Is acupuncture effective for the treatment of chronic pain: a systematic review

Ezzo J, Berman B, Hadhazy V A, Jadad A R, Lao L, Singh B B

Authors' objectives
To assess the effectiveness of acupuncture as a treatment for chronic pain within the context of the methodological quality of the studies.

Searching
MEDLINE (1996 to 1999) and two specialised complementary medicine databases, the Cochrane Complementary Medicine Field registry and the University of Maryland CAMPAIN (Complementary and Alternative Medicine in Pain) were searched. The terms 'acupuncture', 'alternative medicine', 'electroacupuncture', 'moxibustion', 'injections, intramuscular', 'Medicine, Traditional Chinese' were exploded as MeSH terms and searched as textwords. 'Trigger point therapy;' and 'auriculotherapy' were searched as textwords only. The search was limited to human subjects and English publications. Sixty-nine conference proceedings and abstracts were handsearched to locate trials which had not been fully published elsewhere. Conferences were identified by a conference title search of keywords 'pain', 'analgesia', 'acupuncture', 'rheumatology', 'arthritis', 'headache', 'migraine', 'fibromyalgia', 'fibrositis', or 'temporomandibular dysfunction'. Bibliographies from relevant papers were also searched.

Study selection
Study designs of evaluations included in the review
Randomised controlled trials (RCTs) for which a between groups analysis was presented. Only studies that were published in English were included. Multiple publications of the same study were omitted so that each study population was counted only once.

Specific interventions included in the review
Acupuncture that uses needles. Control groups were classified into one of 4 categories: waiting lists; physiologically inert controls e.g. sham TENS, sugar pills, placebo acupuncture; sham acupuncture; and standard medical care e.g. drug therapy or physiotherapy. Placebo acupuncture was defined as a mock acupuncture procedure in which needles were not actually inserted. Sham acupuncture was defined as a mock acupuncture procedure in which needles were inserted in the skin.

Participants included in the review
Individuals who have been in pain for longer that three months. The types of pain reported by included studies were low back, neck, arthritis (osteo and rheumatoid), Raynauds, angina, post herpes, headache, ankylosing spondylitis, pancreatitis, fibromyalgia, mixed, musculoskeletal, face, dysmennorrhea, migraine, cystitis, and jaw pain.

Outcomes assessed in the review
For inclusion studies had to have some measure of pain relief. Follow-up for included studies varied from 1 day to 1 month.

How were decisions on the relevance of primary studies made?
The authors do not state how many of the reviewers performed the selection. The titles, abstracts, and when necessary, the study manuscripts were reviewed to establish whether they met all the inclusion criteria. Authors were contacted for missing information if inclusion was not clear.

Assessment of study quality
Study quality was assessed using a validated instrument (see Other Publications of Related Interest no.1) with the additional items that influence internal validity (see Other Publications of Related Interest no.2). A high-quality score was defined as three or more on a five-point scale. The criteria list for assessing quality included whether randomisation was stated, randomisation method described and appropriate, double-blinding reported, whether double-
blinding was likely to have succeeded, outcome assessor blinded, reason for and number of withdrawals and drop-outs
reported, concomitant treatments eliminated or controlled for, compliance satisfactory, homogeneity adequate, and equivalent
therapeutic time in treatment and control group. Two reviewers independently reviewed all studies. Inter-rater
disagreements were resolved by discussion.

**Data extraction**

Two reviewers independently reviewed all studies. Inter-rater disagreements were resolved by discussion. Data were
extracted on four aspects of the acupuncture treatment:

1. The total number of treatments.
2. Number of acupuncture points needled per session.
3. Whether 'de. qi' was elicited, and confined to the real acupuncture group when sham acupuncture was used as the
control group.
4. Whether formula or individualisation acupuncture treatments were given.

Other data extraction presented in a table, included reference details, type of control, condition being treated number
of participants, results and quality score. When available, proportions responded and the country of the study were also
extracted. Outcomes were defined as positive when acupuncture was significantly more effective than the reference
group; neutral when acupuncture was not significantly different from the reference group; or negative when
acupuncture was significantly less effective than the reference group. A P value greater that 0.05 was used to define a
significant outcome. When pain outcomes were inconsistent, then the measure assessing the most global assessment
was used to determine the outcome of the study. Authors were contacted for missing information.

**Methods of synthesis**

How were the studies combined?

A best evidence synthesis (see Other Publications of Related Interest no.3) method was used to formulate conclusions
on the effectiveness of acupuncture for each type of control group. This method consists of four levels of scientific
evidence and takes into account both quality and the outcome of the studies.

How were differences between studies investigated?

Clinical heterogeneity was discussed in a narrative. Sensitivity analysis was performed using Chi square analysis to
test for statistically significant association between overall trial quality and trial outcomes. The Mann-Whitney U-test
was used to compare the proportions responding in the various control groups. Logistic regression was selected as the
multivariate model to simultaneously control for trials quality while identifying aspects of the acupuncture treatment
associated with positive outcome.

**Results of the review**

Fifty-one RCTs representing 2423 chronic pain patients. The median sample size per group was 18 and the mode was
15.

Substantial clinical heterogeneity from different conditions, control groups, outcomes and types of acupuncture,
precluded statistical pooling. Results were positive in 21 studies, negative in 3 and neutral in 27.

Five RCTs (n=215) compared acupuncture to no treatment (waiting list) and all had positive outcomes, but all were
low quality. Eleven RCTs (n=331) comparing acupuncture to physiologically inert control were rated as 7 low quality
and 4 high quality. The results were inconsistent (6 neutral, 5 positive). Twenty-two RCTs (n=1309) compared sham
to real acupuncture. Of these, 12 were high quality, 10 were low quality. Results were inconsistent (15 neutral, 7
positive). Twelve RCTs (n=539) compared acupuncture to standard care, of which one trial used acupuncture as an
adjunct to standard care vs standard care alone. One trial was high quality and 11 were low quality. The results were
inconsistent (6 neutral, 3 positive, 3 negative).
Two-thirds of the studies (n=34) received a low-quality score and low-quality trials were significantly associated with positive results (P=0.05). High-quality studies were associated with being clustered in designs using sham acupuncture as the control group, where the risk of false negative (type II) errors was high due to large sample size requirements. Six or more acupuncture treatments were significantly associated with positive outcomes (P=0.03) even after adjusting for study quality (0.01, odds ratio 9.2, 95% CI: 1.6 to 53.2). Sensitivity analysis on individual methodological components demonstrated that trials which did not report double blinding, did not report having blinding outcomes assessor, did not report blinding patients, and did not have equivalent amounts of therapeutic time between the experimental groups were significantly associated (p<0.05) with positive outcomes.

**Authors' conclusions**
There is limited evidence that acupuncture is more effective than no treatment for chronic pain; and inconclusive evidence that acupuncture is more effective than placebo, sham acupuncture or standard care. However, we have found an important relationship between the methodology of the studies and their results that should guide future research.

**CRD commentary**
This is a well conducted review that included a clear objective with predefined inclusion/exclusion criteria. A comprehensive search of the literature was undertaken although this did not include a search for unpublished studies and, therefore, the presence of publication bias cannot be ruled out. In addition, only studies published in English were included and therefore some important data may have been missed. However, the implications of these shortcomings are discussed by the authors. Information about the methodology of the review process was clearly presented, although it was not stated how many reviewers were involved in the process of assessing articles for relevancy. A systematic procedure involving two reviewers was used to assess the validity of included trials and to perform data extraction. Some relevant information about included studies was presented in table format, although these data were limited. Heterogeneity between studies was considered along with the quality of included studies (in the sensitivity analysis). A narrative syntheses of the results was appropriate. The authors conclusions appear to follow from the results.

**Implications of the review for practice and research**
Practice: The authors note that there is limited evidence that acupuncture is better than no treatment (waiting list).

Research: The authors make the following recommendations.

1. More high-quality trials are needed to answer the question of effectiveness.
2. Larger trials must be conducted to assess whether real acupuncture is more effective than sham acupuncture.
3. Issues of dosing and an adequate/optimal acupuncture procedure need to be examined.
4. The effectiveness of acupuncture for muscular pain is likely to be a promising area for future research.
5. Future trials should use study groups that are balanced, and a measure of the level of the patient's discomfort should be included so as to ascertain whether the discomfort predicts subjective reports of benefit.
6. Studies should incorporate other dimensions of acupuncture which were not found to be significant in this study, but for which a biological model exists, such as the elicitation of 'de qi', and the number of points stimulated.

**Funding**
Maurice Laing Foundation; National Institutes of Health, grant number 1 R21-RR09327-01.

**Bibliographic details**

PubMedID
10812251

Other publications of related interest


These additional published commentaries may also be of interest.


Haigh C. Review: limited evidence suggests that acupuncture is more effective than no treatment for chronic pain. Evid Based Nurs 2001;4:17.


Indexing Status
Subject indexing assigned by NLM

MeSH
Acupuncture Therapy; Chronic Disease; Humans; Pain Management; Randomized Controlled Trials as Topic; Sensitivity and Specificity

AccessionNumber
12000001179

Date bibliographic record published
31/07/2001

Date abstract record published
31/07/2001

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.