCTURE

Based on the NHMRC-approved proposal, the published RCT protocol, and the information from the trial registration,13,15-17 the acupuncture group was included as a positive control; thus, comparing acupuncture with sham laser acupuncture, a negative control, was not justifiable,3,4,9-12 nor was it a specific aim in any of the proposals. Moreover, the study’s blinding protocol covered only the laser acupuncture and sham laser acupuncture groups. No blinding was performed for treatments given to acupuncture group in comparison to the sham laser acupuncture group, and those 2 interventions do not have comparability in any setting.3,4,9-12 Consequently, the assessment of the effectiveness of acupuncture should have been based only on a comparison with the group receiving conventional care, which was the true control group in the study.

In the current research team’s comparison of the results for the acupuncture group with those of the control group at

**Figure 1.** Pooled Data from Hinman 2014 for the Outcome Measure Overall Pain Score at Week 12

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Active intervention</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
<th>Mean Difference IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hinman 2014a</td>
<td>3.3</td>
<td>2.2</td>
<td>64</td>
<td>4.4</td>
<td>2.4</td>
<td>69</td>
<td>51.0%</td>
<td>-1.10 (-1.88 to 0.32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hinman 2014b</td>
<td>3.4</td>
<td>2.2</td>
<td>65</td>
<td>3.4</td>
<td>2.3</td>
<td>58</td>
<td>49.0%</td>
<td>0.00 (-0.80 to 0.80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>129</td>
<td>127</td>
<td>100.0%</td>
<td>-0.56 (-1.12 to 0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: SD, standard deviation; CI, confidence interval.
In the acupuncture group, only 54 patients received acupuncture treatments. Consequently, a 31.27% dilution occurred in the efficacy assessment for the acupuncture group as a result of the Zelen design. Theoretically, any loss in statistical efficiency can be overcome by increasing sample size; however, because the JAMA report gave a post hoc analysis, that increase was not possible. Considering the small sample size of the study, the current research team decided that an adjustment was needed in the data to compensate for the dilution in the estimation of acupuncture's efficacy for chronic knee pain. Therefore, the team allocated the 10 patients in the acupuncture group who did not receive any acupuncture treatment to the control group that was discussed in the JAMA report.

As Dr Hinman had declined the current research team’s request to provide the original raw data for further analysis, the team has been unable to reconduct the individual, patient-based data analyses to compare the differences among the groups. Nonetheless, the team used a student *t* test to evaluate the efficacy of the acupuncture in comparison with the control group. The data on overall pain for the acupuncture and control groups as shown in the JAMA report were recalculated after the earlier-mentioned data adjustment, shown in Table 2.

### Table 1. Comparisons of the Pain Scores at Week 12 Using Original Data from Hinman et al

<table>
<thead>
<tr>
<th>Group</th>
<th>Baseline M ± SD, n</th>
<th>Week 12 M ± SD, n</th>
<th>Improve (%)</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>5.1 ± 2.1, 71</td>
<td>4.4 ± 2.4, 69</td>
<td>13.37</td>
<td></td>
</tr>
<tr>
<td>Acupuncture</td>
<td>5.3 ± 1.9, 70</td>
<td>3.3 ± 2.2, 64b</td>
<td>37.74</td>
<td>0.58</td>
</tr>
</tbody>
</table>

*Comparisons within the same group from baseline to postintervention.

bCompared with the controls, *t* = 2.75, *P* = .0068.

Abbreviations: M, mean; SD, standard deviation.

week 12, acupuncture did induce statistically significant improvements in the primary outcomes: (1) of overall pain score—difference, -1.1; *P* = .002; ES, 0.58—and (2) the WOMAC Function—difference, -3.9; *P* = .04. Changes in the scores on the secondary outcome measurements were also statistically significant for (1) pain on walking (*P* = .003), (2) pain when standing (*P* = .05), and (3) WOMAC pain (*P* = .05). Although the ES of 0.58 for pain relief in the reanalysis (Table 1) was slightly less than the ES of 0.60 that had been set by Hinman et al for the MCID, that ES was larger than the lower threshold of 0.50 that has been proposed by the National Institute for Health and Care Excellence (NICE) for testing nonsteroidal anti-inflammatory drugs (NSAIDs).

The ES for the MCID that was selected by Hinman et al was based on expert opinions, which can sometimes be invalid and unreliable; selection of an ES = 0.60 for the MCID in the study created a stricter requirement for proof of the efficacy of acupuncture than that set for NSAIDs. As per Cohen, who was a world-renowned statistician and psychologist, an ES of ≤0.20 is small, 0.50 is moderate, and ≥0.80 is large. That classification system for ES is well recognized in the academic community and is specifically recommended by NHMRC for the assessment of nonsurgical management of hip and knee osteoarthritis. Thus, based on the above analysis, the acupuncture had actually been moderately effective, even though its treatment protocol (ie, dose and frequency) was not optimally designed in the original study.

It is worth noting that the control group in the study discussed in the JAMA report was not a pure no-treatment group because a significant portion, if not all, of the patients in that group received conventional care. Many participants in the group used various types of pain medications during the study. Consequently, the results of the study at week 12 indicate that a suboptimal acupuncture treatment was moderately effective, as compared with conventional care, for chronic knee pain in patients aged 50 years and older.

Hinman et al’s judgment on the effectiveness of acupuncture at 1 year was not appropriate and did not make any sense because chronic knee pain is mostly due to knee osteoarthritis, a degenerative process. Using a suboptimal dose of acupuncture and providing only 8 to 12 sessions of treatment would not play any significant role in delaying such degeneration after 1 year. In addition, the trial lacked observations or comparisons of immediate and more short-term efficacy for acupuncture, at which times that acupuncture may be more effective.

### ACUPUNCTURE AFTER DATA ADJUSTMENT

As pointed out by various researchers, the sample size calculation for the study discussed in the JAMA report was not appropriate. For a 4-arm study with a Zelen design, 780 to 800 participants were needed; however, only 282 patients were included in the original study. In addition, for the JAMA report, the data for 10 patients who had never had an acupuncture treatment during the study was still pooled with that of the other participants in the acupuncture group.

In the acupuncture group, only 54 patients received acupuncture treatments. Consequently, a 31.27% dilution occurred in the efficacy assessment for the acupuncture group as a result of the Zelen design. Theoretically, any loss in statistical efficiency can be overcome by increasing sample size; however, because the JAMA report gave a post hoc analysis, that increase was not possible. Considering the small sample size of the study, the current research team decided that an adjustment was needed in the data to compensate for the dilution in the estimation of acupuncture’s efficacy for chronic knee pain. Therefore, the team allocated the 10 patients in the acupuncture group who did not receive any acupuncture treatment to the control group that was discussed in the JAMA report.

As Dr Hinman had declined the current research team’s request to provide the original raw data for further analysis, the team has been unable to reconduct the individual, patient-based data analyses to compare the differences among the groups. Nonetheless, the team used a student *t* test to evaluate the efficacy of the acupuncture in comparison with the control group. The data on overall pain for the acupuncture and control groups as shown in the JAMA report were recalculated after the earlier-mentioned data adjustment, shown in Table 2.

### Table 2. Comparisons at Week 12 for the Pain Score After Data Adjustment

<table>
<thead>
<tr>
<th>Group</th>
<th>Baseline M ± SD, n</th>
<th>Week 12 M ± SD, n</th>
<th>Improve (%)</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>5.1 ± 2.1, 71</td>
<td>4.4 ± 2.4, 79</td>
<td>13.37</td>
<td></td>
</tr>
<tr>
<td>Needle acupuncture</td>
<td>5.3 ± 1.9, 70</td>
<td>3.09 ± 2.2, 54b</td>
<td>41.70</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*Comparisons within the same group from baseline to postintervention.

bCompared with the controls, *t* = 3.20, *P* = .0017.

Abbreviations: M, mean; SD, standard deviation.
Under the current research team’s re-evaluations, the new sample size would be 79 for the control group and 54 for the acupuncture group. The team has assumed that the 10 patients in the acupuncture group who did not receive any acupuncture treatment had the same mean score (4.4) and standard deviation (2.4) as those in the control group, because they did not receive any acupuncture intervention. Assignment of those values would equate them to those participants in the control group who also did not receive any intervention.

The new calculation of means thus needed to consider the influence of these 10 patients after the data adjustment. In the acupuncture group with 64 patients, as in the JAMA report, the total overall pain score was 3.3 × 64 = 211.2, whereas the total overall pain score for the 10 patients that were added to the control group under the adjustment was 4.4 × 10 = 44. Therefore, the total overall pain score for the new acupuncture group with 54 patients should be 211.2 + 44 = 167.2. The new mean would therefore be 167.2 / 54 = 3.09.

The current research team used the new means and standard deviations for the acupuncture group for the comparison of that group with the control group in Table 2. Although a suboptimal acupuncture protocol was used in study discussed in the JAMA report, the acupuncture was effective in decreasing the overall pain score at week 12 (P < .01), and the ES was 0.67 when compared with the controls. An ES of 0.67 is higher than the ES of 0.60 for the MCID that had been set by Hinman et al in the JAMA report, with that report assigning an ES of 0.58, as calculated based on original data prior to the current research team’s adjustments, and is also higher than the NICE requirement of 0.50,20 indicating that even with the ES set as 0.60 for the MCID, acupuncture remained effective.

CONCLUSIONS

Based on the analyses presented, the study discussed in the JAMA report was neither a conventional RCT nor an appropriately randomized trial and, therefore, its results are probably invalid. The use of an ES of 0.6 for the MCID and the resulting interpretation were also inappropriate. Even with an ES of 0.60 for the MCID at week 12, acupuncture remained effective after the current research team’s data adjustment. Consequently, compared with conventional care, acupuncture treatment was moderately effective for chronic knee pain in patients aged 50 years and older. The original trial also lacked observations or comparisons of more immediate and short-term efficacy for acupuncture, in which acupuncture may be more effective.

REFERENCES


ACKNOWLEDGEMENTS

The authors would like to thank Ms Sarah Faggert for her editing support.

AUTHOR DISCLOSURE STATEMENT

The authors are independent researchers and declare that they have no competing interests.

Fan—Acupuncture and Chronic Knee Pain