Complementary Therapies in Medicine (1997) **5**, 8–12

Issues to consider in determining an adequate treatment in a clinical trial of acupuncture

S. Birch

SUMMARY. Many clinical trials of acupuncture have administered inadequate treatments, or have inadequately described the treatments used, making it difficult to assess what has actually been tested, and to interpret study results. In this article, issues such as the generalizability of results from a study, what constitutes enough treatment in a study, whether enough treatment was administered in a study, and problems with source documentation of the tested treatment are examined.

INTRODUCTION: EXPLORING THE PROBLEM

Several authors have discussed the issues and difficulties involved in performing clinical trials of acupuncture.[1-7] Significant contributions have already been made to this topic both by reviewers and researchers. While many clinical trials of acupuncture have been performed, and a number of reviews written, not much attention has been given yet to the issues involved in selecting and administering the treatments in clinical trials of acupuncture. A clinical trial will be a fair test of acupuncture only if an adequate treatment is administered. However, little effort has been made to date to ensure that adequate treatment is administered in clinical trials of acupuncture. This paper will focus on this specific topic as it may be important for future studies and further interpretations of published studies. This discussion is relevant both for selection of the 'true' or 'active' treatments and the 'sham', 'control' or 'less active' treatments in controlled studies.

A recently published trial highlights some of the problems. Wyon et al tested the efficacy of acupuncture in the treatment of menopausal hot flashes.[8] This study found significant improvements in both treatment groups without much difference between their test treatment (needling points with *de qi* and additional low-frequency electrical stimulation) and their control treatment (shallow needling of the same points with no *de qi* or electrical stimulation). It is an important study because of its novel use of the measurement of urinary neuropeptide excretions by radioimmunoassay methods as an outcome, and because it is the first published controlled clinical trial of acupuncture for menopausal hot flashes.[8] However, in a review of the clinical literature describing the treatment of menopausal hot flashes, a problem was found relative to this study. In four texts that describe the use of techniques similar to the study's test treatment,[9-12] few of the nine points treated in the study were recommended for the treatment of menopausal hot flashes (1–3 points), while, in three texts that describe the use of techniques similar to the study's control treatment,[13-15] more of the points treated in the study were recommended for the treatment of menopausal hot flashes (3–6 points). In other words, the literature itself would suggest that, what Wyon et al considered as a control treatment, was possibly a more effective treatment for the condition for which it was supposed to be ineffective. It is therefore not surprising that they found little difference in treatment effect between the two treatment groups. The authors do not cite a primary source for the origins of their selected treatment, instead referring to a study using the same points for dysmenorrhea. Further, while they have data supporting the possible use of shallow needling as a control in some pain studies, they do not have specific data supporting the use of this technique at the treatment points as a control treatment. While an excellent study in other regards, it does exemplify problems commonly encountered in clinical trials of acupuncture. Inadequate treatment is applied, inappropriate needle controls are used, treatment is rarely rooted in the clinical literature, or a combination of these.

Examples of these problems can be found in many studies. Coan et al conducted a study in the treatment of low back pain,[16] but did not describe the treatments administered.[4] They also did not cite sources

Stephen Birch LicAc, Doctoral Candidate, University of Exeter, Centre for Complementary Health Studies, Exeter, UK.

Correspondence to the author, CRU, 914 1/2 Howard Avenue, The APT Foundation, New Haven, CT 06519, USA.

describing where the treatments came from,[17] instead stating that treatments were based on the experience of the practitioners. Not describing the treatment makes the study unreproducible. This factor, and basing the treatment only on the experience of the practitioner, also diminishes generalizability of the results. In a review of 18 chronic back and neck pain studies, 67% were found to not cite the sources of the tested treatments.[17] Obviously, the fact that a study cites no sources does not itself carry critical weight, but when sources are not cited and clearly inadequate treatments are administered, it suggests a plausible reason for that inadequacy. Of these 18 studies, a study conducted by Emery and Lythgoe needled only two points for chronic back pain,[18] while one conducted by Laitinen also needled only two points in the treatment of chronic neck pain.[19] The same review found that considerably more points need to be treated for chronic pain conditions and that, of the 18 studies, 94–100% may have treated an inadequate number of points. The determination of what constituted an 'adequate' number of points was based on a review of a number of acupuncture texts and papers that described the treatment of back and/or neck pain. In the review, 18 sources were examined for back pain, yielding an average range of points treated (9–13.8); 12 sources were similarly examined for neck pain, yielding an average range of points treated (10.5–13.8). The lower number of each range was taken as the 'minimum number' of points necessary and was assumed to represent an 'adequate' treatment. Of the 18 reviewed studies, one by Fox and Melzack applied only two treatment sessions,[20] while the study by Emery and Lythgoe applied only three treatment sessions for chronic back pain.[18] The same review found that considerably more treatment sessions need to be administered in the treatment of chronic pain, and that, of the 18, 56–94% may have administered an inadequate number of treatment sessions.[17] In this review, the same sources were examined to determine the minimum number of treatment sessions recommended for the condition being treated; this number was then assumed to represent an 'adequate' treatment. The Emery, Laitinen and Fox studies are important and were included in ter Riet et al's 1990 meta-analysis, but with no analysis of these inadequacies.[4]

The main problem in Wyon et al's study appears to have been the selection of appropriate 'active' and 'control' treatments. The selection of treatment technique is very important when designing studies with a needle-control group. Ideally, the control treatment would have no therapeutic effect beyond placebo, a condition very difficult to attain. Short of this, the control treatment should have minimal effects beyond placebo.[5] Having an appropriate control needling technique is also important when dealing with the confounding factor of diffuse noxious inhibitory control,[4] which needs to be allowed for in a study that is to test the specificity of acupuncture points.[21] An inappropriate needle-control technique can create difficulty in the interpretation of results, as we have seen in the Wyon et al study. In the same review of 18 back and neck pain studies,[17] six used a needle-control condition.[18,22–26] Of these, 83% have problems with the appropriateness of the needle-control techniques, 50% using electroacupuncture techniques as part of the control condition.[22–24] Further, 50% of these six studies also equated their control needling with placebo.[22,25,26]

The general question of what would constitute an adequate or appropriate treatment in a clinical trial of acupuncture is deceptively difficult to answer. An obvious solution to this question would be simply to look up the relevant treatment in an appropriate source and apply that protocol. This, however, is fraught with problems:

Sources. Which sources would one use? Why one source and not another? Is there empirical evidence to support what the sources describe? If not, what can one rely on instead? These issues are compounded by the fact that most researchers do not cite appropriate sources for the treatment they tested.

Generalizability. How do we know if the effects of treatment based upon what that source describes can be generalized to what other sources describe or even sources in the same tradition?

Adequacy. Are enough points being treated? Are appropriate techniques being applied? Are enough treatments being administered? Is a treatment adequate if it applies a standardized treatment rather than a variable treatment, achieved by adjusting treatment according to some traditionally based 'individualized' assessment method? Is the use of acupuncture alone an adequate treatment, or must it be combined with some other therapy such as herbal therapy, as part of a larger discipline, for example traditional Chinese medicine?

While some of these issues have been touched on by other authors,[16] they have never been systematically examined before. In this paper, the author summarizes current work exploring these issues.

THE PRIMARY ISSUES

Sources

Acupuncture is a culturally diverse system of medicine. There are many explanatory models of what it is and how it works, many of which are contradictory.[27–29] While many acupuncture researchers tend to test what they believe to be *the* correct treatment, this assumption is almost never based on an extensive review of the literature. This weakens the generalizability of results from a study, and makes the foundation of the

test treatment dependent upon untested assumptions: that the selected test treatment is the best or an appropriate one, and that it is representative of a model of practice or the general practice of acupuncture. Further, when a control needle condition is used in a study, it is dependent on the assumption that the control treatment has no effect beyond placebo, or only minimal effects. Experimental evidence from pilot studies is rarely available as documentation for selection of the test or control treatment in acupuncture studies.

Referencing of sources as support for the selection of the test treatment is complicated by a number of issues. Research in orthodox medicine tends to rely on sources of knowledge that are supported by empirical evidence, ideally from clinical trials. However, unlike much of the orthodox medical literature, textbooks on acupuncture are not based on empirical evidence from clinical trials. Thus, acupuncture researchers should not expect to use the existence of empirical evidence as a way of validating the sources they choose to use. Other criteria are necessary to help establish the validity of textbooks that are to be utilized. One must be sure that the sources of information upon which the test treatment are based are reliable and accurate. Here, too, a number of issues emerge. Within the same school of practice, where the same methods are supposedly applied, there can be wide variation. A review of a number of written sources each claiming to describe the differential treatment of low back pain according to 'traditional Chinese medicine' reveals a wide variety in diagnosis and treatment selection. Two sources typify this variety. Kwok Chi Yau[30] describes seven patterns associated with low back pain, while Liu Yan-chi describes four patterns,[11] two of which overlap with the seven of the first source, and with little overlap in point selection. It is also important to determine whether what one author describes can be taken as accurately representing the system they purport to describe. Here, significant scholarly and philological issues exist. Unschuld and Wiseman have perhaps the best discussions of these issues.[31–32] Authors of texts often inadequately cite the sources they use and leave open the question of how accurately they have translated those sources.[33–34] These problems are compounded by intraschool and interschool text variability (explanatory models, treatment techniques and treatment-selection methods). While some of these issues lie outside the domain of the clinical researcher, requiring other scholarly skills, they must ultimately be tackled if the generalizability of results from clinical trials of acupuncture is to be improved, and if researchers are to develop confidence in the literature on which their treatments are based. A recent literature survey found few reliable English-language texts,[35] especially prior to the early 1980s, and this has probably been another confounding factor in the source documentation of studies.

Generalizability

Usually, in clinical trials the term generalizability refers to the ability to generalize results to a particular patient population. In this paper, the term is used instead to refer to the ability to generalize the results of a study to either the whole field of acupuncture or to a particular school of thought within the field. The problem of generalizability of results in clinical trials of acupuncture has been virtually ignored and is significant. It is not only a matter of determining if the results of method X can be assumed to represent what would be found when applying methods Y or Z. Lack of source documentation undermines the generalizability of studies. Issues relating to the accuracy and validity of cited sources also impact on the generalizability of results. Techniques need to be developed to help increase the generalizability of results. One solution may be to look for congruence between sources among many sources.[17]

Adequacy of treatment

Establishing the adequacy of treatment is not simply a matter of determining a 'standard' treatment, number of points or treatment sessions for each condition, rather of deciding some minimum that would be more appropriate. Naturally, there is variability in treatment criteria: what points are recommended, how many points, what techniques are applied, and how many treatment sessions should be given. For a test treatment to be accepted as a fair test of the condition it is being applied for, it is essential that basic guidelines be established that are generally agreed upon. The author proposes that developing accepted minima for each of these criteria is a viable solution. This process is outlined above. Reviewing multiple sources that describe treatments of the condition being treated in the clinical trial yields ranges of average numbers of points and treatment sessions recommended. The 'minimal' treatment is the lower end of these average ranges. In terms of the numbers of points treated, of the 18 back-pain sources reviewed, 61% or more agreed with the minimal number established by this method, while 67% or more of the 12 neck-pain sources agreed with the minimal number so established. While not including the recommendations of all sources consulted, this method does include the majority of sources. Any treatment that does not meet the minimum so established should be tested in pilot studies before being used in larger clinical trials.[17] Further, if the process of developing these guidelines for minimal treatment involves reviewing a broad range of acupuncture literature, this can also improve generalizability of the study results.[17] There will, however, be exceptions to these guidelines. The use of a single point as an antiemetic treatment does not match what the acupuncture literature typically recommends for emesis, and has been criticized as inadequate.[36] However,

there are now substantial experimental data supporting the antiemetic effects of a single point.[37] This evidence justifies the selection of the point for future studies, but it also illustrates difficulties and exceptions in setting criteria for the selection of test treatments, perhaps stressing the importance of using pilot studies to address these issues.

Some authors have insisted that applying a standardized or 'formulary' treatment is not a real or adequate application of acupuncture; rather, one must apply some individualized method based upon a traditional explanatory model.[16,36,38–39] While this is a commonly held opinion among proponents of particular traditional approaches, it is, however, not a necessity dictated by historical or modern evidence.[40] It may be important to tailor treatment according to some model of practice when testing that model of practice, but this then raises the not uncomplicated issues of textual inconsistencies within particular models of practice, and issues of test–retest and inter-rater agreements in those decision-making processes.

Some proponents of traditional Chinese medicine have argued that acupuncture, as an integral part of this larger medicine, cannot be separated from the rest of the medicine (especially herbal medicine) and that therefore to test acupuncture alone might be inadequate. However, such an opinion is not supportable by the large body of historical literature where acupuncture and herbal medicine, for example, were routinely practiced separately. Nor is it reflective of the general practice of acupuncture where, in countries such as Japan, acupuncture is almost always practiced separately from other traditional medicines of Chinese origin.[41] Further, the current situation of training and practice in China appears to be different from what is claimed. Ergil discusses how, since the mid 1980s, acupuncture and herbal medicine have been taught as specialized medical systems in mainland China.[42–43] Acupuncture may be combined with other traditional therapies, but it is not a necessity for all acupuncture studies, nor is it clear that it is a necessity within the broad field of Chinese traditional medicine.

There are many problems that need to be carefully addressed, and many issues in need of greater scholarship and attention. Designing studies of acupuncture is complicated not only because of difficulties fitting a hands-on therapy into a pharmaceutical research model, but also because of insufficient knowledge of and thus consideration of the broad range of acupuncture. The majority of published clinical trials have been plagued by this latter problem. It has undermined both the selection of the 'test' or 'active' treatments and the 'control' or 'less active' treatments. Acupuncture is not a homogeneous therapy and, without adequate consideration of its complex nature, the unwary researcher can fall into the trap of performing inadequate therapy that is inadequately informed and inappropriately controlled.

ACKNOWLEDGEMENTS

The author wishes to thank Art Margolin, Kelly Avants and Simon Mills for their critical feedback on the manuscript.

REFERENCES

1. Ernst E. Acupuncture research: where are the problems? Acup Med 1994; 12(2): 93–97.
2. McLellan TA, Grossman DS, Blaine JD, Haverkos HW. Acupuncture treatment for drug abuse: a technical review. J Subst Abuse Treat 1993; 10: 569–576.
3. Patel M, Gutzwiller F, Paccaud F, Marazzi A. A meta-analysis of acupuncture for chronic pain. Int J Epidem 1989; 18(4): 900–906.
4. ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. J Clin Epidem 1990; 43(11): 1191–1199.
5. Vincent CA. Acupuncture as a treatment for chronic pain. In: Lewith GT, Aldridge D, eds. Clinical research methodology for complementary therapies. London: Hodder and Stoughton, 1993; pp 289–308.
6. Vincent CA, Lewith GT. From Beijing to Belgrade: making sense of acupuncture research. APS J 1994; 3(2): 89–91.
7. Vincent CA, Richardson PH. The evaluation of therapeutic acupuncture: concepts and methods. Pain 1986; 24: 1–13.
8. Wyon Y, Lindgren R, Lundeberg T, Hammar M. Effects of acupuncture on climacteric vasomotor symptoms, quality of life, and urinary excretion of neuropeptides among postmenopausal women. Menopause: J North Am Menop Soc 1995; 2(1): 3–12.
9. Chen RJ, Wang N. Acupuncture case histories from China. Seattle: Eastland Press, 1988.
10. Soulie De Morant G. Chinese acupuncture (L'acuponcture Chinoise). Brookline, MA: Paradigm, 1994.
11. Liu YC. The essential book of traditional Chinese medicine, Vol. 2. New York: Columbia University Press, 1988.
12. Tianjin Chinese Medical College. Shinkyu rinsho no riron to jisai. Translated by S Ikegami. Tokyo: Kokusho Publishing Assoc, 1988.
13. Manaka Y. Kyu to hari. Tokyo: Shufu no Tomosha, 1971.
14. Nagahama Y, Kinoshita H, Nakamura R. Shinkyu chiryo no shin kenkyu, 3rd edn. Osaka: Sogensha, 1983.
15. Shiroda B. Chikenrei woshu toshita; shinkyu chiryo no jisai, 12th edn. Vol. 2. Osaka: Sogen Sha, 1967.
16. Coan R, Wong G, Ku SL et al. The acupuncture treatment of low back pain: a randomized controlled treatment. Am J Chin Med 1980; (8) 2: 181–189.
17. Birch S. Testing the clinical specificity of needle sites in controlled clinical trials of acupuncture. I. The importance of validating the 'relevance' of 'true' or 'active' points and 'irrelevance' of 'control' or 'less-active' points – proposal for a justification method. Proc. 2nd Symposium of the Society for Acupuncture Research. Washington DC: Society for Acupuncture Research, 1995.
18. Emery P, Lythgoe S. The effect of acupuncture on ankylosing spondylitis. Br J Rheum 1986; 25(1): 132–133.
19. Laitinen J. Treatment of cervical syndrome by acupuncture. Scand J Rehab Med 1975; 7: 114–117.
20. Fox E, Melzack R. Transcutaneous stimulation and acupuncture: comparison of treatment for low back pain. Pain 1976; 2: 141–148.
21. Birch S. Testing the clinical specificity of needle sites in controlled clinical trials of acupuncture. II. The problem of 'diffuse noxious inhibitory control' and how to control for it. Proc. 2nd Symposium of the Society for Acupuncture Research. Washington DC: Society for Acupuncture Research, 1995.
22. Edelist G, Gross AE. Treatment of low back pain with acupuncture. Can Anaesth Soc J 1976; 23(3): 303–306.
23. Ghia JN, Mao W, Toomey TC, Gregg JM. Acupuncture and chronic pain mechanisms. Pain 1976; 2: 285–299.
24. Matsumoto T, Levy B, Ambruso V. Clinical evaluation of acupuncture. Am Surg 1974; 400–405.
25. Mendelson G, Selwood TS, Kranz HK, Loh TS, Kidson MA, Scott DS. Acupuncture treatment in chronic back pain: a double-blind placebo-controlled trial. Am J Med 1983; 74: 49–55.

26. Thomas M, Eriksson SV, Lundeberg T. A comparative study of Diazepam and acupuncture in patients with osteoarthritis pain: a placebo controlled study. Am J Chin Med 1991; (19) 2: 95–100.
27. Baldry PE. Acupuncture, trigger points and musculoskeletal pain. Edinburgh: Churchill Livingstone, 1989.
28. Fukushima K. Meridian therapy. Tokyo: Toyo Hari Medical Assoc, 1991.
29. So JTY. Treatment of disease with acupuncture. Brookline, MA: Paradigm, 1987.
30. Kwok CY, Wang PZ, Wong R, Yang LG, Zhang DZ. Applying mathematical models to traditional Chinese medicine for the diagnosis and treatment of low back pain. Am J Acup 1991; 19(2): 129–135.
31. *See for example*: Unschuld PU. Traditional Chinese medicine: some historical and epistemological reflections. Soc Sci Med 1987; 24(12): 1023–1029.
32. Wiseman N, Boss K. Introduction to the Glossary of Chinese medical terms and acupuncture points. Brookline, MA: Paradigm, 1990.
33. Birch S. Letter to the editor. Compl Ther Med 1995; 3(1): 57.
34. Birch S. Letter to the editor. Compl Ther Med 1995; 3(4): 259.
35. Birch S, Tsutani K. A bibliometrical study of English-language materials on acupuncture. Compl Ther Med 1996; 4: 172–177.
36. Maciocia G. Letter to the editor. Compl Ther Med 1993; 1(4): 221–222.
37. Dundee JW, McMillan C. Positive evidence for P6 acupuncture antiemesis. Postgrad Med J 1991; 67: 417–422.
38. Bensoussan A. Contemporary acupuncture research: the difficulties of research across scientific paradigms. Am J Acup 1991; 19(4): 357–365.
39. Diebschlag F. Placebo acupuncture. Eur J Orient Med 1993; 1(2): 12–17.
40. Birch S. Letter to the editor. J Alt Compl Med 1995; 1, 3: 221–222.
41. Birch S. Acupuncture in Japan: an introductory survey. 4 parts. Review. Part 1 1990; 6: 12–13; Part 2 1990; 7: 16–20; Part 3 1991; 8: 21–26; Part 4 1991; 9: 28–31, 39–42.
42. Ergil MC. Letter to the editor. J Acup Soc New York 1993; 1(2 and 3): 23–26.
43. Ergil MC. Medical education in China. CCAOM News 1994; 1, 1: 3–5.