

# A trial design that generates only “positive” results

Ernst E, Lee MS

Complementary Medicine,  
Peninsula Medical School,  
Universities of Exeter and  
Plymouth, 25 Victoria Park  
Road, Exeter EX2 4NT UK

## Correspondence:

E Ernst  
E-mail:  
Edzard.Ernst@pms.ac.uk

Received : 28-03-08  
Review completed : 12-05-08  
Accepted : 22-05-08  
pubMed ID : 18626172  
J Postgrad Med 2008;54:214-6

## ABSTRACT

In this article, we test the hypothesis that randomized clinical trials of acupuncture for pain with certain design features (A + B versus B) are likely to generate false positive results. Based on electronic searches in six databases, 13 studies were found that met our inclusion criteria. They all suggested that acupuncture is effected (one only showing a positive trend, all others had significant results). We conclude that the 'A + B versus B' design is prone to false positive results and discuss the design features that might prevent or exacerbate this problem.

**KEY WORDS:** Acupuncture, bias, clinical trial, false positive, trial design

Randomized clinical trials (RCTs) are designed to minimize bias in comparative tests of therapeutic effectiveness. However, they are not always free of bias. In recent years, we have seen a plethora of RCTs adopting a design where patients are randomized to receive either usual care (the control group) or usual care plus the experimental treatment. Schematically this design could be depicted as 'A + B versus B'. At first glance, such comparisons may seem reasonable. However, on closer inspection, doubts emerge regarding whether such RCTs are fair scientific tests of the experimental intervention. These doubts originate from the theoretical view that 'A plus B' will invariably amount to more than 'B' alone. Even in cases where treatment A is a pure placebo, its placebo and other nonspecific effects could lead to a better outcome in the experimental group than in the control group. This would be particularly likely if i) the experimental treatment is associated with sizable nonspecific effects, ii) a subjective outcome measure is used, and iii) the experimental intervention 'A' causes a deterioration of the condition being treated.

Here we evaluate RCTs with the 'A + B versus B' design to determine whether any of these trials produced 'negative' findings, i.e. a result where the experimental group experiences outcomes which are either the same or worse than those of the control group. As hundreds of such studies exist, we focus on one particular situation which fulfils the above listed three criteria: acupuncture as a treatment of pain.

## Materials and Methods

Electronic databases were searched from their respective inception up to 10 October 2007: Medline, AMED, British

Nursing Index, CINAHL, EMBASE, and PsycInfo. The search terms used were: [acupuncture AND pain AND (usual care OR routine care OR standard care OR wait-list)] OR [acupuncture AND pain AND random]. In addition, screening of reference lists of all located articles and reviews was performed and our departmental files were hand searched.

Articles were considered if they reported an RCT in which human patients with any type of pain were treated with any type of acupuncture. Specifically, we included studies that fulfilled the following criteria: randomized patient allocation; pain as one outcome measure; control group receiving usual care; experimental group receiving usual care plus acupuncture. No language restrictions were imposed. We excluded feasibility studies, trials testing the effectiveness of TENS or laser treatment and studies that failed to include sufficient detail for assessment.

Hard copies of all articles were obtained and read in full by both authors. When an RCT had more than two treatment groups, we evaluated only the two groups ('A + B' and 'B') mentioned above. Both authors independently extracted key data [Table 1]. The methodology quality of all RCTs was assessed using the modified Jadad score.<sup>[1]</sup>

## Results

We found 200 RCTs, 13 of which met our inclusion/exclusion criteria [Table 1].<sup>[2-14]</sup> Eleven RCTs had a parallel group design,<sup>[2,5-14]</sup> and two followed a partial crossover design.<sup>[3,4]</sup> Nine trials included two groups,<sup>[3,4,6,8,10-14]</sup> three had three groups<sup>[5,7,9]</sup> and one had four groups of patients [Table 1].<sup>[2]</sup>

Table 1: Summary of randomized clinical trials comparing acupuncture plus usual care with usual care only

First author (Year)	Design, quality score, <sup>a</sup> allocation concealment	Pain condition (sample size) <sup>b</sup>	Control treatment (C)	Experimental treatment (E)	Main results
Helms (1987) <sup>[2]</sup>	4 parallel, patient blind	Dysmenorrhea ( $n = 48$ )	Analgesics as required	C + 9 acupuncture sessions	Non sign trend for pain reduction in E compared to C
Berman (1999) <sup>[3]</sup>	4 (1 + 1 + 1 + 0 + 1), NR Partial crossover, assessor blind, 4	Osteoarthritis pain (knee) ( $n = 73$ )	Analgesics and NSAID	C + 16 acupuncture sessions	Sign more pain reduction in E compared to C
Meng (2003) <sup>[4]</sup>	(1 + 1 + 1 + 1 + 0), NR Partial crossover, assessor blind, 4	Chronic low back pain ( $n = 55$ )	Usual care (NSAIDs, analgesics, back exercises)	C + 10 acupuncture sessions	Sign more pain reduction in E compared to C
Tukmachi (2004) <sup>[5]</sup>	4 (1 + 1 + 1 + 1 + 0), adequate	Osteoarthritis pain (knee) ( $n = 30$ )	Analgesic and NSAID	C + up to 10 acupuncture sessions	Sign more pain reduction in E compared to C
Vickers (2004) <sup>[6]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), adequate	Chronic headache ( $n = 401$ )	Usual care	C + 12 acupuncture sessions	Sign more pain reduction in E compared to C
Witt (2005) <sup>[7]</sup>	3 parallel, assessor and patient blind, 5 (1 + 1 + 1 + 1 + 1), NR	Osteoarthritis pain (knee) ( $n = 300$ )	NSAID as required	C + up to 12 acupuncture sessions	Sign more pain reduction in E compared to C
Witt (2006) <sup>[8]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), n.r.	Osteoarthritis pain (knee or hip) ( $n = 712$ )	Conventional treatment (n.r.)	C + up to 15 acupuncture sessions	Sign more pain reduction in E compared to C
Brinkhaus (2006) <sup>[9]</sup>	3 parallel, patient blind 4 (1 + 1 + 1 + 0 + 1) NR	Chronic low back pain ( $n = 298$ )	NSAID as required	C + 12 acupuncture sessions	Sign more pain reduction in E compared to C
Thomas (2006) <sup>[10]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), NR	Persistent low back pain ( $n = 241$ )	Usual care (n.r.)	C + 10 acupuncture sessions	Sign more pain reduction in E compared to C
Witt (2006) <sup>[11]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), adequate	Low back pain ( $n = 3093$ )	Usual care (n.r.)	C + up to 15 acupuncture sessions	Sign more pain reduction in E compared to C
Witt (2006) <sup>[12]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), adequate	Chronic neck pain ( $n = 3766$ )	Usual care	C + up to 15 acupuncture sessions	Sign more pain reduction in E compared to C
Goertz (2006) <sup>[13]</sup>	2 parallel, assessor blind 4 (1 + 1 + 1 + 1 + 0), adequate	Military personnel with acute pain syndromes ( $n = 100$ )	Emergency medical care (prescription medication and other appropriate treatments)	C + 1 auricular acupuncture placement and retained 4 to 6 days	Sign more pain reduction in E compared to C
Reinhold (2007) <sup>[14]</sup>	2 parallel, open, 3 (1 + 1 + 1 + 0 + 0), NR	Osteoarthritis pain (knee or hip) ( $n = 489$ )	Conventional routine medical care (n.r.)	C + up to 15 acupuncture sessions	Sign more pain reduction in E compared to C

GP: General practitioner; NR: Not reported; NSAID: Non-steroidal anti-inflammatory drugs; Sign: Significant; <sup>a</sup>Quality score: Jadad score (randomization + appropriate randomization method + describing withdrawals and dropouts + assessor blinding + patient blinding, maximum 5 points);

<sup>b</sup>Sample size is randomized sample number

The RCTs differed in many aspects: they recruited patients with different types of pain, employed a range of methods for quantifying pain, used different types of acupuncture administered according to different treatment schedules, employed different types of usual care, and varied in terms of Jadad score and sample size. Despite these differences, all studies generated results that suggest significantly better outcomes in the experimental ('A + B') compared to the control group ('B'). The only exception is the RCT by Helms *et al.*,<sup>[2]</sup> which only showed a non-significant trend in that direction, most likely due to a Type II error caused by its very small sample size ( $n = 12$ ).

### Discussion

Randomized clinical trials with the 'A + B versus B' design are currently popular, particularly for pragmatic studies. Our findings suggest that, in the realm of acupuncture for pain, this

design is likely to generate false positive results. Alternatively acupuncture could, of course, be a highly effective intervention for reducing pain. However, recent RCTs which rigorously control for placebo effects by employing non-penetrating sham-devices as control interventions suggest that acupuncture is not superior to sham-acupuncture for pain.<sup>[15]</sup> Moreover, several RCTs with patients suffering from conditions other than pain (for which acupuncture has not been proven to be effective) indicate that this phenomenon may not be confined to studies of pain.

The reason for the inability of the 'A + B versus B' design to generate negative results (under the above-stated three conditions) seems obvious: even in the absence of any specific therapeutic effect, the results of such studies would be positive due to nonspecific effects such as a placebo-effect, the additional care given to patients, the therapist-patient relationship or social desirability. A further contributor could

be the disappointment experienced by patients of the control groups when not receiving the experimental treatments they may have hoped for.

The design issues highlighted here are well known to research methodologists. They are, however, not always appreciated by clinicians who read published articles and may tend to equate RCTs with the highest level of reliability. Often the authors of these articles fail to discuss the drawbacks of the ‘A + B versus B’ design in sufficient detail. Certainly they were not emphasized by the authors of the studies included in this review.

Studies with three or more groups are usually designed to answer more than one research question. Therefore such RCTs might be valuable even if they include a comparison of ‘A + B versus B’. Conclusions based specifically on any ‘A + B versus B’ comparison within such trials can nevertheless be criticized on the basis of the above arguments. Similarly, “pragmatic” trials that adopt a different design (for instance ‘A versus B’) are not affected by our criticism.

It would, of course, be an over-interpretation of our results to state categorically that RCTs with the ‘A + B versus B’ designs can, in principle, only generate positive results. We have shown this to be likely only for acupuncture as a treatment for pain. In other therapeutic areas, negative trials with that design may exist. Crucially, however, these RCTs do not fulfill the three additional criteria outlined in the introduction of this article. Interventions that are less prone to generating false positive findings when tested in a ‘A + B versus B’ study include those that not only do not improve but worsen the condition in question and treatments which do not generate sizable placebo-effects. Also, studies that employ objectively measurable endpoints might, in some cases, offer a protection against such false positive findings.

In conclusion, our systematic review of RCTs of acupuncture for pain control with the ‘A + B versus B’ suggests that this trial methodology is likely to produce false positive results. Such studies may therefore not be adequate scientific tests of the effectiveness of therapeutic interventions.

## References

1. White AR, Ernst E. A systematic review of randomized controlled trials of acupuncture for neck pain. *Rheumatol* 1999;38:143-7.
2. Helms JM. Acupuncture for the management of primary dysmenorrhea. *Obstet Gynecol* 1987;69:51-6.
3. Berman BM, Singh BB, Lao L, Langenberg P, Li H, Hadhazy V, *et al.* A randomized trial of acupuncture as an adjunctive therapy in osteoarthritis of the knee. *Rheumatology* 1999;38:346-54.
4. Meng CF, Wang D, Ngeow J, Lao L, Peterson M, Paget S. Acupuncture for chronic low back pain in older patients: A randomized, controlled trial. *Rheumatology* 2008;42:1508-17.
5. Tukmachi E, Jubb R, Dempsey E, Jones P. The effect of acupuncture on the symptoms of knee osteoarthritis: An open randomised controlled study. *Acupunct Med* 2004;22:14-22.
6. Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith C, Ellis N *et al.* Acupuncture for chronic headache in primary care: Large, pragmatic, randomised trial. *BMJ* 2004;328:744.
7. Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S *et al.* Acupuncture in patients with osteoarthritis of the knee: A randomised trial. *Lancet* 2005;366:136-43.
8. Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN. Acupuncture in patients with osteoarthritis of the knee or hip: A randomized, controlled trial with an additional nonrandomized arm. *Arthritis Rheumat* 2006;54:3485-93.
9. Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, *et al.* Acupuncture in patients with chronic low back pain: A randomized controlled trial. *Arch Intern Med* 2006;166:450-7.
10. Thomas KJ, MacPherson H, Thorpe L, Brazier J, Fitter M, Campbell MJ, *et al.* Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. *BMJ* 2006;333:623-6.
11. Witt CM, Jena S, Selim D, Brinkhaus B, Reinhold T, Wruck K, *et al.* Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. *Am J Epidemiol* 2006;164:487-96.
12. Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN. Acupuncture for patients with chronic neck pain. *Pain* 2006;125:98-106.
13. Goertz CM, Niemtow R, Burns SM, Fritts MJ, Crawford CC, Jonas WB. Auricular acupuncture in the treatment of acute pain syndromes: A pilot study. *Mil Med* 2006;171:1010-4.
14. Reinhold T, Witt CM, Jena S, Brinkhaus B, Willich SN. Quality of life and cost-effectiveness of acupuncture treatment in patients with osteoarthritis pain. *Eur J Health Economics* 2007;(epub: 19 7 2007).
15. Ernst E. Acupuncture: A critical analysis. *J Intern Med* 2006;259:125-37.

Source of Support: Nil, Conflict of Interest: Not declared.