

Matching acupuncture clinical study designs to research questions

**Karen J. Sherman, Lixing Lao,
Hugh MacPherson, George Lewith,
Val Hopwood, Kenji Kawakita**

DEFINING THE RESEARCH QUESTION

Karen J. Sherman,
Center for Health Studies,
Group Health Cooperative,
Seattle WA 98101, USA.

Lixing Lao,
Complementary Medicine
Program, University of
Maryland School of
Medicine, Baltimore,
Maryland 21207-6697,
USA.

Hugh MacPherson,
Foundation for Traditional
Chinese Medicine, York
YO24 4EY, UK.

George Lewith,
University of Southampton
School of Medicine,
Southampton SO16 6YD,
UK.

Val Hopwood, University
of Southampton School of
Medicine, Southampton
SO16 6YD, UK.

Kenji Kawakita,
Department of Physiology,
Meiji University of Oriental
Medicine, Hiyoshi, Funai,
Kyoyo 629-0392, Japan.

Correspondence to:
Karen J. Sherman, PhD,
MPH, Assistant Scientific
Investigator, Center for
Health Studies, Group
Health Cooperative, 1730
Minor Ave Suite 1600,
Seattle WA 98101.
Tel: (206) 287-2426;
Fax: (206) 287-2871;
ksherman@
u.washington.edu

Ensuring that study designs for clinical research in acupuncture are appropriate for answering the research question of interest is a fundamental concern when beginning to plan a trial. A framework for considering this broad issue can be embodied in the question: 'How does **acupuncture** compare with **control group X** for changing **outcome Y** for **condition Z**?' The bolded words are place-holders for a variety of terms that allow us to specify a number of design features crucial for conducting any type of trial.

This question can be defined more precisely by considering the example of how acupuncture given bi-weekly for 12 weeks, using a fixed point prescription compares to needling in non-channel, non-point locations for pain relief in patients with fibromyalgia (as defined by a tender point examination). Acupuncture is defined as the use of a specific point prescription given bi-weekly for 12 weeks. The control group is an invasive sham treatment where needles are placed in 'non-channel, non-point' locations. The outcome would be pain and a specific pain measure (e.g., global rating or visual analogue scale) would be selected to measure changes at several time points. Finally, the condition in this example is fibromyalgia as defined for research by Wolfe et al.¹ Thus, the central question is whether a particular acupoint prescription performs better than another prescription that is unlikely to be helpful for fibromyalgia. Such a study

resembles the more traditional 'explanatory' clinical trial, which is usually designed to determine an intervention's effectiveness compared to placebo and is commonly used for testing new medications.²

Yet another reasonable research question might be to ask whether acupuncture, as performed by competent acupuncturists over a period of up to three months, is more likely to relieve pain than the current conventional treatments for patients with fibromyalgia. In this case, the treatment could be anything within an acupuncturist's scope of practice that would be appropriate for fibromyalgia or it could be more predetermined. The number and frequency of treatments might be left up to the practitioner or the patient. The control would be an active treatment and the purpose of the study would be to evaluate how acupuncture, as it is actually practiced, compares with another treatment, i.e. the relative efficacy of acupuncture or of a specific acupuncture treatment. Such a study is often termed a pragmatic trial and is useful for answering questions with a more practical orientation.²

DESIGNING THE TRIAL TO MATCH THE QUESTION

The term acupuncture can refer to different styles or systems of acupuncture (e.g. Traditional Chinese Medicine or Toyo Hari), to different doses (e.g. bi-weekly for 12 weeks), and to different practitioners (e.g. those who have trained traditionally and practiced at least 10 years versus those who use tender

point or segmental treatments). Treatment protocols can differ widely in acupuncture studies, depending on the condition under investigation and the skill of the acupuncturist. Standardized treatment protocols tend to be used in studies that are more analogous to trials for testing new medications. While such protocols are not always appropriate for evaluating acupuncture, they may be reasonable options under certain circumstances.^{3,4} Whatever treatment is used in a clinical trial, the rationale and the exact treatment used must be clearly described within the method section.

The term 'control group' refers to an appropriate comparison, which may include no treatment (or a wait-list control), standard care, a non-invasive or non-penetrating sham, an invasive or penetrating sham (e.g. 'misplaced needling') or a combination of treatments.⁵ The inference available from each comparison will differ substantially. Comparisons with no-treatments or wait-list controls factor out the effect of the natural history of the condition under study, while comparisons with standard care or another treatment evaluate the relative merits of two packages of treatment, i.e. effectiveness. Comparisons with non-invasive/non-penetrating or invasive/penetrating sham treatments begin to tease apart some of the 'non-specific' and the 'specific' effects of treatment, with the precise type of control determining exactly what aspect of the specific effects is being evaluated.

The assessment of outcomes evaluates whether and how much the patient has changed with respect to various health measures. Outcome measures (either subjective, objective, or both) should include those that are typically recommended for clinical trials of the condition regardless of the intervention (e.g. see Bombardier⁶ for recommendations on outcomes for lumbar pain). Although other outcomes can be included that may reflect a more global or holistic approach to health, the use of primary (and secondary) outcomes that are identical to those in other clinical trials will facilitate comparisons of effect sizes and improve the quality of conclusions from systematic reviews.

Typically, the condition under study will be a biomedically defined condition. From a practical perspective, it is critical to test a condition where either practitioners see positive results with patients in their practice or practitioners find that a large proportion of the patients they treat present with the condition (e.g. low back pain). It is important to specify the condition under study carefully and in some cases by degree of severity (e.g. mild depression vs major depression) and what aspect of the condition is being treated.

Because acupuncture may be more effective for some groups of conditions than others, it may be

helpful to think in terms of broader categories, such as chronic vs acute conditions, pain vs non-pain, musculoskeletal vs. internal, or self-regulation vs symptom control or as a general disease category (e.g., osteoarthritis). For example, Berman et al.⁷ have shown acupuncture to be helpful for the relief of osteoarthritis of the knee. Based on these results, it may be reasonable to hypothesize that acupuncture can relieve pain of osteoarthritis of the shoulder or hip.

However, such assumptions may not always be accurate. For example, acupuncture appears useful for cocaine addiction,⁸ but apparently has no effect on smoking cessation⁹ although both conditions are considered 'addictions'. Such inconsistencies may be due to differences in study design, the relative effectiveness of the treatment protocols (e.g. daily treatment in a cocaine study vs weekly treatment in a smoking study), the severity of the patients' addictions in the two studies (e.g. casual cocaine users vs. inveterate smokers), or the underlying physiological/psychological mechanisms of the addiction process.

For acupuncture, the most variability is likely to come from the plethora of treatment and control options available. These issues are more constrained in studies evaluating conventional medications. Such options permit a range of study questions that vary between narrowly focused 'explanatory' trials (e.g. what is the specific effect of needling point X compared to point Y) to broadly defined 'pragmatic trials' (e.g. how does acupuncture treatment compare with standard care). The issue of control groups, which are particularly challenging in acupuncture studies, are considered in greater detail in the accompanying article on non-specific effects of acupuncture.

PHASING THE RESEARCH PROCESS

It is recommended that research follows a multi-step approach of increasing complexity, as outlined below, that broadly mimics the phasing of drug trials for US regulatory approval:

1. Review clinical literature
2. Consult practitioners and examine practitioner records
3. Choose outcome measures
4. Design/perform small, uncontrolled trial ('Phase I')
5. Design/perform small RCT ('Phase II')
6. Test feasibility of control (e.g. credibility)
7. Design/perform full-scale 'explanatory' trial ('Phase III')

8. Design/perform 'pragmatic' trial
9. Design/perform follow-on studies.

By using this process, resources will be utilized in a more cost-effective manner than if full-scale trials, or even pilot trials, were begun immediately, without a firm foundation justifying the treatment and control. Typically studies in steps 4, 5 and 7 use protocols that are narrowly focused so that the research can be replicated and the findings can be generalized. However, when the appropriate treatments for a given biomedical condition differ enormously or are dependent on the East Asian medical diagnosis, it can be argued that a more pragmatically oriented trial should be conducted prior to an efficacy trial. Even if that were done, we recommend preliminary work (steps 1 through 5) be performed. If the pragmatic trial demonstrates a positive result, the more constrained explanatory trial can then be used to identify whether there was a specific effect attributable to one or more components of the treatment or if the entire treatment was needed to create an effective response.

There are several other instances where it may be of value to consider a pragmatic trial prior to the more conventional explanatory trials. Since many studies of acupuncture involve individualizing treatment as well as the use of ancillary modalities such as moxibustion, cupping, or herbs, an argument can be made for performing an appropriately powered pragmatic trial that explicitly permits variability of treatment prior to testing a more constrained intervention. Pragmatic trials may also be more appropriately performed first when the question of interest is whether the incorporation of a service into the health care system would be cost-effective and safe.¹⁰ In these circumstances, step 5 may directly lead to step 8 and steps 6 and 7 would follow if the pragmatic trial were positive. For the purposes of this report, we discuss pragmatic trials after the Phase III trials.

Literature review and consultation with practitioners

One of the most important questions for any study is the appropriateness of the treatment protocol. A firm basis for a treatment protocol should exist, rather than just the experience of a single practitioner. We recommend a review of the clinical literature supplemented by a study of practitioner records for the condition to document the variability in treatment approaches and to offer guidelines for the selection of treatments. The use of more comprehensive approaches, including expert practitioner panels, is also recommended.^{4,11}

Selection of outcome measures

Outcome measures should be chosen after a review of the literature and prior to the commencement of any patient treatment. Consultation with experts in the condition under study is recommended. Although acupuncture is a relatively safe procedure,¹² information on adverse reactions should be collected. A standardized, systematic questionnaire combined with open-ended questions is recommended for this process.^{13,14}

The 'Phase I' trial

The first clinical trial should be a relatively small, uncontrolled open trial with only one arm, in which patients are treated according to a protocol based on steps 1 and 2. Such studies are typically designed to examine dosage and safety. Outcomes are measured before, during and after treatment to ascertain how long the treatment should be. The reports by Berman et al.¹⁵ and MacPherson et al.¹⁶ present examples of Phase I trials. If no changes are seen in the outcomes after treatment, the project should be reassessed. If changes in outcomes appear to be of potential clinical importance additional studies are warranted.

The 'Phase II' trial

Assuming that the results are promising, the second trial would be a small, randomized controlled trial, including somewhere between 30 and 50 people per group. It would normally compare the treated individuals to a group receiving no treatment. Alternatively, if the treatment were an addition to standard care, the intervention group would receive standard care plus acupuncture and the control group would receive standard care alone. Besides providing limited data on whether the treatment is superior to no treatment, such studies collect important information needed for calculating sample size for a more definitive trial. A pilot study of acupuncture to treat osteoarthritis⁷ is an example of a Phase II trial.

The 'Phase III' trial

For a full-scale RCT, a sham control is typically used and the credibility of that control procedure would need to be pre-tested using acupuncture-naïve patients with the condition.¹⁷ The requirement for acupuncture-naïve controls may be relaxed when using invasive sham controls, depending on whether an appropriate control treatment could be created that would appear plausible to individuals who have some experience with acupuncture. Sometimes several control groups may be appropriate.

Pragmatic trials

Since acupuncture is usually practised in an individualized manner, a pragmatic trial evaluating acupuncture as it is commonly practised might follow a more constrained trial. This would typically allow a more open and individualized protocol, and might well involve more than needle acupuncture.¹⁸ Thomas¹⁹ and Vickers²⁰ describe the rationale and design of two large pragmatic trials currently underway in the UK. If a pragmatic trial is successful, additional studies trying to optimize the treatment would then be appropriate, regardless of whether the pragmatic trial was conducted prior or subsequent to a Phase III explanatory trial.

REFERENCES

1. Wolfe F., Smythe H.A., Yunus M.B., Bennett R.M., Bombardier C., Goldenberg D.L., Tugwell P., Campbell S.M., Abeles M., Clark P. et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum* 1990;33:160–172.
2. Jadad A.R. *Randomised controlled trials*. London: BMJ Publishing Group, 1998.
3. Lao L., Ezzo J., Berman B.M., Hammerschlag R. Assessing clinical efficacy of acupuncture: Considerations for designing future acupuncture trials, pp 187–210. In: Stux G., Hammerschlag R. eds. *Clinical Acupuncture: Scientific basis*. Berlin: Springer, 2001.
4. Sherman K.J., Hogeboom C.J., Cherkin D.C. How traditional Chinese medicine acupuncturists would diagnose and treat chronic low back pain: results of a survey of licensed acupuncturists in Washington state. *Complement Ther Med* 2001;9:146–153.
5. Hammerschlag R. Methodological and ethical issues in clinical trials of acupuncture. *J Altern Complement Med* 1998;4:159–171.
6. Bombardier C. Outcome assessments in the evaluation treatment of spinal disorders: summary and general recommendations. *Spine* 2000;25(24):3100–3103.
7. Berman B.M., Singh B.B., Lao L., Langenberg P., Li H., Hadhazy V., Baretta J., Hochberg M. A randomized trial of acupuncture as an adjunctive therapy in osteoarthritis of the knee. *Rheumatol (Oxford)* 1999;38:346–354.
8. Avants S.K., Margolin A., Holford T.R., Kosten T.R. A randomized controlled trial of auricular acupuncture for cocaine dependence. *Arch Intern Med* 2000;160(15):2305–2312.
9. NIH Consensus Development Panel on Acupuncture. Acupuncture. *JAMA* 1998;280:1518–1524.
10. MacPherson H. Out of the laboratory and into the clinic: acupuncture research in the real world. *Clin Acupunct Oriental Med* 2000;1:97–100.
11. Sherman K.J., Cherkin D.C., Hogeboom C.J. The diagnosis and treatment of patients with chronic low back pain by Traditional Chinese Medical acupuncturists. *J Altern Complement Med* 2001;7:641–650.
12. Ernst E., White A.R. Prospective studies of the safety of acupuncture: a systematic review. *Am J Med* 2001;110:481–485.
13. Lao L., Bergman S., Hamilton G., Langenberg P., Berman B. Evaluation of acupuncture for pain control after oral surgery. *Arch Otolaryngol: Head and Neck Surg* 1999;125(5):567–572.
14. List T., Helkimo M. Adverse events of acupuncture and occlusal splint therapy in the treatment of craniomandibular disorders. *J Craniomandibular Practice* 1992;10(4):318–326.
15. Berman B.M., Lao L., Greene M., Anderson R.W., Wong R.H., Langenberg P., Hochberg M.C. Efficacy of traditional Chinese acupuncture in the treatment of symptomatic knee osteoarthritis: a pilot study. *Osteoarthritis Cartilage* 1995;3:139–142.
16. MacPherson H., Gould A.J., Fitter M. Acupuncture for low back pain: results of a pilot study for a randomized controlled trial. *Complement Ther Med* 1999;7(2):83–90.
17. Sherman K.J., Cherkin D.C., Hogeboom C.J. Developing protocols for acupuncture studies: an efficacy trial of low back pain. *Altern Ther* 2001;7:111.
18. Cherkin D.C., Eisenberg D., Sherman K.J., Barlow W., Kaptchuk T.J., Street J., Deyo R.A. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Arch Intern Med* 2001;161(8):1081–1088.
19. Thomas K.J., Fitter M., Brazier J., MacPherson H., Campbell M., Nicholl J.P., Roman M. Longer-term clinical and economic benefits of offering acupuncture to patients with chronic low back pain assessed as suitable for primary care management. *Complement Ther Med* 1999;7:91–100.
20. Vickers A., Rees R., Zollman C., Smith C., Ellis N. Acupuncture for migraine and headache in primary care: a protocol for a pragmatic, randomized trial. *Complement Ther Med* 1999;7:3–18.