THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE
Volume 12, Number 10, 2006, pp. 943–953

ROUNDTABLE DISCUSSION

Controversies in Acupuncture Research: Selection of Controls and Outcome Measures in Acupuncture Clinical Trials

Chairperson: HÉLÈNE M. LANGEVIN, M.D., L.Ac.[1]

Panelists: RICHARD HAMMERSCHLAG, Ph.D.,[2] LIXING LAO, Ph.D., C.M.D. (China), L.Ac.,[3] VITALY NAPADOW, Ph.D., Lic.Ac.,[4] ROSA N. SCHNYER, Lic.Ac.,[5] and KAREN J. SHERMAN, Ph.D., M.P.H.[6]

This paper summarizes the scientific session and roundtable discussion on acupuncture research held during the North American Research Conference on Complementary and Integrative Medicine, sponsored by the Consortium of Academic Health Centers for Integrative Medicine (CAHCIM), at Edmonton, Alberta, Canada, May 24–27, 2006. The session panelists, co-sponsored by the Society for Acupuncture Research, focused on two main challenges in the design of acupuncture clinical trials: selection of appropriate controls and outcome measures. What follows are highlights of each presentation, as provided by the panelist, as well as an edited transcript of the discussion.

SUMMARIES OF PRESENTATIONS

What is the "Right" Control Group for Acupuncture Trials?

Karen J. Sherman, Ph.D., M.P.H.

Center for Health Studies, Group Health Cooperative, Seattle, WA.

The "right" control group for acupuncture studies will depend on the precise question being asked. In 1997, the NIH Consensus Development Conference asked two broad questions regarding acupuncture that have implications for selecting appropriate comparison groups. These questions were: (1) What is the efficacy of acupuncture compared with placebo or sham acupuncture? and (2) What is the place of acupuncture in comparison or in combination with other interventions?

Answering the first question examines the issue of whether or not acupuncture works by looking for the "specific effects" of acupuncture, which is often taken to mean the effects of needling. While such an approach is typically straightforward in a study of medications or herbs, there are a number of issues that make it more challenging to apply to acupuncture.

In any treatment, there are a variety of factors that could account for improvement in symptoms, including the natural history of the condition, extra attention, enthusiasm of the provider, positive expectations of the patient, the "meaning response," physiological effects of needling anywhere, and physiological effects of needling in the "proper" locations. Typically, a placebo intervention would control for natural history, extra attention, enthusiasm of the provider, positive expectations of the patient, and the "meaning response" and would look exactly like the treatment.

[1]Departments of Neurology, Orthopaedics and Rehabilitation, University of Vermont, Burlington, VT.
[2]Research Department, Oregon College of Oriental Medicine, Portland, OR.
[3]Center for Integrative Medicine, University of Maryland School of Medicine, Baltimore, MD.
[4]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA.
[5]Osher Institute, Harvard Medical School, Boston, MA; and New England School of Acupuncture, Watertown, MA.
[6]Center for Health Studies, Group Health Cooperative, Seattle, WA.

While a variety of controls, including sham laser, sham TENS [transcutaneous electrical nerve stimulation], noninsertion of needles, and "placebo" needles, have been used, ostensibly to control for these effects, each of them has its limitations. For example, sham laser and sham TENS do not look like needles, whereas noninsertion of needles (e.g., poking the blunt end on the skin) or using "placebo" needles, which retract into the handle of the needle, may not be inert treatments.

Insertive sham controls have been controversial because needle insertion, regardless of location, may exert physiological effects. Nonetheless, to control for needle-related factors, such as location and/or stimulation and/or depth of insertion, studies have needled "inappropriate" points, inserted needles superficially, and/or failed to stimulate the needles. Some authors have suggested that a sham control should include superficial needling of "inappropriate" points without needle stimulation.

As a result of controversies over acupuncture research methods identified at the Consensus Development Conference, the NIH [National Institutes of Health][1] issued a Program Announcement requesting applications for pilot studies to address methodological issues in acupuncture research.

Six studies were funded, and five of those dealt with defining appropriate control groups. A study of depression in pregnant women found that "specific" acupuncture was better than "nonspecific" acupuncture and a massage attention control.[2] Among patients with carpal tunnel syndrome, "specific" acupuncture was no better than "irrelevant meridian" or nonchannel, nonpoint acupuncture.[3]

Two studies of fibromyalgia failed to find acupuncture better than the control conditions, which were nonchannel, nonpoint needling with or without stimulation in one study[4] and either acupuncture for another condition, nonchannel, nonpoint acupuncture, or simulated acupuncture in the other study.[5] In the last study comparing two kinds of "real" acupuncture with irrelevant point acupuncture, simulated (noninsertive) acupuncture, or a self-care book, acupuncture was better than the self-care book, there was little difference between any needling group, and the noninsertive simulated acupuncture treatment was intermediate.[6] Collectively, these studies suggest that needling location and stimulation may not be critical for pain conditions, but show conflicting results regarding the value of insertion itself.

The general conclusion was that for placebo-controlled trials, a three-arm design should be considered to anchor the findings in clinical relevance and that fMRI [functional magnetic resonance imaging] studies may shed light on the optimal "placebo" control groups for clinical trials.

The second question looks at the practical issue of where acupuncture is a useful intervention in current clinical practice, while ignoring the question of what components might contribute to acupuncture's effectiveness. In the pragmatic trials undertaken to answer this question, the appropriate comparison group would be another treatment for the condition under study. In the UK [United Kingdom], two large pragmatic trials have compared acupuncture to usual care and found acupuncture superior for back pain and for headache. In the US [United States], one large pragmatic trial compared acupuncture with massage and with self-care materials and found massage superior to acupuncture.

References

1. NIH Consensus Conference in Acupuncture. JAMA 1998; 280:1518–1524.
2. Manber R, Schnyer RN, Allen JJ, et al. Acupuncture: A promising treatment for depression during pregnancy. J Affect Disord 2004;83:89–95.
3. Weinstein A, Pan J, Richardson P, et al. A controlled pilot trial of acupuncture for carpal tunnel syndrome. Presented at the Ninth Annual Symposium of the Society for Acupuncture Research, Seattle, WA, October 18–19, 2002.
4. Harris RE, Tian X, Williams DA, et al. Treatment of fibromyalgia with formula acupuncture: Investigation of needle placement, needle stimulation, and treatment frequency. J Altern Complement Med 2005;11:663–671.
5. Assefi NP, Sherman KJ, Jacobsen C, et al. A randomized clinical trial of acupuncture compared with sham acupuncture in fibromyalgia. Ann Intern Med 2005;143:10–19.
6. 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;9:54–60.

Invasive Sham Controls: Does the Evidence Support Their Use? A Critical Review of the Literature 1997–2006

Rosa N. Schnyer,[1,2] Maria Broderick,[2] Hudson Doyle,[2] Lauren Hansen,[2] Helen Ly, Jonathan Stegall, Ryan Milley,[3] Robert Scholten,[1] and Paul Bain[4]

[1]Osher Institute, Harvard Medical School, Boston, MA.
[2]New England School of Acupuncture, Watertown, MA.
[3]Oregon College of Oriental Medicine, Portland, OR.
[4]Harvard Medical School, Boston, MA.

At the 1997 NIH Consensus Conference on Acupuncture,[1] conclusive evidence on the efficacy of acupuncture was found only for two conditions, dental pain and postoperative nausea. The final statement of the Consensus Report included a recommendation to design and conduct well-pow-

ered, well-designed controlled acupuncture trials. Almost 10 years later, the evidence for the efficacy of acupuncture remains inconclusive or contradictory for a large number of conditions. This may be due at least in part, to methodological challenges specific to conducting clinical trials of multivariate, complex interventions such as acupuncture. One of those challenges involves the selection of appropriate controls.

In order to evaluate whether invasive sham controls have produced equivocal or contradictory results more often than noninvasive controls in clinical trials of acupuncture, we are currently conducting a critical review of literature to include all randomized controlled trials of acupuncture conducted in 1997–2006. Our purpose is to identify those studies that have used acupuncture-like sham controls, either invasive or noninvasive. Invasive sham controls are defined by the use of devices that penetrate the skin and that aim at controlling the selection of points (i.e., nonspecific/sham versus acupoints), the type of stimulation (i.e., shallow versus standard needling) or both. So far, we have identified a total of 107 randomized controlled trials of acupuncture that used an acupuncture-like sham control.

During 1997–2004, 43 studies met this criterion (10 studies with sample size 50–99; 7 studies $>$ 100). Thirteen used a traditionally based style of acupuncture (TBSA), all of which used TCM [Traditional Chinese Medicine], and one of which also included Japanese acupuncture. Six (6) used individualized treatments, 1 combined electroacupuncture with manual stimulation, and 1 combined acupuncture with herbs. More than half (59%, 10 studies) found acupuncture to be effective over the sham-controlled intervention (5, n = 100 or more); 5 studies focused on pain-related conditions; 90% of these studies used an invasive sham control.

For studies conducted between January 2005 and through April 2006, we identified 29 randomized controlled trials of acupuncture; 14 met our criteria (five studies n = 50–99; 9 studies n = 100 or more). Eleven (11) studies used a TBSA (10 used TCM; 1 used Korean); 6 studies used individualized treatment, and all used manual stimulation, noncombined acupuncture with herbs. Of all 14 studies, only 1 study found acupuncture to be significantly more effective than the sham-control intervention[2] in the primary outcome, and only 1 study found [that] acupuncture outperformed the sham controlled in secondary outcomes.[3] Ninety percent of those studies used an invasive sham control. Six of seven studies (n = 100 or more), which found acupuncture to be no more effective than a sham control, aimed at evaluating acupuncture in the treatment of pain-related conditions.

Although these results are very preliminary, it is interesting to note that most recent large randomized controlled trials of acupuncture (n = 100 or more) have used an invasive sham control and most have failed to find a significant difference between acupuncture and an acupuncture-like sham intervention.

In order to design an appropriate intervention that controls for the specific effects of acupuncture, it is necessary to first understand the mechanism by which acupuncture works. Appropriate controls in clinical trials of acupuncture need to derive from a clear understanding of the mechanism of action of acupuncture (e.g., the needling process, and the specificity of acupuncture points). Acupuncture sham controls must be inert as defined by the Chinese Medicine explanatory model; invasive sham controls may be an inadequate control for clinical trials of acupuncture.

References

1. NIH Consensus Conference in Acupuncture. JAMA 1998; 280:1518–1524.
2. Witt C, Jena, S, Brinkhaus B, et al. Acupuncture in patients with osteoarthritis of the knee: A randomised trial. Lancet 2005; 366:36–43.
3. Emmons S, Otto, L. Acupuncture for overactive bladder: A randomized controlled trial. Obstet Gynecol 2005;106:138–143.

Sham/Placebo Controls in Acupuncture: The Evidence from Neuroimaging

Vitaly Napadow, Ph.D., Lic.Ac.

Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital & Harvard Medical School, Boston, MA.

The processing of acupuncture stimuli by the brain can be encapsulated by a combination of expectation (prior), sensation (during), and contextualization (post). While sham forms of acupuncture have been designed to control for supposed "nonspecific" effects, these controls also influence expectation, sensation, and contextualization. Thus, there is potential overlap in effects of both sham and verum acupuncture.

Furthermore, as the "specific" mechanisms of verum acupuncture have not been unequivocally identified, designing a sham control that stimulates only nonspecific effects has been difficult. Top–down descending pain modulation networks may be stimulated to a variable degree by both verum and sham acupuncture. Neuroimaging is a potential approach to disentangle these effects, as there may exist a specific brain response attributable to verum compared with sham acupuncture interventions.

Neuroimaging (e.g., fMRI [functional magnetic resonance imaging], PET [positron emission tomography], etc.) noninvasively explores brain activity in humans in response to various stimuli, and past studies have provided important clues regarding how the brain responds to different sham acupuncture stimuli. Furthermore, neuroimaging can demonstrate potential differences between patient populations and healthy subjects in their response to both sham and verum acupuncture.

Most previous acupuncture neuroimaging studies have been performed on healthy subjects. Controls in these studies have included insertive nonacupoint stimulation, such as minimal electro-acupuncture and superficial pricking. Additionally, noninsertive acupoint stimulation has been performed with "placebo-needles"[1] or, since the fMRI environment precludes visual feedback, with von Frey monofilaments.[2–4] These procedures attempt to control for somatosensory and/or cognitive aspects of acupuncture.

Insertive, nonacupoint stimulation has been explored with fMRI. Minimally sensible, shallow-insertion electroacupuncture was found to produce fMRI activation in both somatosensory (thalamus) and cognitive (dorsolateral prefrontal cortex, DLPFC) brain regions.[5] Wu et al.[6] also found that superficial pricking with an acupuncture needle at a nonacupoint produces activity not only in somatosensory (thalamus, SI, SII) and cognitive (DLPFC) brain regions, but also in affective/motivational pain processing areas (anterior cingulate cortex, ACC).

Noninsertive sham acupuncture at an acupoint has been explored with both PET and fMRI. Pariente et al. used PET in osteoarthritis patients to compare LI-4 [acupuncture point Large Intestine–4] stimulation with verum acupuncture, Streitberger-needle, or overt placebo.[7] They found that placebo-needle stimulation produced greater activation than overt placebo in both cognitive (DLPFC) and pain-modulatory (rostral ACC, rACC) brain regions. Our group has used fMRI to explore differences in verum and sham acupuncture processing in chronic pain (carpal tunnel syndrome, CTS) patients compared with healthy controls (HC).[8] Sham acupuncture was delivered with von Frey monofilament to LI-4, and subjects were instructed to expect "different forms" of acupuncture. A conjunction analysis found that both CTS and HC responded to sham acupuncture with activation in somatosensory brain regions (thalamus, SII, left posterior insula) and deactivation in the hippocampus. However, CTS patients demonstrated greater fMRI activation than HC not only in somatosensory brain regions (SI), but also in important cognitive (DLPFC) and affective/motivational pain modulatory areas (rACC/ventromedial prefrontal cortex, VMPFC).

The ACC, which has been implicated in several sham acupuncture neuroimaging studies, may be an especially pertinent brain region, because it has been activated in a host of placebo analgesia studies. The pregenual/subgenual ACC is an area of the brain rich in endogenous opioid receptors and is intimately interconnected with other brain regions that process pain, reward, autonomic, and neuroendocrine functions. Zubieta et al.[9] used PET to localize placebo drug infusion–induced μ-opioid neurotransmission to this ACC subregion. Furthermore, placebo analgesia using the Streitberger needle was found to also be localized to a similar rACC region in a recent fMRI study.[10] Moreover, brain activity in the ACC can be used as real-time biofeedback to modulate pain levels in patients with chronic pain.[11] Also, fMRI activity in a more dorsal ACC subregion has been observed during pain empathy in subjects who were alternately subjected to a pain stimulus or observed their loved one experiencing the same pain stimulus.[12]

Ultimately, the mind is a powerful modulator of pain and other subjective clinical or subclinical symptoms. Most forms of sham acupuncture have biophysical effects and can impact expectation, sensation, and contextualization. Thus, brain processing in affective and cognitive brain regions responsible for top–down descending analgesic control may help explain the powerful placebo effects observed in several recent large acupuncture randomized controlled trials. In choosing appropriate controls, it may be best to stay away from more invasive sham procedures, which have been shown to affect pain processing regions. Furthermore, chronic pain patients may process even noninvasive sham acupuncture to a greater degree in affective/cognitive brain regions compared with healthy controls. Hence, the caveat above may be even more pertinent for trials with clinical pain populations.

References

1. Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. Lancet 1998;352:364–365.
2. Hui KK, Liu J, Makris N, et al. Acupuncture modulates the limbic system and subcortical gray structures of the human brain: Evidence from fMRI studies in normal subjects. Hum Brain Mapp 2000;9:13–25.
3. Napadow V, Makris N, Liu J, et al. Effects of electroacupuncture versus manual acupuncture on the human brain as measured by fMRI. Hum Brain Mapp 2005;24:193–205.
4. Yoo SS, Teh EK, Blinder RA, Jolesz FA. Modulation of cerebellar activities by acupuncture stimulation: Evidence from fMRI study. Neuroimage 2004;22:932–940.
5. Wu MT, Sheen JM, Chuang KH, et al. Neuronal specificity of acupuncture response: A fMRI study with electroacupuncture. Neuroimage 2002;16:1028–1037.
6. Wu MT, Hsieh JC, Xiong J, et al. Central nervous pathway for acupuncture stimulation: Localization of processing with functional MR imaging of the brain—preliminary experience. Radiology 1999;212:133–141.
7. Pariente J, White P, Frackowiak RS, Lewith G. Expectancy and belief modulate the neuronal substrates of pain treated by acupuncture. Neuroimage 2005;25:1161–1167.
8. Napadow V, Kettner N, Liu J, et al. Hypothalamus and amygdala response to acupuncture stimuli in carpal tunnel syndrome. Pain 2006; in press.
9. Zubieta JK, Bueller JA, Jackson LR, et al. Placebo effects mediated by endogenous opioid activity on mu-opioid receptors. J Neurosci 2005;25:7754–7762.
10. Kong J, Gollub RJ, Rosman IS, et al. Brain activity associated with expectancy-enhanced placebo analgesia as measured by functional magnetic resonance imaging. J Neurosci 2006;26:381–388.
11. deCharms RC, Maeda F, Glover G, et al. Control over brain activation and pain learned by using real-time functional MRI. Proc Natl Acad Sci U S A 2005;102:18626–18631.

12. Singer T, Seymour B, O'Doherty J, et al. Empathy for pain involves the affective but not sensory components of pain. Science 2004;303:1157–1162.

Sham Acupuncture Controls: What Can We Learn from Animal Studies?

Lixing Lao, Ph.D., C.M.D. (China) L.Ac.

Center for Integrative Medicine, University of Maryland School of Medicine, Baltimore, MD

There is now a large body of published acupuncture studies, many of them well-designed, but results are inconclusive. One large clinical trial on the effect of acupuncture on osteoarthritis of the knee showed that traditional Chinese acupuncture is superior to sham acupuncture control in reducing pain and improving function.[1] On the other hand, another clinical trial on the same condition showed no significant difference between acupuncture and control,[2] perhaps due largely to the challenge of devising invasive sham acupuncture controls. However, animal studies that explore acupuncture mechanisms often show positive results compared to sham control.

Animal studies and clinical trials present different challenges. In human pain studies, the placebo effect plays an important role in treatment response and may be triggered by the patient's psychology or prior knowledge of the intervention. Furthermore, patients with different diseases and pain conditions may respond differently to acupuncture. Challenges in animal studies include nonspecific effects due to experimental procedures (e.g., from restraints or noxious stimuli that elicit spinal reflex and transient reactions, lack of clarity on the relevance of animal acupoints to those in humans, and variations in severity of disease among models). Nevertheless, animal studies yield much information on such issues as "dosage" and point specificity.

Animal studies show that acupuncture effects are dose-dependent, where "dosage" includes the intensity of electroacupuncture (EA) at acupoints[3] and the amount of therapeutic material injected into a point.[4,5] For example, Kim et al.[4] showed in a rat model that bee venom at ST36 [Stomach 36] suppressed formalin-induced inflammation in a dose-dependent manner. Another laboratory reported that an *Ephedra sinica* extract injected into ST36 produced a dose-dependent inhibition of TNF-α [tumor necrosis factor-α] and COX-2 [cyclooxygenase-2] expression in PMA/LPS*-induced inflammation.[5]

Additional animal studies have shown that acupuncture effects are point specific. Koo et al.[6] reported that treatment at SI6 [Small Intestine 6] produced significant analgesia on a rat model of ankle sprain pain compared to treatment at a nearby point. Lee and Beitz[7] showed that EA on normal rats at ST36 activated brainstem regions but had no such effect at the hamstring muscle. Lao et al.[3] reported similar results in a persistent inflammatory rat model. Tjen-A-Looi et al.[8] investigated the effect of acupuncture on cardiovascular response in a cat model, finding that EA at P5-P6 [Pericardium 5-6], but not LI6–LI7 [Large Intestine 6–Large Intestine 7] or K1-B67 [Kidney 1–Bladder 67], caused cardiovascular or rostral ventrolateral medulla [rVLM] neuronal effects.

In a rat study investigating the effect of acupuncture on ethanol withdrawal, Zhao et al.[9] reported that acupuncture at H7 [Heart 7], not P6 [Pericardium 6] or the tail, significantly prevented release of dopamine during the ethanol challenge.

The high degree of point specificity for given conditions in animal studies contrasts to human clinical trial results showing similar effects from sham treatment at nonspecific acupoints and intended acupuncture intervention.

This difference may be due to several factors:

1. Animal studies use naïve animals with no prior exposure to acupuncture.

2. The controllable pathology and severity of the model make it possible to configure conditions adequate to test acupuncture effects. For example, an adequate dosage of an inflammatory agent will produce a model that allows researchers to differentiate acupuncture effects from those of sham treatment.[10]

3. Experimental animals are often genetically homogeneous, with little variation from individual to individual, which narrows the standard deviation in statistical analysis.

4. Animal studies also may be underreporting animals that fail to show the desired effect, as the intention-to-treat concept has not been widely applied to statistical analysis in animal studies. However, this may be justifiable as, unlike human clinical trials, the major, if not sole, purpose of animal studies is to explore mechanisms of action. A reliable animal model capable of reproducible results is critical for animal basic science.

*PMA/LPS = phorbol myristate acetate/lipopolysaccharide.

References

1. Berman BM, Lao L, Langenberg P, et al. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: A randomized, controlled trial. Ann Intern Med 2004; 141:901–910.
2. Scharf HP, Mansmann U, Streitberger K, et al. Acupuncture and knee osteoarthritis: A three-armed randomized trial. Ann Intern Med 2006;145:12–20; summary for patients in: Ann Intern Med 2006;145:I17.
3. Lao L, Zhang R-X, Zhang G, et al. A parametric study of electroacupuncture on persistent hyperalgesia and *Fos* protein expression in rats. Brain Res 2004;1020(1–2):18–29.
4. Kim HW, Kwon YB, Ham TW, et al. Acupoint stimulation using bee venom attenuates formalin-induced pain behavior

and spinal cord fos expression in rats. J Vet Med Sci 2003; 65:349–355.
5. Yeom MJ, Lee HC, Kim GH, et al. Anti-arthritic effects of Ephedra sinica STAPF herb-acupuncture: Inhibition of lipopolysaccharide-induced inflammation and adjuvant-induced polyarthritis. J Pharmacol Sci 2006;100:41–50.
6. Koo ST, Park YI, Lim KS, et al. Acupuncture analgesia in a new rat model of ankle sprain pain. Pain 2002;99:423–431.
7. Lee JH, Beitz AJ. The distribution of brain-stem and spinal cord nuclei associated with different frequencies of electroacupuncture analgesia. Pain 1993;52:11–28.
8. Tjen-A-Looi SC, Li P, Longhurst JC. Medullary substrate and differential cardiovascular responses during stimulation of specific acupoints. Am J Physiol Regul Integr Comp Physiol 2004;287:R852–R862.
9. Zhao RJ, Yoon SS, Lee BH, et al. Acupuncture normalizes the release of accumbal dopamine during the withdrawal period and after the ethanol challenge in chronic ethanol-treated rats. Neurosci Lett 2006;395:28–32.
10. Zhang R-X, Lao L, Wang X, et al. Electroacupuncture attenuates inflammation in a rat model. J Altern Complement Med 2005;11:135–142.

Translational Research Strategy for Developing Objective Outcome Measures Reflecting Traditional Chinese Medicine

Hélène M. Langevin, M.D., L.Ac.

Departments of Neurology, Orthopaedics and Rehabilitation, University of Vermont, Burlington, VT

Testing the effectiveness of a treatment is difficult if one does not understand the pathophysiology of the condition being treated. For example, testing the effect of a new drug for diabetes would be very difficult if one were measuring how the patient feels, instead of measuring the blood sugar.

In Traditional Chinese Medicine [TCM], the condition being treated may be a syndrome such as "Spleen *Qi* deficiency" or "Liver *Qi* stagnation," but the pathophysiological basis underlying these syndromes remains unknown (Spleen *Qi* deficiency is associated with symptoms such as loose stools, poor muscle tone, and prolapse of tissues, while Liver *Qi* stagnation is associated with oppression in [the] chest, abdomen, and throat, and emotional frustration and anger).

Let us consider a hypothetical clinical trial of acupuncture for myofascial pain. Increasingly, investigators conducting such trials are placing patients into subgroups according to TCM diagnoses for the purpose of treatment selection.

Subjects recruited to participate in this trial thus may include some with Spleen *Qi* deficiency and some with Liver *Qi* stagnation. So far, however, the outcome measures used in such trials are still related to the Western diagnosis (i.e., myofascial pain) rather than the TCM condition. If specific pathological abnormalities related to myofascial pain and associated with TCM syndromes could be identified, more-specific clinical trials could be performed using objective outcome measures that reflect not just pain, but also TCM.

A translational research strategy aimed at developing objective outcome measures that reflect TCM, therefore, can be outlined as follows:

First, identify physiological processes related to Spleen and Liver function that may be relevant to myofascial pain (e.g., maintenance of normal tissue composition, structure, viscoelasticity, and water content). Second, compare tissue structure, viscoelasticity, etc. . . . in groups of patients with Spleen *Qi* deficiency to those with Liver *Qi* stagnation. Finally, conduct a clinical trial of acupuncture in patients with chronic myofascial pain and Spleen *Qi* deficiency or Liver *Qi* stagnation, using these tissue measurements as outcome measures (in addition to subjective "whole person" outcomes).

Such a translational approach "from the ground up" first seeks a more thorough understanding of the condition being studied, then moves on to use this basic knowledge to objectively measure efficacy. Although this type of approach is likely to take much time, the rewards at the end of the road may be well worth the effort.

Do Sham-Controlled Trials Underevaluate the Real-World Effectiveness of Acupuncture?

Richard Hammerschlag, Ph.D.

Research Department, Oregon College of Oriental Medicine, Portland, OR

Arguably, a case can be made that inclusion of a sham acupuncture arm in a randomized controlled trial (RCT) will lead to an underestimation of treatment efficacy. This seeming anomaly, inherent in the acupuncture versus sham acupuncture research design, may be a net result of a *decreased* benefit to the treatment group, relative to the putative benefit of the same acupuncture treatment provided in an unblinded manner, and an *increased* benefit to the sham (control) group, relative to the benefit of a placebo response alone.

Reduced benefit to the treatment group may be an unappreciated confounder in all placebo or sham-controlled trials if findings from sequential trials, in which the same dose of a new opioid analgesic was compared first to a placebo and then to a standard-care analgesic, are generalizable.[1] The explanation offered as to why the new drug was found more effective when compared with standard care than when compared with placebo was based on presumptive, consent form–induced expectations. Participants learned during consenting that they stood a 50% chance of receiving active treatment in the first trial, but that either arm in the second trial would bring analgesic benefit.

Enhanced benefits from sham acupuncture, which are especially likely from *invasive* sham needling (even at nonacupuncture point sites), have been suggested to result

from locally induced changes in microcirculation, immune system responses, and even nonspecific pain modulation mechanisms.[2] More generally, because of the paucity of our understanding of how acupuncture works (e.g., what the needle stimulates, what signals are generated, and how homeostatic regulation is triggered), sham needling may inadvertently activate (albeit less robustly) the same regulatory mechanisms that are triggered by needling at acupuncture points.[3] Just as an appropriate sham should mimic all aspects of practitioner delivery of treatment, it should not mimic any aspect of the physiological action of the treatment. Even *noninvasive* sham procedures, such as tapping an empty needle guide tube on the skin surface, may trigger physiological regulatory responses, especially if the tube is tapped at acupuncture points.[4]

Thus, to the extent that design features of sham-controlled RCTs may reduce the effects of the true acupuncture treatment as well as enhance the effects of the sham treatment, such trials are prone to type 2 (false-negative) errors. It follows that such trials may not adequately assess the effectiveness of real-world acupuncture treatments. One suggestion that meets this concern is for future sham-controlled RCTs to randomize participants to three arms: acupuncture and sham acupuncture (as in the currently accepted manner that purports to ensure blinding) and a third arm in which patients receive the same acupuncture protocol *and are told they are receiving real treatment*. This approach allows both explanatory (blinded) and pragmatic (unblinded) assessments of acupuncture and may reveal whether acupuncture effectiveness is being systematically underevaluated in the current design of sham-controlled RCTs.

References

1. Skovlund E. Should we tell patients that they might receive placebo? Lancet 1991;337:1041.
2. Birch S. A review and analysis of placebo treatments, placebo effects and placebo controls in trials of medical procedures when sham is not inert. J Altern Complement Med 2006;12: 303–310.
3. Hammerschlag R, Zwickey H. Evidence-based complementary and alternative medicine: back to basics. J Altern Complement Med 2006;12:349–350.
4. Berman BM, Lao L, Langenberg P, et al. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: A randomized, controlled trial. Ann Intern Med 2004;141: 901–910.

ROUNDTABLE DISCUSSION

Hélène Langevin, M.D., L.Ac.: I hope that these presentations have helped illustrate what makes the field of acupuncture research both interesting and challenging. We will now open up the discussion to members of the audience.

Ronald Glick, M.D., (University of Pittsburgh): I'm involved in acupuncture research and my observation is that when I use tiny, skinny needles and place the needles superficially in nonacupuncture points, rarely, in less than 100 insertions, does the patient or the subject have anything beyond a very mild sensation of a needle going through the skin. So if *de qi* is the operative phenomenon, then how does one interpret that observation?

Rosa N. Schnyer, Lic.Ac.: *De qi* is defined in different ways depending on who is interpreting the classic text. In TCM, the practitioner is supposed to feel *de qi*, and the patient is also supposed to feel *de qi*. In Japanese acupuncture, as a practitioner, you can perceive *de qi* just by feeling with the hand that's not your needling hand but the patient does not necessarily have to feel it. So we don't really know that *de qi* is the defining aspect of acupuncture. It could be that *de qi* as assessed in TCM denotes a specific effect elicited by stronger needle stimulation, different from that of Japanese acupuncture, which consists of lighter needling stimulation.

Since I became a researcher, I'm very critical of what I do as a clinician. I'm constantly observing in the clinic the specificity of what I'm doing. I can tell you, with 20 years of clinical experience, having seen hundreds if not thousands of patients, that there is specificity to what I do. It doesn't always work and I know how to manipulate what I do to make it work. And we just need to try to understand where the specificity actually does lie.

Jun Mao, M.D., (University of Pennsylvania School of Medicine): Every speaker thus far has alluded to the importance of psychosocial factors in acupuncture research; however, I don't think that we are actually measuring these factors adequately in large clinical trials of acupuncture. Systematically measuring these variables across prospective studies will allow us to develop either a predictive or an explanatory model to inform the biobehavioral mechanism of acupuncture. Having similar measures across studies would then also allow us to pull the data for future meta-analysis or systematic analysis. Important components such as patient expectation of outcomes, the perceived experience of the *de qi*, and aspects of the patient/provider relationship can and should be measured.

A second point I would like to make is that with any therapy, whether conventional or complementary, there are responders and nonresponders. As an acupuncturist, I see people who respond really well and people who don't respond at all. Understanding for what type of population acupuncture works will help guide clinical care to provide patients with the right types of therapies. How are we going to address this issue in clinical trial design? In the end, the most clinically important question is not whether acupuncture works better than sham acupuncture; instead it should be for whom and for what conditions does acupuncture—sham or real—work, either through psychosocial factors or via specific needling.

Karen J. Sherman, Ph.D., M.P.H.: In the absence of hypotheses in advance, what one typically does is analyses of responders and nonresponders, trying to figure out what predicts what. Then, in the future, you can do trials where if you can try to predict in advance, you can stratify people and see if you get some type of effect modification.

Lixing Lao, Ph.D., C.M.D. (China), L.Ac.: Your question about responders and nonresponders is interesting. I have observed that the same patient can respond well when treated for one condition and not respond when treated for another condition. So I'd be careful when defining a person as a nonresponder, because severity of disease and condition also matter.

Dr. Langevin: That is a good point, because we don't want to just lump everything together. When we ask: "Is the patient responding to acupuncture?" in the real world—in the clinic—the acupuncturist is treating the patient for a specific condition or diagnosis, not simply for response to acupuncture in general.

Remy R. Coeytaux, M.D., Ph.D., (University of North Carolina): This question is for Dr. Langevin. I agree and appreciate your argument that we need to spend more effort in looking at the pathophysiology, the mechanisms of conditions, and markers. But there are many clinical trialists who favor the use of clinical outcomes or patient-oriented outcomes that matter. Objective outcomes are maybe more directly relevant than surrogate measures. For instance, the example that you gave of diabetes, one could not understand the mechanism of the disease itself but look at outcomes, such as hospitalizations or complications, that develop. So I'd like for you to comment on the possibility of either in parallel, or instead of, focusing on mechanisms when those might be difficult to really establish, to look at objective outcomes that are measurable but are not specifically related to the mechanism.

Dr. Langevin: I think you put it very well. The most pragmatic way we could look at this is to say: "Does this save health care dollars?" Some of the pragmatic trials are getting at that, looking at the population and comparing people having acupuncture versus people having some other form of health care. You see how many times they have to go to the emergency room or how many MRI scans they need. I think we need to do both, clearly. I think they are both ends of a very long spectrum, and we cannot look only at one end. I totally agree.

Charlotte Paterson, M.R.C., (University of Bristol, United Kingdom): I think what we need to question is this rush into sham acupuncture trials that's happening, and whether this is the right way to go. To do a sham acupuncture trial, as Karen says, is all about deconstruction; deconstructing the therapy into different bits and putting some in some boxes and some in others. But if we believe acupuncture to be a holistic therapy, and certainly its theory base would suggest that it is, then maybe it can't be deconstructed without losing what it's all about. The whole point of a holistic therapy is the synergy and that things relate to each other.

In this vein, I draw upon a lot of qualitative work that I've done interviewing patients who've had acupuncture. Certainly, from their perspective, they describe a lot of [things that are] in actuality about holism. Taking Richard's examples, for example, that what we do with a sham acupuncture trial actually decreases the specific effect of real acupuncture, well I've certainly got a lot of examples of that. For example, in real life, the needling that goes on develops over a long period of time after the initial diagnosis and emergent diagnosis, changing needling according to what the patient says about new illnesses that arrive. It's a whole emergent process, whereas in a strictly experimental acupuncture trial the needling can't be like that.

The other side of sham acupuncture is that it also increases the benefit in the control group. So qualitative research does support this argument and does make us question whether a sham control design is actually appropriate for a holistic therapy.

Richard Hammerschlag, Ph.D.: As a quantitative researcher, I'm always delighted to learn more about the benefits and, in this case, the support from qualitative research.

Dr. Sherman: Of course, I agree with you. But I also think it would be very interesting, in the process of attempting some amount of deconstruction with its various problems, to actually see what fraction of acupuncture is more or less due to the different pieces that are hypothesized to be part of the healing benefits from acupuncture treatments. I have no idea what the answer is, although I have some thoughts on it and I think that actually would be quite fascinating. You could do a study with even more treatment groups, [such as] probably 10 arms, where you have a lot of these different pieces and actually get some estimates of how relatively important the different pieces are. Holism might turn out to be enormous, or it might turn out to be a relatively small aspect once you actually look at what fraction contributes to different aspects of the final outcome.

Richard Harris, Ph.D., (University of Michigan): I agree that we are at a crossroads with the data coming in from the clinical trials of acupuncture, especially in the pain conditions, showing that there's very little added benefit of real acupuncture over sham acupuncture. I am wondering if we might want to step aside and think of other possible research designs.

What does the panel think of an acupuncture trial that, instead of fighting the sham or the placebo effect, one had two arms [that] had opposing effects of the treatment? Si-

multaneously, [you would] be having an increased effect size, because you are comparing two conditions [that] have opposite effects, you would be reducing placebo effects, and [you would] be assessing acupuncture's efficacy or methods, because you designed the treatment arms to have opposing effects. I realize this couldn't be used in all conditions.

Maybe you wouldn't want to have a condition such as pain where you're both simultaneously reducing pain and then increasing pain. But in other conditions [such as] alertness or fatigue, where you have a slight reduction in fatigue versus a slight increase in fatigue, you may be able to show efficacy.

Ms. Schnyer: We explored that possibility when I was at the University of Arizona and did a very small pilot on that. You know, acupuncture has an adaptative effect. It's very difficult to control whether you're going to be stimulating alertness or producing relaxation. It is extremely difficult, because it has to do with the state in which the person comes into the treatment, the amount of stimulation you provide, and the combination of points that you're using. I don't know that you could easily design a trial in that way and consciously produce opposite effects.

Dr. Harris: I'm thinking in terms of Chinese medicine where *yin–yang* balance and oppose each other, and excess and deficiency are treated. Chinese medicine might be amenable to that type of a trial design.

John Longhurst, M.D., Ph.D., (University of California, Irvine): In response to a statement made by an earlier questioner, there is truly a group that doesn't respond to acupuncture. We find that in our animal studies, and we find in our human studies that about 30% of people and animals simply do not respond to acupuncture. We have to take that into account. I think we know some of the mechanisms behind that, but I don't think that we know all of the mechanisms. Whether it's due to cholecystokinin or some other mechanism, I'm not sure, but that certainly is part of the process.

As a scientist, as a clinician, I find that one of the things that really puts off most of the Western allopathic community with regard to acupuncture, or any of the other areas of [complementary and alternative medicine] or [integrative] [CAM] or medicine, is the fact that we don't know much about how they work. So I would like to make the argument that it's as important for us to understand the mechanisms as well as anything else.

I also have concerns with regard to using imaging to study acupuncture. One of the difficulties has to do with the influence of acupuncture on the central nervous system. Acupuncture's effect there is as much a decrease in activity as it is an increase in activity. The release of opioids and the influence of GABA [γ-aminobutyric acid], which actually quiets down activity, is very important. But to look at this decrease in activity I think is much harder to do with the imaging studies. So I think we need to realize that's a limitation that the imaging studies will always have.

Dr. Hammerschlag: I am aware of the reports of poor responders versus good responders for acupuncture analgesia and your allusion to the cholecystokinin system acting to counter endogenous opioid effects. There's certainly excellent evidence both in animals and humans for that phenomenon. There's also been some elegant rat work with molecular biological techniques of converting poor responders into good responders. One would love to have some acupuncture techniques that, in themselves, could convert poor responders into good responders. But my question for you is: In your acupuncture research on the cardiovascular system, do you see the same phenomena in animals and humans of poor responders and good responders?

Dr. Longhurst: You see exactly the same thing with our cardiovascular work, whether they're anesthetized or nonanesthetized, in humans and in animals.

Dr. Hammerschlag: When we talk about acupuncture mechanism, we often think by analogy to the nervous system where we know, for example, what the receptor is that mediates a drug response. We also know the initial transduction event, and what the signal is. But we don't know with confidence what these analogous phenomena are for acupuncture. Although fMRI and other imaging studies are giving us remarkable insights of neural correlates of acupuncture, we don't really have a clear understanding of what happens between insertion of the needle and the brain activity that results in a functional MRI image. We don't know what the pathway is.

With a drug that has an effect on the immune system or endocrine system, we can measure changes in relevant biomarkers as well as correlated changes in brain function. We describe the results as an immune response or an endocrine response that the brain monitors. But with acupuncture, we seem to be assuming it is neurally mediated because we observe neural signals in the brain.

Vitaly Napadow, Ph.D., Lic.Ac.: Neuroimaging is not the "be-all and end-all" of physiological testing of acupuncture in animal or human models. It has its limitations just like any other technique. But it happens to be one of the best ways that we can study brain functions in humans. Some of the more invasive techniques that can be done in animals cannot be done in humans.

Having said that and taking the point made about deactivation to heart, there actually has been quite a series of studies that have been published showing deactivation in certain brain regions in response to acupuncture. All of

these studies have shown signal deactivation or decreases from neuroimaging in the brain in response to the acupuncture stimulus. So there have been a whole host of nonacupuncture studies that have tied deactivation in functional MRI to actual neuronal deactivation and decreases in signaling. A lot of this work is, of course, ongoing. Functional brain imaging has not been around anywhere nearly as long as electrophysiology has, but I think it's unfair to say that neuroimaging cannot measure decreases in brain activity.

Dr. Langevin: Perhaps we could ask one question of the panel. If you had your wish, if you could have the answer to one question besides "does acupuncture work or not?" what would that question be?

Ms. Schnyer: I guess it depends on how you define whether it works or it doesn't work. For me, the question is: "How is it that acupuncture can produce such a transformative effect as described by patients and as observed by practitioners? What exactly is happening that acupuncture can produce this transformative effect?"

Dr. Lao: I am interested in the long-term effect of acupuncture. From a practical point of view I see many patients who not only show temporary pain relief after treatment, but also show progressive recovery after a period of time. But we don't know whether that is because of how specific the treatment itself is or because of all of the treatment factors taken together. So I think there is a healing process. In the study we conducted recently with Dr. Berman in osteoarthritis, we found that at follow-up with patients at 4 week or 8 weeks, there was no difference between treatment and sham control. But we did show a difference after 14 weeks and 26 weeks. So there may be some underlying healing process.

Dr. Langevin: So you [would] like to know what the healing process is—that's great. Thank you. I would like to know: "How does *qi* move? How do you make it move? How does it get blocked? And what makes it not move?" I'm not even asking what *qi* is. I just want to know how it moves.

Dr. Hammerschlag: My question would be: "How does such a tiny physiological nudge, such a tiny physiological action, cause such large change?" It's similar to what Rosa asked, but I think she's almost at the whole-person level, and I'm still at the level of physiology. What is the system that acupuncture is activating so we can have that huge ramp-up in effect from such a small nudge?

Dr. Sherman: I want to know what to be able to tell people with confidence when they call me up on the phone in my Center and say: "OK, I'd like to try acupuncture. Do you think it will help me, and whom should I see?" Then I'd also like to be able to predict who is going to respond and who isn't.

Dr. Napadow: To me it's important to realize that acupuncture is not a "religion." It has this rich history, dating back thousands of years. You can look back at canonical texts that were written hundreds, thousands of years ago. But I don't think these books should be read as a "Bible." I think acupuncture has always been changing and growing throughout history, and I don't see why it should stop now. I don't think we should just study acupuncture as this relic of the past. I think we should take what we've learned about its mechanism and see if we can even, dare I say it, augment it. Can we use modern scientific methods to take some of the efficacy of acupuncture and maybe even increase it?

Dr. Langevin: Even though all of us here on this panel work hard at trying to answer practical questions and to design trials, the common theme was that we want to understand better some of the deep philosophical aspects—the meaning—of acupuncture.

Kevin W Chen, Ph.D., M.P.H., (University of Medicine and Dentistry, New Jersey): Actually I may answer your question from my perspective as well. My question is: "Is there any study in acupuncture with control of intention?"

Dr. Hammerschlag: When I've tried to read about this phenomenon, there's a lot about it in English in the martial arts literature but very little that I can find in the acupuncture literature. The challenge is how to include practitioner intention as a variable in clinical trial design.

Dr. Chen: The rule of *yi* or intention may not have been studied in any acupuncture trials. Almost all Traditional Chinese Medicine students have realized that the same acupuncture points, the same technique being used by their teachers or by themselves, may produce different healing results. So they go ask their teachers: "Why can you produce that result, but I cannot?" The teachers respond: "What [are] you thinking when you're doing it?" "I'm just thinking of putting the needle into the right point." "No, you have to have the right intention with a visualization process. . . . "

They don't teach that in acupuncture school any more, especially with a 150-hour M.D. program. I teach *qigong*. Some of my students are acupuncturists. Once they learn the basic principle of applying *yi* or intention in their practices, their clinical results go up quickly. Sometimes in extreme cases, the good acupuncturists don't need to use real needles. With visualization and imagery, they produce the same therapeutic results as other students who use needles. What they do is just imagine putting the needle in the point, and they induce the *qi* at a blockage. It's

very effective for pain relief. We have students who can use simple visualization and reduce the pain or completely eliminate the pain during *qigong* practice. The *yi* involves visualizing how the *qi* goes, where the stuck *qi* will go, where the good *qi* will be supplemented. That process can be achieved through *qigong* training, not just by acupuncture training.

Do we have any trials where you introduce the variable of intention as a scale or index to determine whether the therapist has the ability of visualization to make the pictures and change the pictures? If we can measure that, it will be very highly correlated with the result of clinical effect. Do you have any comments on this?

Dr. Lao: As a researcher, I want to bring it back to the basis of research. All the practitioners have intention, whether they are practicing *qigong*, *t'ui na*, or acupuncture. So we have to answer the question: "How important is the intention itself on top of the intervention method?" How can we separate them? If the intention is very important, then the acupuncture needle may be less important. So we have to think of a more scientific basis to distinguish the physiologic component and the psychologic component. But I agree that for treatment, intention is an important part of what we provide to the patient.

Dr. Langevin: I agree with what you're saying, but it would not preclude specific needle effects. There could still be some important effects via the needle. If you can do it another way, without a needle, then effects also possibly could happen. But the needle effects may still be important if that's what you happen to be using.

Dr. Lao: But as a mechanism it would be different.

Dr. Chen: The needle is a carrier for the healing intention.

Dr. Langevin: Thank you all for your participation. Today's presentations and discussion have illustrated very eloquently how much the field of acupuncture research has accomplished, and how much it has left to go.

For further discussion of these important issues, please join us next year on November 9–11, 2007 as the Society for Acupuncture Research (SAR) will hold a special conference entitled: **The Status and Future of Acupuncture Research: 10 Years Post-NIH Conference. For details, please visit the SAR Web site at: www.acupunctureresearch.org**

Address reprint requests to:
Karen Ballen
Mary Ann Liebert, Inc.
140 Hugenot Street, 3rd Floor
New Rochelle, NY 10801

E-mail: kballen@liebertpub.com