

Placebo controls for acupuncture studies

Charles Vincent MPhil PhD¹ George Lewith MRCP MRCGP²

J R Soc Med 1995;88:199-202

Keywords: *acupuncture; placebo; credibility; clinical trials*

SUMMARY

Many studies of acupuncture treatment are seriously flawed by methodological problems¹⁻³. Poor design, inadequate measures and statistical analysis, lack of follow-up data and sub-standard treatment are all too common. However, the major problem, which many investigators consider to be still unresolved⁴, is the definition of an appropriate placebo control. The use of inappropriate placebo controls has bedeviled acupuncture research and led to serious misinterpretation of the results of clinical trials⁵. While a number of different solutions have been proposed there is, as yet, no agreed way of assessing the adequacy of control conditions or of deciding which placebo to use in a particular trial. We propose that assessing the credibility of treatments and control conditions may provide a way forward to a more rigorous, consensus approach.

SOURCES OF BIAS IN CONTROLLED TRIALS

There are various possible sources of bias in all controlled trials⁶. Two are particularly important in trials of acupuncture. First, trials of acupuncture have to be single blind. Inevitably, the clinician giving the acupuncture treatment is aware of which is the true treatment and which is the control, and may inadvertently communicate different expectations to the patients in the treatment and control groups. Any advantage shown by a true treatment may then be due to factors other than the specific effect of the needles.

A second difficulty with any control condition, particularly if it is of a different form from the true treatment, is that it may have a different psychological impact. Some trials of acupuncture have used mock transcutaneous nerve stimulation (mock TENS) in which electrodes are applied as usual but no current is passed. If mock TENS has a lesser psychological impact than acupuncture, then a significant difference between treatments might simply mean that acupuncture was the more powerful placebo. Whichever control condition is used, the psychological impact of the true treatment and the control need to be assessed if we are to be confident that the trial is not favouring either the real or control treatment.

Both these problems arise in controlled trials of acupuncture or any other skilled, physical treatment. They do not mean the trials are necessarily flawed or that the

methodology is unsatisfactory, but they do suggest that great care must be taken in the choice of placebo control.

PLACEBO CONTROL CONDITIONS USED IN ACUPUNCTURE TRIALS

A bewildering variety of control procedures have been used in acupuncture trials. In some all acupuncture procedures are matched with those in the true treatment group except that needles are not inserted; instead they are rubbed against the skin⁷ or glued to it⁸. These are not really credible, as even patients with no experience of acupuncture treatment are likely to know that needle insertion is involved.

In the most commonly used control treatment needling is carried out at theoretically irrelevant sites, away from the classical point locations. Depth of insertion and stimulation are the same; only location differs. This procedure, which is termed 'sham' acupuncture, has been used as a placebo in a great many studies⁹⁻¹¹. Sham acupuncture was initially assumed by most investigators to be ineffective, and therefore ideal as a placebo. However, in 1983 Lewith and Machin¹ pointed out that sham acupuncture appeared to have an analgesic effect in 40-50% of patients, in comparison with 60% for real acupuncture. Experimental work suggests that stimulation at many different sites, whether or not they be classical point locations, may produce analgesia, possibly via diffuse noxious inhibitory control (DNIC)¹²⁻¹⁶. Controlled trials have also shown significant therapeutic benefits from both classical and non-classical locations^{9,15,17}. It is now clear that sham acupuncture cannot be considered a placebo. Real versus sham acupuncture trials for pain therefore provide information only about the role of point location^{5,18}. If precise point location is not important there will be no

¹Academic Department of Psychiatry, St Mary's Hospital Medical School, Praed St, London W2 1NY and ²Centre for the Study of Complementary Medicine, 51 Bedford Place, Southampton, Hampshire SO1 2DG, UK

Correspondence to: Dr Vincent

difference between groups even if the true treatment does have specific effects.

The argument with respect to the treatment of non-painful conditions such as the use of P6 to treat nausea is different. Here the clinical trial evidence suggests that point location is important and that acupuncture away from P6 has little effect on nausea and is primarily a placebo. Real P6 acupuncture or acupressure shows a consistent 60–70% response rate, whereas sham acupuncture or acupressure only a 25–30% response rate, consistent with it being primarily a placebo^{19,20}. It is probable that in non-painful conditions the underlying physiological mechanism is different to that in pain and so sham acupuncture can then act as a valid placebo control treatment¹⁸. Nevertheless, in the interests of standardizing the evaluation of acupuncture, one of the options considered below might be preferable as a control condition.

SOLUTIONS TO THE PROBLEM OF THE ACUPUNCTURE CONTROL

The basic problem is to find a control condition with small or non-existent specific physiological effects.

Mock TENS

The first plausible solution was the introduction of mock transcutaneous nerve stimulation (TENS) as a control condition in acupuncture trials. In this procedure transcutaneous electrical nerve stimulators are used in the usual way, except that no current actually passes between the electrodes. Patients are sometimes told that they are receiving subliminal pulse therapy and they will therefore not feel the current. This control was developed as a placebo comparison in trials of TENS itself, but first used in 1983 in a trial of acupuncture by Macdonald²¹. Mock TENS has also been used in a number of trials of acupuncture including post herpetic neuralgia²² and migraine²³.

Minimal acupuncture

In minimal acupuncture^{24,25} needles are placed away from classical or trigger points, inserted only 1–2 mm and stimulated extremely lightly. This procedure minimizes the specific effects of the needling while maintaining its psychological impact: it can be almost exactly matched to the real treatment. Minimal acupuncture has been used as a control condition in several studies, though not necessarily described in this way. It is possible that minimal acupuncture might have some therapeutic effect but, even if there is a small effect, the trial is not invalidated; it is just slightly harder to demonstrate a difference between treatment and control.

Either minimal acupuncture or mock-TENS may be appropriate as a control, depending on whether it is likely to

seem plausible to the patients involved. The choice may depend on the condition being treated, the expectations of the patient and the nature of the real treatment. For instance if very light stimulation is being assessed, minimal acupuncture may be too close to the true treatment and mock TENS should be employed. In other situations it may be preferable to simulate the true treatment very closely and minimal acupuncture may be preferred. The most important matter though, seldom considered in clinical trials, is to find a way of assessing the adequacy of whichever control is chosen. This means ensuring that the psychological impact of the true treatment and the control are equivalent, in essence that they have equivalent placebo power.

THE CONCEPT OF THE PLACEBO

Placebo effects are seldom studied in their own right, usually being treated simply as a nuisance variable to be eliminated so that the specific treatment effects can be discerned. This may not matter too much when the placebo has, as in a drug trial, the same form as the true treatment. It is of greater concern with skilled physical treatments on conscious patients where changes to the treatment may be noticed by the patient. In these cases we need to pay a little more attention to the nature of the psychological effects involved.

As Richardson²⁶ has pointed out, the placebo is a 'portmanteau' concept, involving the use of a single term to describe a set of quite disparate phenomena. The power of placebo effects is influenced by treatment characteristics (more 'serious' treatments being more powerful) and by therapist characteristics (status, style of treatment administration). A number of different psychological processes may be involved.

For example many psychological processes influence pain perception. Effective placebo analgesia could conceivably be achieved through the manipulation of any of these processes. One placebo may divert the patient's attention (e.g. mock TENS) while another may reduce anxiety and reassure the patient (e.g. traditional inert tablet or injection)...

Assessing the power of a placebo, or comparing the placebo effects of two different treatments, or a treatment and a control condition is therefore not a simple matter. A host of non-specific factors may influence response to treatment and it is impossible to assess all potentially relevant factors; there are simply too many variables to take into account. The only solution is to select and assess one of the more important aspects of the placebo response. One of the most valuable approaches has been the assessment of the credibility of a treatment, and hence, indirectly the strength of the patient's expectations of improvement. The credibility of a treatment appears to be an important aspect of its

power, more credible treatments tending to have greater therapeutic effects²⁶.

VALIDATING PLACEBO CONTROLS IN ACUPUNCTURE STUDIES

The treatment credibility scale was originally conceived and employed by Borkovec and Nau²⁷ in a study of the credibility of different forms of psychological treatment; it has since been used in a variety of other psychological treatments as well as in studies of acupuncture²⁸⁻³⁰. The main questions identified by Borkovec and Nau in their 'Credibility of treatment rating scale' were:

- 1 How confident do you feel that this treatment can alleviate your complaint?
- 2 How confident would you be in recommending this treatment to a friend who suffered from similar complaints?
- 3 How logical does this treatment seem to you?
- 4 How successful do you think this treatment would be in alleviating other complaints?

The scale is sometimes given a slightly different title such as 'attitudes to acupuncture'. The form of the questions can be slightly amended to take account of the condition being treated and other circumstances of the trial²⁷.

It is usual to ask patients to rate their response to the four questions on a five point Likert scale (strongly agree ... strongly disagree). Vincent³⁰ assessed the psychometric properties in a sample of patients receiving acupuncture treatment and found that it had good internal consistency and re-test reliability; the questions all relate to the central concept of credibility and patients answer consistently on different occasions.

Petrie and Hazleman³¹ were the first to use the credibility scale in an acupuncture study. They assessed the credibility of acupuncture and mock-TENS on their study population before embarking on the clinical trial. Acupuncture and mock TENS were considered equally credible treatments for neck pain when demonstrated to patients before treatment began, so justifying the use of mock TENS as a placebo control. The trial demonstrated that acupuncture was significantly more valuable than mock TENS (ref) in providing pain relief, but placebo credibility was not re-assessed at the end of the study.

Vincent³⁰ used the scale in a controlled trial of the treatment of migraine by acupuncture in which real acupuncture was compared with a minimal acupuncture control. The credibility scale was given to patients at the end of the second and fifth treatments. There were no significant differences between true and sham treatments on any of the credibility measures. The true treatment was seen as slightly

more credible (though not significantly so) by the fifth week. In the trial the true acupuncture proved more effective than the control: the difference in credibility late in treatment probably reflects the fact that by then patients receiving the true treatment were deriving greater benefit.

A similar assessment was utilized by Bayreuther and Lewith²⁰ in a double blind crossover study of acupressure for the treatment of early morning sickness (EMS)²⁰. Twenty-three patients were given instructions to use a real or sham acupuncture point in a random order. The study showed a significant effect of real over sham treatment (66% of patients versus 33%). The credibility of the real and placebo points was evaluated using two questions from the credibility scale (Nos 1 and 3) at the start of the trial, and a further two at the end of the study (Nos 2 and 4). At the outset, the women were equally confident that acupressure would work at both positions. At the end of the study, they were significantly more confident in the real rather than the sham point. It seemed therefore that the sham point was a credible placebo at the outset. It produced a clinical response compatible with that of a placebo and the opinions of the women changed in response to an effective treatment.

All these studies support the contention that the credibility scale accurately reflects patients' beliefs about the authenticity and efficacy of the acupuncture treatment they received. This in turn suggests that the scale is a useful index of the psychological impact of acupuncture treatment and therefore the credibility of placebo controls within acupuncture studies.

IMPLICATIONS FOR RESEARCH: ROUTINE ASSESSMENT OF THE ADEQUACY OF CONTROL CONDITIONS

There are at least two viable control conditions for acupuncture trials, and the choice of control may vary according to the particular nature of the trial. Whatever the choice of control group it is valuable to check its adequacy. It is not feasible to assess every psychological variable that may be of importance, but it is possible to make some assessment of the adequacy of whichever control procedure one is using. The credibility measure is introduced as a check that the treatment and control are equivalent in their psychological impact. If they prove to be equally credible, this increases our confidence that the control procedure is adequate. If they prove to be different, the difference can be introduced as a variable in the statistical analysis of the results.

The fact that minimal acupuncture, acupressure and mock TENS are equally credible to acupuncture or acupressure in one study does not necessarily mean that they will be in all. A different clinician, a different group of patients, and a different setting may all influence the perception of the respective treatments or control

procedures. As acupuncture becomes more widely used patients will be more aware of the sensations of correct treatment and so more liable to detect variations introduced in control procedures. The recommendation must be that credibility, or a similar index of psychological impact, be routinely assessed in trials of acupuncture, in fact in all controlled trials of any physical treatment. Only then can we be sure, in any particular trial, that the treatment and control are adequately matched. Credibility is only one possible parameter of assessment, but it has already been shown to be a simple and useful measure. Routine measures of treatment credibility in trials of acupuncture would mean that arguments about placebo controls in acupuncture trials could in future rely less on speculation and more on evidence.

REFERENCES

- 1 Lewith GT, Machin D. On the evaluation of the clinical effects of acupuncture. *Pain* 1983;16:111-27
- 2 Vincent CA, Richardson PH. The evaluation of therapeutic acupuncture: concepts and methods. *Pain* 1986;24:1-13
- 3 Ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria based meta-analysis. *J Clin Epidemiol* 1990;11:1191-9
- 4 Tavola T, Gala C, Conte G, Invernizzi G. Traditional Chinese acupuncture in tension-type headache: a controlled study. *Pain* 1992;48(3):325-9
- 5 Vincent CA. Acupuncture. In: Lewith GT, Aldridge DA, ed. *Clinical Research Methodology*. London: Hodder & Stoughton, 1993:289-308
- 6 Kramer MS, Shapiro SH. Scientific challenges in the application of randomized controlled trials. *J Am Med Assoc* 1984;252:2739-45
- 7 Borglum-Jensen L, Melsen B, Borglum-Jensen S. Effect of acupuncture on headache measured by reduction in number of attacks and use of drugs. *Scand J Dent Res* 1979;87:373-80
- 8 Gallachi G. Acupuncture for cervical and lumbar syndrome. *Schweiz Med Wschr* 1981;111:1360-6
- 9 Gaw AC, Chang LW, Shaw LC. Efficacy of acupuncture on osteoarthritic pain: a double blind controlled trial. *N Engl J Med* 1975;293:375-8
- 10 Godfrey CM, Morgan P. A controlled trial of the theory of acupuncture in musculoskeletal pain. *J Rheumatol* 1978;5:121-4
- 11 Henry P, Baille H, Dartigues JF, Jogeix M. Headache and acupuncture. In: Pfaffenrath V, Lundberg PO, Sjaastad O, ed. *Updating in Headache*. Berlin: Springer Verlag, 1985
- 12 Chapman CR, Wilson ME, Gehrig JD. Comparative effects of acupuncture and transcutaneous stimulation on the perception of painful dental stimuli. *Pain* 1976;2:265-83
- 13 Reichmanis M, Becker RO. Relief of experimentally induced pain by stimulation at acupuncture loci: a review. *Compar Med East West* 1977;5:281-8
- 14 Stewart D, Thompson J, Oswald I. Acupuncture analgesia: an experimental investigation. *B M J* 1977;:67-70
- 15 Melzack R. Acupuncture and related forms of folk medicine. In: Melzack R, Wall PD, ed. *Textbook of Pain*. London: Churchill Livingstone, 1984
- 16 Le Bars D, Villaneuva L, Willer JC, Bouhassira D. Diffuse noxious inhibitory control (DNIC) in animals and man. *Acup Med* 1991;9(2):47-57
- 17 Richardson PH, Vincent CA. Acupuncture for the treatment of pain: a review of evaluative research. *Pain* 1986;24:15-40
- 18 Lewith GT, Vincent CA. On the evaluation of the clinical effects of acupuncture: a problem reassessed and a framework for future research. *Am Pain Soc J* 1994 (in press)
- 19 Dundee JW, McMillan CM. P6 acupressure and postoperative vomiting. *Br J Anaes* 1992;68:225-6
- 20 Bayreuther J, Lewith GT, Pickering R. Acupressure for early morning sickness: a double blind, randomized controlled crossover study. *Comp Ther Med* 1994;2(2):70-4
- 21 Macdonald AJR, Macrae KD, Master BR, Rubin AP. Superficial acupuncture in the relief of chronic low back pain. *Ann R Coll Surg Engl* 1983;65:44-6
- 22 Lewith GT, Field J, Machin D. Acupuncture compared with placebo in post-herpetic pain. *Pain* 1983;16:361-8
- 23 Dowson D, Lewith GT, Machin D. The effects of acupuncture versus placebo in the treatment of headache. *Pain* 1985;21:35-42
- 24 Vincent CA. The methodology of controlled trials of acupuncture. *Acup Med* 1989;6:9-13
- 25 Vincent CA. A controlled trial of the treatment of migraine by acupuncture. *Clin J Pain* 1989;5:305-12
- 26 Richardson P. Placebos: their effectiveness and mode of action. In: Broome AK, ed. *Health Psychology: Processes and Applications*. London: Chapman & Hall, 1989
- 27 Borkovec TD, Nau SD. Credibility of analogue therapy rationales. *J Beh Ther Exp Psychiat* 1972;3:257-60
- 28 McGlynn FD, McDonnel RM. Subjective ratings of credibility following brief exposure to desensitization and pseudotherapy. *Behav Res Ther* 1974;12:141-6
- 29 Shapiro DA. Comparative credibility of treatment rationales: three tests of expectancy theory. *Br J Clin* 1981;20:111-22
- 30 Vincent CA. Credibility assessment in trials of acupuncture. *Comp Med Res* 1990;4(1):8-11
- 31 Petrie J, Hazelman B. Credibility of placebo transcutaneous nerve stimulation and acupuncture. *Clin Exp Rheumatol* 1985;3:151-3

(Accepted 16 September 1994)