Spinal manipulation: a systematic review of sham-controlled, double-blind, randomized clinical trials
Ernst E, Harkness E

Authors' objectives
To assess whether spinal manipulation (SM) conveys more than a placebo effect for various conditions.

Searching
MEDLINE, EMBASE, CISCOM, AMED and the Cochrane Library were all searched from their inception to the end of 1998 for eligible trials reported in any language. The search terms used were 'chiropractic', 'spinal manipulation', 'osteopathy' and 'controlled clinical trials'. Additional material was obtained by contacting experts and chiropractic organisations, and by examining bibliographies and the authors' own files were consulted. Bibliographies were screened.

Study selection
Study designs of evaluations included in the review
The included studies had to be sham-controlled clinical trials that were double-blind to both the patients and evaluators.

Specific interventions included in the review
SM as practised by chiropractors or other healthcare professionals, with sham SM as the control group treatment (i.e. controls receiving an intervention mimicking true SM, deemed by the investigators to be inactive). Studies of treatment packages that included other interventions were excluded, as were studies of spinal mobilisation and trials with other types of placebo control treatments. The treatment was undertaken by a chiropractor in all of the identified trials.

Participants included in the review
Participants with any condition were eligible. The conditions treated included chronic lower back pain (LBP), primary childhood nocturnal enuresis, both chronic and childhood asthma, phobia and primary dysmenorrhoea.

Outcomes assessed in the review
The inclusion criteria were not specified for the outcomes, which varied according to the condition under study. The measures included: pain, function or disability and depression for LBP; lung function for asthma; pulse rate and emotional discomfort for phobia; and pain and menstrual distress for dysmenorrhoea. In addition, each trial was critically appraised.

How were decisions on the relevance of primary studies made?
The authors do not report how the papers were selected for the review, or how many of the reviewers performed the selection.

Assessment of study quality
Validity was assessed using the method of Jadad et al. (see Other Publications of Related Interest), supplemented by assessments of other methodological factors such as whether participant 'de-blinding' was checked. Both authors made independent judgements of validity. Any disagreements were settled through discussion.

Data extraction
The two authors extracted the data independently in a standardised manner. Information on trial methodology, treatment schedule, outcome measures and significance of the results was extracted.

Methods of synthesis
How were the studies combined?
Each trial was described separately in a narrative summary.

How were differences between studies investigated?
Each study was described in detail but no comparisons were made. Between-study differences could only be evaluated by examining the tables and text.

**Results of the review**

Eight studies were included: there were 3 studies (n=19,209,18) of chronic LBP, 1 study (n=46) of primary nocturnal enuresis associated with spinal subluxations, 1 study (n=31) of chronic asthma, 1 study (n=91) of childhood asthma, 1 study (n=20) of phobia, and 1 study (n=138) of primary dysmenorrhoea.

Three trials (two of back pain and one of enuresis), whose authors concluded that they provided evidence in favour of chiropractic, were assessed as being seriously methodologically flawed. Their conclusions were criticised on the basis of reanalysis and interpretation of their results. The third trial of back pain found a significant difference for the Oswestry Disability Score, but not for pain measured by the visual analogue scale. While this was a large study, there was a drop-out rate of 44% with no intention to treat analysis, so the review authors assessed the evidence as unreliable.

Of the remaining studies, the three studies that were assessed to be methodologically sound (two for asthma and one for primary dysmenorrhoea) found no significant difference between the SM and control groups, although one of the asthma trials could have failed to detect a true difference due to its small sample size. The trial for phobia, which had a small sample size but was otherwise of a fair quality, found a significant difference for emotional discomfort (visual analogue scale) but not for pulse rate.

**Cost information**
The cost was mentioned in the discussion section, but the included studies provided no data on the costs.

**Authors' conclusions**
Sham-controlled, double-blind, randomised controlled trials of SM are feasible. Few exist and some of the existing studies have serious methodological shortcomings. The three most rigorous trials did not suggest that SM is associated with specific therapeutic effects.

**CRD commentary**
The review question was clear but the range of conditions that it included, and the small number of eligible studies, meant that its conclusions were very limited. It was generally well-conducted and had a good search strategy and full validity assessment. Details of the individual trial characteristics were well-summarised. However, the results were lacking and it was not possible to assess them fully; the reported results were confined to whether they reached statistical significance, and no effect sizes or associated probabilities or confidence intervals were given.

While the choice of narrative methods was appropriate, the trials were assessed individually with only the broadest synthesis of results from those trials for the same conditions (i.e. asthma and back pain). In a narrative synthesis, the studies are usually grouped according to methodological rigour and/or other factors such as participant characteristics, treatment regimes or type of outcome, and comparisons are made of salient features. Some synthesis was conducted in relation to validity assessments.

The authors' conclusions with respect to validity appear justified. However, the conclusion about the suggested lack of an association between SM and therapeutic effects was not well-expressed. Firstly, it can only be applied to the conditions of asthma and primary dysmenorrhoea studied in the more rigorous trials, and it should be noted that there was only one trial of reasonable size and quality for each condition. Secondly, a lack of effectiveness is not proven by a lack of statistical significance; it can only be said that these results provide no reliable evidence that there is a treatment effect for these conditions.
The question of the effectiveness of SM remains open for the other conditions studied, including LBP. Some effects were found but since the studies were inadequate in relation to their conduct, analysis or sample size, the significance or non significance of these results has not been proved.

**Implications of the review for practice and research**

**Practice:** The authors did not state any implications for practice.

**Research:** The authors state that further sham-controlled, double-blind, randomised controlled trials of good quality and of sufficient size are required to assess the effectiveness of SM in relation to conditions that it aims to treat, particularly LBP.

**Bibliographic details**

**PubMedID**
11576805

**Other publications of related interest**


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Double-Blind Method; Humans; Manipulation, Spinal; Placebo Effect; Randomized Controlled Trials as Topic; Treatment Outcome

**AccessionNumber**
12001002296

**Date bibliographic record published**
30/11/2002

**Date abstract record published**
30/11/2002

**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.