Karen J. Sherman, Japanese Acupuncture and Moxibustion, 2010; Vol.6(1):45-52

The 2 nd JSAM International Symposium on Evidence-based Acupuncture

Clinical trials of acupuncture for low back pain: what have they taught us about acupuncture's effectiveness?

Karen J. Sherman
Group Health Center for Health Studies
Seattle, USA

Abstract

Background: Substantial controversy exists surrounding the effectiveness of acupuncture for low back pain, largely due to poor quality studies. Recent large, well-designed European trials suggest that acupuncture shows similar short term benefits for verum and sham acupuncture needling, but that needling is superior to usual medical care. We designed a two-site randomized controlled trial to evaluate the importance of needle location and insertion on therapeutic benefits of acupuncture.

Methods: 638 patients with chronic back pain were randomized to receive individual acupuncture, standardized acupuncture, simulated acupuncture or care as usual. Ten treatments were given over 7 weeks by experienced acupuncturists. Back-related function and symptoms were assessed at baseline, 8, 26, and 52 weeks by telephone interviewers unaware of treatment group.

Results: Compared to those receiving usual care, patients receiving verum or simulated acupuncture were more likely to show clinically important improvements in dysfunction (60% vs. 39%, $p<0.0001$) or symptom bothersomeness (50% vs. 32%; $p=0.0004$) at the end of treatment, with attenuated benefits persisting for dysfunction over the one year follow-up period.

Conclusion: Acupuncture was more beneficial than care as usual, but neither customizing needle location to the patient nor inserting needles were related to these benefits. These findings extend those of other large trials. They may be interpreted from at least 3 different perspectives (i.e., of medical efficacy trials, of traditionally trained acupuncturists, of whole system researchers). Suggestions for further research and clinical recommendations are provided in light of our knowledge regarding effective treatments for chronic back pain.

Key words: *USA, acupuncture, chronic low back pain, clinical trials, treatment*

Background

Evaluating the therapeutic effectiveness of acupuncture for specific medical conditions has proven remarkably challenging[1)]. This is due, in part, because acupuncture has been traditionally practiced within a non-biomedical framework, encompasses a multiplicity of styles (with variations in diagnostic techniques and theoretical orientations, needling methods, and use of adjuncts), and typically includes multi-modal treatments that are customized to the individual. These features, coupled with difficulties in masking participants and providers to treatment group, have lead to strong controversies regarding the best methods for evaluating acupuncture trials and have rendered assessment of extant trials challenging.

Many of these concerns were raised at the Consensus Development Conference on Acupuncture convened in

Corresponding author: Karen J. Sherman
Group Health Center for Health Studies, 1730 Minor Ave., Suite 1600, Seattle WA 98101
E-mail: sherman.k@ghc.org

1997 by the US National Institutes of Health to examine a number of important questions related to the therapeutic value of acupuncture[2]. After the non-advocacy multidisciplinary panel concluded that the scientific literature could not provide clear conclusions because of methodological limitations, the NIH solicited grant proposals to address such challenges. My colleagues and I began a program of research to evaluate the value of acupuncture needling for chronic low back pain in direct response to this initiative. At that time, there were no meta-analyses or systematic reviews of acupuncture for low back pain, although several would be published before the end of the 1990's[3,4].

Given the tension between the gold standard of randomized double-blind placebo controlled trials and the realities of acupuncture practice, we chose to specifically address issues of appropriate and adequate acupuncture treatment, defining the appropriate comparison group and masking patients and practitioners to treatment. While the evaluation of a new medication would first evaluate how the medication compared to placebo and then evaluate how the medication compared to usual care, we thought a more complex approach was most appropriate for evaluating acupuncture. In particular, we thought there were five critical questions that would comprehensively address the value of acupuncture for a particular clinical condition, in our case chronic low back pain[5]. Collectively, these questions would also enhance our understanding of how acupuncture worked, which we believed was critical for determining more efficient designs for future clinical trials.

Those questions were:

1) Is acupuncture superior to usual care? This question is important for patients and physicians looking for treatments that may be more effective than usual care alone. It is also of value for insurers who need to decide what treatments to cover.
2) Is individualized acupuncture more effective than standardized acupuncture? Acupuncturists typically believe that optimal therapeutic results are obtained when treatments are customized to the patients. However, researchers favor standardized treatments that can be applied to everyone so that the trial results are replicable. Conceivably, similar results could have been found regardless of whether participants received a standardized treatment (wherein everyone received the same point prescription) or received a customized treatment. Answers to this question would have important implications for acupuncture research as well as practice and training.
3) Is there a "specific effect" of acupuncture attributable to the needling procedure itself? If so, persons receiving acupuncture needling would have better results than those receiving a non-insertive treatment stimulating the same acupuncture points. Although most acupuncturists insert acupuncture needles into the body, some do not and there are a variety of other methods of stimulating acupuncture points that might be used. Therefore, we did not automatically assume that persons receiving a non-insertive treatment were getting a "placebo" or "sham" control. Understanding this question would provide important insights into the mechanisms undergirding acupuncture treatment.
4) Is the needling of acupoints that are considered helpful for a specific condition more effective than needling of acupuncture points not used for this condition? This question is of major importance in the design of acupuncture studies because it would shed light on the value of traditional acupuncture theory for predicting benefits from needling. However, it would not answer the question of whether acupuncture functioned strictly as a placebo or caused a beneficial systemic response.
5) Is needling acupoints considered ineffective for a particular condition more effective than non-insertive stimulation? Answering this question would give insights into the relative importance of systemic therapeutic effects that result from acupuncture needling.

Our original intention was to design a five arm trial with two treatment (individualized, standardized) and three control ("misplaced needling", non-insertive acupoint stimulation, and usual care) arms. In this trial, the "misplaced needling" group would receive needling in "inappropriate acupuncture points" for treating back pain, but located in the same anatomic regions of the body. The non-insertive acupoint stimulation group would receive stimulation of acupuncture points used in the standardized group using a non-insertive method. However, findings from our initial project actually suggested that we could not identify "clearly inappropriate" acupuncture points on the back for treating persons with

low back pain. Thus, we omitted this arm from the large clinical trial we subsequently conducted.

Previous studies investigating acupuncture have used a wide variety of "misplaced needling controls", including minimal stimulation of acupuncture and non-acupuncture points, needling "near" the acupuncture points used in the verum treatment, use of a point prescription for another condition, and use of "non-channel, non-point locations"[6]. The wide variety of controls makes interpretation of these studies problematic.

Two recently published large trials from Germany have compared acupuncture with "misplaced" needling acupuncture for persons with chronic back pain[7,8]. Both reported that verum and "sham" acupuncture, the latter defined as superficial needling of non-acupuncture points, were equally effective. The most recent systematic review[9] suggested that, for persons with back pain, acupuncture was better than no treatment and an effective adjunct to biomedical care, in the short term.

Although undertaken prior to the publication of the large German trials, our study extends those findings by including a non-insertive method of stimulating acupuncture points, including two different types of acupuncture treatment; and assessing patient outcomes at 8, 26, and 52 weeks after randomization. Our large 4-arm trial specifically addressed the first three questions described earlier: is acupuncture superior to usual care?; is acupuncture needling superior to non-insertive stimulation?; and is individualized acupuncture superior to standardized acupuncture?

Methods

We conducted a randomized controlled trial comparing individualized, standardized and non-insertive acupuncture point stimulation with each other and with usual care for chronic mechanical low back pain using 638 patients from two integrated health care systems. Our trial protocol[10] and main results[11]have been previously published. Therefore, we give only the most pertinent details of the study here, emphasizing the unique features of each acupuncture treatment group. The study was approved by the ethics boards of both health care organizations. Informed consent was obtained from all study participants prior to their enrollment in the trial. Although non-insertive stimulation of acupoints was described as "simulated acupuncture" in reporting the protocol and main results for a biomedical audience[10,11], the term "non-needle" acupuncture is used here to more clearly describe the type of acupoint stimulation used.

All trial participants were between 18 and 70 years of age, had back pain for at least 3 months and rated it at least 3 on a 0 to 10 scale of pain bothersomeness, had symptoms and history consistent with mechanical low back pain, were not planning on seek other medical care for their back pain, and had never had acupuncture for any reason. Eligible and willing individuals completed the baseline questionnaire via telephone and were then randomized to one of the four treatment groups. We described the study as a study of "three different methods of stimulating acupuncture points", but did not provide further information on the treatments. All trial participants received a self care book, The Back Pain Helpbook[12] and continued access to the care they would have received from their insurance plan.

All participants assigned to one of the needle or non-insertive acupuncture arms were asked to attend 10 treatment sessions, twice weekly for 3 weeks and weekly for an additional 4 weeks. To maintain maximum masking possible in this trial, we used two acupuncturists per treatment and asked participants to lie prone on the table with an eye cover and their head in the face cradle. At each visit, a Diagnostician acupuncturist evaluated each participant using traditional Chinese medical techniques and then prescribed an individualized point prescription, which was used only for individuals randomized to the individualized treatment group. The Diagnostician remained masked to treatment group. A Therapist acupuncturist, who interacted minimally with the participant, actually performed each treatment. Trial compliance was good, with 84% to 90% of persons randomized to one of the acupuncture treatment groups attending a full course (at least 8 of 10) of treatments.

Participants assigned to the individualized acupuncture group received the treatment prescribed by the Diagnostician. These typically included about 11 needles, retained for 15 to 20 minutes, used in acupuncture points on the back and legs.

In the standardized and non-needle groups, the point prescription included Du-3, Bladder-23, a low back ashi

point, Bladder-40 and Kidney-3. For the standardized group, we used 32 gauge needles, retained for 20 minutes. Needle stimulation was obtained by twirling the needles after insertion, at 10 minutes and again just prior to needle withdrawal. For the non-needle group, a toothpick in a needle guidetube was used to mimic needle insertion, while toothpicks were used to simulate needle stimulation in the middle of treatment and before withdrawal. Needle stimulation was simulated by twirling the toothpicks. A more complete description is given elsewhere[13].

In addition to baseline, trial outcomes were assessed at 8, 26, and 52 weeks using computer assisted telephone interviews by interviewers unaware of treatment group. Primary outcomes were back related functional status, measured using the Roland-Morris Disability Questionnaire, and symptom bothersomeness, measured using a 0 (not at all bothersome) to 10 (extremely bothersome) scale[10]. In addition, we defined clinically important improvement as a decrease of 3 points from baseline on the Roland Scale and a decrease of 2 points on the symptom bothersomeness scale. Follow-up rates exceeded 90% at all time points.

Our analyses were based on intent-to-treat, with analysis of covariance used for continuous outcome measures. We adjusted for baseline values of the outcome measure, gender, age group, and geographic location. Separate analyses were conducted for each follow-up period. A Chi-square test was used to look at differences in categorical outcomes.

Results

As described in our main paper[11], 62% of study participants were women, 68% were Caucasian and 53% had graduated from college. Their mean age was 47 years. About two-thirds reported pain lasting at least a year and having used medications in the previous week and a quarter reported reduced activities in the previous 4 weeks because of back pain. Their back related dysfunction and bothersomeness scores prior to randomization were consistent with moderate levels of dysfunction and pain (Tables 1 and 2).

At the end of the treatment period, all groups had both improvement in their back-related function and decreased pain, with greater improvements for the three acupuncture groups (4.1 to 2.5 points for function; 1.6 to 1.9 points for pain) than for the usual care group (2.1 points for function; 0.7 points for pain). Although all three acupuncture groups were superior to usual care, there were no significant differences among them.

Table 1: Mean Roland disability score and percent improved participants over time by randomized treatment assignment

	Treatment group				
	Individualized Acupuncture (n=157)	Standardized Acupuncture (N=158)	Non-Needle Acupuncture (n=162)	Usual Care (n=161)	P value
Roland score, mean (SD)					
Baseline	10.8 (5.2)	10.8 (5.6)	9.8 (5.2)	11.0 (5.2)	
8 weeks	6.4 (5.3)	6.3 (5.7)	5.4 (4.9)	8.9 (6.0)	P <0.0001*
26 weeks	6.8 (5.5)	6.7 (5.8)	6.4 (6.0)	8.4 (6.0)	P = 0.003*
52 weeks	6.0 (5.4)	6.0 (5.8)	6.2 (5.8)	7.9 (6.5)	P = 0.001*
% Improved Participants¶					
8 weeks	60	60	59	39	P<0.001
26 weeks	62	58	58	44	P=0.01
52 weeks	65	65	59	50	P=0.02

* Adjusted for baseline outcome measure, site, age group (18-29, 30-39, 40-49, 50-59, 60-70), and gender

¶ Improvement defined as 3 or more points lower than baseline score

Compared to those receiving usual care, patients receiving any type of acupuncture were more likely to show clinically important improvements in dysfunction (60% vs. 39%, p<0.0001; Table 1) or symptom bothersomeness (50% vs. 32%; p=0.0004; Table 2) at the end of treatment. These differences correspond to a number needed to treat of 5 (95% CI=3.3 to 8.5) for back related function and 6 for pain reduction (95% CI=3.7 to 10.6).

At the 6 and 12 month follow-ups, both back related function and bothersome levels remained similar in the acupuncture groups, but had improved in the usual care group. As a result by 52 weeks, the benefits of acupuncture diminished slightly for back related function (the number needed to treat increased to 8). For pain bothersomeness, there was no advantage for the acupuncture groups by 52 weeks.

Conclusions

Our study found that compared with usual care, stimulation of acupuncture points was associated with clinically important improvements in symptoms and function at the end of treatment. Improvements in function persisted for at least one year. However, we did not find additional benefits for individualizing acupoint selection or for inserting needles.

Our findings are consistent with those of other large trials, which collectively have demonstrated that acupuncture is superior to usual care or best conventional care[7-9,14], but that that needle location and depth of insertion appears unimportant for eliciting these benefits. Our findings extend those of these other trials by showing that customizing acupuncture to the individual or inserting needles is unnecessary to achieve therapeutic benefits and by showing that these findings are consistent over the course of a year.

Three distinct perspectives can be used to interpret these results: that of double-blind placebo controlled medication trials; that of traditional East Asian acupuncture and that of a "whole systems" perspective. Each perspective is shaped by substantially different a priori assumptions about the nature of the comparison and treatment groups. This discussion provides sufficient background to detect the biases of each of these interpretations. Finally, suggestions are provided for further

Table 2 Mean Symptom bothersomeness score and percent improved participants over time by randomized treatment assignment

	Treatment group				
	Individualized Acupuncture (n=157)	Standardized Acupuncture (N=158)	Non-Needle Acupuncture (n=162)	Usual Care (n=161)	P value
Bothersomeness Score, Mean (SD)					
Baseline	5.0 (2.5)	5.0 (2.3)	4.9 (2.4)	5.4 (2.4)	
8 weeks	3.4 (2.7)	3.3 (2.5)	3.0 (2.4)	4.7 (2.6)	P <0.0001*
26 weeks	3.8 (2.5)	3.7 (2.6)	3.5 (2.7)	4.4 (2.6)	P = 0.04*
52 weeks	3.7 (2.6)	3.5 (2.7)	3.4 (2.7)	4.1 (2.6)	P = 0.12*
% Improved Participants¶					
8 weeks	45	55	50	32	P<0.001
26 weeks	49	44	48	41	P=NS
52 weeks	52	49	50	47	P=NS

* Adjusted for baseline outcome measure, site, age group (18-29, 30-39, 40-49, 50-59, 60-70), and gender
¶ Improvement defined as 2 or more points lower than baseline score

research in light of our knowledge regarding effective treatments for chronic back pain.

The biomedical approach to evaluating medications uses randomized double blind placebo-controlled trials to compare a verum medication to a placebo pill that has the same appearance as the real medication, but lacks the active ingredient. Double blinding is important to prevent beliefs about the treatment from interfering with the measurement of treatment effects. Applying this interpretation to acupuncture trials, the active ingredient of acupuncture would be needle insertion into acupuncture points. Therefore, a comparison treatment that resembled the real treatment but lacked needle insertion would clearly be an inactive treatment. Thus, the non-insertive acupoint stimulation treatment would serve as an inactive treatment under this interpretation. Our finding that verum needling had equivalent results to this inactive treatment would clearly indicate that acupuncture needling is not better than sham. The finding that all three acupoint stimulation treatments were better than usual care could easily be explained because of the time and attention given to those individuals in the course of the trial. Using this theoretical framework, acupuncture is ineffective for back pain. Researchers holding this perspective would marshal further support for their view by noting that studies comparing verum acupuncture to "misplaced" needling and conventional care or waitlist also find no difference between verum and 'misplaced needling', which they would argue is evidence that acupuncture theory is not supported by evidence.

The second interpretation stems from the perspective of Traditional East Asian medicine, which recognizes a variety of methods to stimulate acupuncture points and meridians[15]. These include external qi gong, acupressure, non-insertive needling as practiced by Toyo Hari Meridian Therapy, acupuncture using needle insertion with various depths of insertion ranging from superficial to deeper needling, moxibustion, and newer methods of acupoint stimulation including electrical stimulation, laser stimulation, acupoint injection and ultrasound stimulation. This panoply of techniques suggests that collectively acupuncturists believe that there are many means to stimulating acupuncture points. While not all acupuncturists and styles of acupuncture would use all techniques, there is sufficient diversity that the non-insertive comparison would not automatically be considered a "sham" technique. Indeed, many traditionally trained acupuncturists commonly believe that sham acupuncture techniques simply do not exist. Thus, according to this interpretation, the study compared three viable acupuncture treatments and found all of them equally effective for persons with chronic back pain. Researchers holding this perspective would also note that, for back pain, the stimulation of painful areas - no matter their location - discovered by palpation and called ashi points is a common part of traditional acupuncture practice[16]. Thus, it is not surprising that treatments using either bonafide acupoints or other locations on the back would yield similar results, as has been demonstrated by the recent high quality trials in Europe.

The third perspective arises from the study of "whole systems of medicine"[17]. From this perspective, researchers should acknowledge that acupuncture, which includes a unique diagnostic system, a variety of therapeutic elements beyond needling, and lifestyle advice comes from an entirely different theoretical orientation. Our trial, which required two acupuncturists per participant, was highly artificial: restricting the interactions between patient and practitioner and substantially limiting the therapeutic tools available to the practitioner to educate and assist the patient in healing. Thus, from this perspective, the treatment delivered in our study was an extremely simplified and unnatural version of a robust system of medicine, which can be used to engage patients in their own healing. As a result, our findings are likely to underestimate the true value of acupuncture in the context of a traditional system of medicine. In fact, nearly all other trials suffer from the same limitations and acupuncture has not really been evaluated as it is actually practiced in the West.

Distinguishing between these perspectives may not be entirely possible, as they may partially reflect philosophical differences in the approach to healing. However, studies comparing acupuncture practiced as a whole system to best conventional biomedical treatments for different groups of patients with chronic back pain should give us the most optimistic estimate of acupuncture's benefit for relieving low back pain. To date, no large published studies for back pain have actually done this.

To assess the validity of the first two perspectives requires a better understanding of the mechanisms of ac-

tion undergirding acupuncture. While a number of mechanisms have been investigated including opioid release in the brain and spinal cord, signaling via connective tissue[18] and a variety of responses in the brain[19] - with some of these suggesting different brain mechanisms activated in response to insertive and non-insertive stimulation[20], more work is required to understand how subtle differences between the context surrounding acupoint stimulation may impact mechanistic findings.

In the meantime, the favorable safety profile of acupuncture and the lack of clearly best therapeutic options for patients with chronic back pain suggest that acupuncture remains a "reasonable treatment option" for these patients.

References

1) Lewith GT, White PJ, Kaptchuk TJ. Developing a research strategy for acupuncture. Clin J Pain 2006 Sep; 22(7): 632-8.
2) NIH Consensus Development Panel. NIH Consensus Conference. Acupuncture. JAMA 1998 1998 Nov 4; 280(17): 1518-24.
3) Ernst E, White AR. Acupuncture for back pain: a meta-analysis of randomized controlled trials. Arch Intern Med 1998 1998 Nov 9; 158(20): 2235-41.
4) van Tulder MW, Cherkin DC, Berman B, Lao L, Koes BW. The effectiveness of acupuncture in the management of acute and chronic low back pain. A systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 1999 1999 Jun 1; 24(11): 1113-23.
5) Sherman KJ, Cherkin DC. Developing methods for acupuncture research: rationale for and design of a pilot study evaluating the efficacy of acupuncture for chronic low back pain. Altern Ther Health Med 2003 2003 Sep-Oct; 9(5): 54-60.
6) Birch S, editor. Testing the clinical specificity of needle sites in controlled clinical trials of acupuncture. Society of Acupuncture Research - Volume II; 1994 1995; Georgetown University: Society for Acupuncture Research.
7) Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. Arch Intern Med 2006 Feb 27; 166(4): 450-7.
8) Haake M, Muller HH, Schade-Brittinger C, Basler HD, Schafer H, Maier C, et al. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. Arch Intern Med 2007 Sep 24; 167(17): 1892-8.
9) Yuan J, Purepong N, Kerr DP, Park J, Bradbury I, McDonough S. Effectiveness of acupuncture for low back pain: a systematic review. Spine 2008 Nov 1; 33(23): E 887-900.
10) Cherkin DC, Sherman KJ, Hogeboom CJ, Erro JH, Barlow WE, Deyo RA, et al. Efficacy of acupuncture for chronic low back pain: protocol for a randomized controlled trial. Trials 2008;9:10.
11) Cherkin DC, Sherman KJ, Avins AL, Erro JH, Ichikawa L, Barlow WE, et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. Arch Intern Med 2009 May 11; 169(9): 858-66.
12) Moore JL, K; Von Korff, M; Gonzalez, VM;Laurent, DD. The Back Pain Helpbook. Reading, MA: Perseus Books; 1999.
13) Sherman KJ, Hogeboom CJ, Cherkin DC, Deyo RA. Description and validation of a noninvasive placebo acupuncture procedure. J Altern Complement Med 2002 2002 Feb; 8(1): 11-9.
14) Thomas KJ, MacPherson H, Thorpe L, Brazier J, Fitter M, Campbell MJ, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. BMJ 2006 Sep 23; 333(7569): 623.
15) Birch SJ, Felt RL. Understanding Acupuncture. New York: Churchill Livingstone; 1999.
16) Birch S, Sherman K. Zhong Yi acupuncture and low-back pain: traditional Chinese medical acupuncture differential diagnoses and treatments for chronic lumbar pain. J Altern Complement Med 1999 1999 Oct; 5. 5(5. 5): 415-25; discussion 27-8, -25; discussion 27-8.
17) Fonnebo V, Grimsgaard S, Walach H, Ritenbaugh C, Norheim AJ, MacPherson H, et al. Researching complementary and alternative treatments--the gatekeepers are not at home. BMC Med Res Methodol 2007; 7:7.
18) Langevin HM, Churchill DL, Cipolla MJ. Mechanical signaling through connective tissue: a mechanism for the therapeutic effect of acupuncture. FASEB J 2001 2001 Oct; 15(12): 2275-82.

19) Dhond RP, Kettner N, Napadow V. Neuroimaging acupuncture effects in the human brain. J Altern Complement Med 2007 Jul-Aug; 13(6): 603-16.
20) Pariente J, White P, Frackowiak RS, Lewith G. Expectancy and belief modulate the neuronal substrates of pain treated by acupuncture. Neuroimage 2005 May 1; 25(4): 1161-7.