Ruddock, J. K., Sallis, H. M., Ness, A. R., & Perry, R. E. (2016). Spinal Manipulation vs Sham Manipulation for Nonspecific Low Back Pain: A Systematic Review and Meta-Analysis. Journal of Chiropractic Medicine, 15(3), 165–183. DOI: 10.1016/j.jcm.2016.04.014

Spinal manipulation versus an effective sham for non-specific low back pain: a systematic review and meta-analysis.

Jay K Ruddock M.Ost [a], Hannah Sallis MSc [b, c], Andy Ness PhD [d], Rachel E Perry MPhil

ABSTRACT

Purpose: Non-specific low back pain (NSLBP) is a common and costly condition which will affect the majority of people in their lifetime. Successful treatment of this condition would be of great benefit to the general population. Spinal manipulation (SM) has been suggested as an effective treatment. However, there is still debate over whether the supposed benefit is due to specific treatment effects or a non-specific 'placebo effect'. Issues around safety of the technique have also been raised.

This review seeks to systematically identify and critically evaluate all randomised controlled trials (RCTs) of SM in the treatment of NSLBP utilising an effective sham manipulation**. Methods:** Five electronic databases were searched from their inception to March 2015 to identify all relevant trials. Reference lists of retrieved articles were hand-searched. All data was extracted by two independent reviewers and risk of bias was assessed using the Cochrane back group Risk of Bias tool (2009).

Results: Nine RCTs were included in the systematic review and four were found to be eligible for inclusion in a meta-analysis. Participants in the SM group had improved symptoms compared with participants receiving sham treatment (SMD=-0.36; 95% CI -0.59,-0.12). The majority of studies were of low risk of bias, however several of the studies were small, the practitioner could not be blinded, some studies did not conduct intention-to-treat (ITT) analysis and had a high level of drop outs. **Conclusions:** Thus given the small number of studies included in this analysis, we should be cautious of making strong inferences based on these results.

Key words: Low back pain, spinal manipulation, placebo, review

INTRODUCTION

Over the course of a lifetime approximately 80% of people will experience low back pain (LBP). Non-specific low back pain (NSLBP) is the second most common reason for worker absenteeism[1,2] and is the most common reason to attend a manual therapy clinic.[3,4]

NSLBP is characterised by pain in the posterior lumbar spine, sacral spine or para-spinal tissues which may be accompanied by decreased range of motion.[5] The aetiology is unclear and a definitive cause remains elusive for researchers.[6] Several different approaches to treatment have been identified, with mixed evidence for their success.[7,8,9] One of the treatments widely used is spinal manipulation (SM).

Spinal Manipulation and Mobilisation

SM can be defined as “treatments that use high velocity/low amplitude (HVLA) to move a joint that is exhibiting somatic dysfunction through its restrictive barrier.” Several models suggest that this technique would be able to produce a hypo-analgesic effect, either by structural[11,12,13] or neurological processes,[14,15] whilst others have postulated it acts through non-specific or ‘placebo’ effects.[16,17]

In contrast, spinal mobilisation uses low velocity/low amplitude cyclical techniques (non-thrust mobilisation); it has been argued that this method of action differs from that of HVLA techniques, thus mobilisation and manipulation should be investigated separately.[18] SM can have serious (although very rare) adverse outcomes such as

intervertebral disc prolapse and fracture,[19] whereas there are no reported adverse events reported from receiving non-thrust spinal mobilisation.[5] If it could be established that there were no specific treatment benefits from HVLA techniques on NSLBP then it would be inappropriate to perform them on patients.

Controlling the Placebo Effect in Trials of SM

In order to exclude possible placebo effects in trials of SM, the control group must either be screened for previous experience of SM,[20] or be exposed to an effective sham intervention.

There is little agreement among experts as to what constitutes an effective sham manipulation.[21] However, there is some evidence as to what may be acceptable as an effective sham manipulation of the lumbar spine. Hancock et al (2006)[21] demonstrated the most credible sham procedure was Maitland's 'log roll.'[22] This procedure comprises 'placing the patient in a side-lying position and placing the physiotherapist's hands over the over the lower ribs and ilium. The pelvis and trunk are then rolled together so no lumbar inter-vertebral motion occurs' (Hancock 2006 p136).

Fulda et al (2007)[23] showed participants videos of side-lying SM, light touch or ultrasound to gauge patients' perceptions of treatments for lumbar spine pain. The participants viewed SM as the therapy most likely to reduce pain and improve function, suggesting that a sham needs to physically resemble a SM technique for it to be believable. Hawk and Long (2000)[24] and Machado et al (2008)[25] also identified the importance of equalization of the non-specific effect of physical touch between participants. The use of an indistinguishable placebo should counteract any subtle differences between groups shown to influence treatment outcomes.[26,27] Other active therapies are not considered a viable control as they can lead to erroneous

interpretation due to varied contextual factors which produce a placebo effect or specific treatment effects.[28] Thus, for a sham manipulation to be an effective control it should physically resemble a HVLA technique and be performed so as to eliminate subtle differences between the intervention group and the control group. For the purpose of this review the term “effective” sham control is used to denote control groups that met these criteria.

Previous reviews have compared SM to sham manipulations, however they have either included papers that did not utilise an effective sham[7] or they permitted techniques that were not solely HVLA.[7,29,30,31]

Bronfort et al’s (2010)[7] review compared SM to a sham intervention. However, one included study[88] used an inappropriate sham intervention by using gluteal massage. Rubenstein et al’s (2011)[30] review of SM for chronic low back pain included one study[32] which used several techniques (HVLA, muscle energy techniques, soft tissue manipulation, fascial manipulation and cranio-sacral) in their treatment group. Rubenstein et al’s (2012)[31] review of SM for acute back pain only included one study of SM versus a sham intervention. None of these reviews distinguished between SM and mobilisation.

Ernst and Harkness’s (2001 p. 887)[17] review of SM for a range of conditions identified three trials[33,34,89] and recommended that ‘the specific efficacy of SM for low back pain must await adequately designed sham-controlled trials’. The most recent systematic review[35] examined SM, mobilisation and exercise as separate interventions against shams in NSLBP sufferers of various durations. Five studies were analysed in groups determined by similarity of patients, interventions, comparisons and outcomes. However, no meta-analysis was performed.

This systematic review critically evaluates data from randomised controlled trials (RCTs) using HVLA techniques for people with NSLBP. The aim is to assess SM in isolation rather than as part of a treatment package of care. Any specific treatment effects or adverse events that are identified can be isolated to SM. To be eligible, the comparison group had to be an effective sham.

METHODS

The following databases were searched from their inception to March 2015: MEDLINE and AMED (via Ovid), Web of Science and Central via Cochrane library, using a combination of MeSH and key word terms (see Appendix 1 for the search strategy). No restrictions were applied regarding language or date. Reference lists of all full-text articles and all relevant systematic reviews were hand-searched for additional studies. A protocol was produced and can be found at http://www.crd.york.ac.uk/PROSPERO/, reg number CRD42014008886.

Study selection

All titles and abstracts retrieved from the searches were assessed for eligibility. Articles that appeared to meet the inclusion criteria were retrieved in full and independently considered for inclusion by two reviewers (JR, RP). Disagreements were resolved through discussion (see Fig 1 for flow diagram). The following inclusion criteria were predefined:

Type of Participant Participants of either gender and >18 years with NSLBP

Type of Intervention Randomised controlled trials (RCTs) which used HVLA SM as an intervention. Studies which either screen for subject expectation of SM, or assess for effective blinding after the intervention, were also included.

Type of Comparator Studies which have an effective sham control i.e. the physical act of the sham manipulation must be credible.

Type of Outcome Studies that had a perceived measure of pain as an outcome (e.g. VAS pain scores, standardised questionnaires)

We excluded studies that were not randomised, that used participants with radicular symptoms, history of lumbar spine surgery, osteoporosis, spinal stenosis and spondylolisthesis or were pregnant. We also excluded any studies that used other therapies, drugs, exercise, advice or information as a control or did not include sufficient details of the blinding process in the text. A table of excluded studies, with reasons for exclusions, can be found in Appendix 3.

Only completed RCTs were included (reports of ongoing trials were excluded (e.g. protocol papers)). The primary outcome was any measure of pain (both standardised and non-standardised). The secondary outcome was any adverse event mentioned. Data from included studies were extracted independently by two reviewers (JR, RP) using a form with pre-defined criteria.

The risk of bias (RoB) of all included RCTs was evaluated independently by two reviewers (JR, RP) using the Risk of Bias tool of the Cochrane Back Review Group (CBRG)[36](Appendix 2). Studies are rated as having a low RoB when 'at least 6 of the CBRG criteria have been met and the study has no serious flaws' (Furlan et al 2009 p1932). Disagreements were resolved through discussion with the 3rd reviewer (HS). The manuscript was developed using the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) checklist.[37]

Meta-analysis

Meta-analyses were performed in Stata 13 (Stata, College Station, TX), using the user-contributed commands metan [38] and metafunnel.[39] Standardised mean differences (SMD) and corresponding standard errors were calculated for included studies using means, standard deviations and sample sizes reported in the relevant publications. Due to the small sample sizes of some studies, Hedges' g, an extension of Cohen's d adjusted for small sample bias, was calculated.[40] A negative SMD corresponds to a lower pain score being associated with the SM group.

Effect estimates were pooled using a random effects model. Unlike a fixed effects model, which assumes that each study estimates the same effect size, a random effects model assumes that each study estimates a different effect, but that these are drawn from some common distribution.[41] Thus, in addition to random sampling error, differences may also be due to dissimilarities between study populations and designs. The I^2 statistic was also calculated, which measures inconsistency between estimates and is independent of sample size.[42]

Post-hoc Sensitivity analysis

Initially, only studies reporting results at follow up were included in the main analysis, using the last follow up time point as comparison. However to maximise comparability, the 1 month follow up data from Senna et al's (2011)[43] study was used in this analysis (as opposed to the last time point at 10 months). One sensitivity analysis included only studies which collected pain measurements immediately post-treatment. A further

sensitivity analysis investigated the effect of including studies assessed as being at a high risk of bias.[33]

Dealing with missing data

Where possible, we extracted the number of participants randomised to a treatment arm, the mean pain score, and standard deviation for each group. In some cases, only means and standard errors were reported, in which case the group sizes were used to estimate the corresponding standard deviation.[43] In one case, we were unable to extract relevant information from the initial publication, but instead results reported in a subsequent review which included this study.[33] Where there was insufficient information we contacted authors.

RESULTS

The literature search identified 1625 potentially relevant titles and abstracts. After screening, nine RCTs were identified that met the inclusion criteria for this review. The studies were published from 1986 to 2014. Five originated from USA,[33,44,45,46,89] one from Canada,[47] one from France,[48] one from Egypt[43] and one from Germany.[49] Eight were in English and one in French.[48] The total number of participants enrolled was 646 (252 male, 394 female), the sample size varied from 6 to 145 and age ranged from 18-65 years. The mean age of participants was incalculable due to incomplete reporting in one of the trials.[44] Four types of pain outcome measure were used. To assess pain levels directly either a Visual Analogue Scale (VAS) or Numerical Rating scale (NRS) were used. To assess physical function due to pain either the Oswestry Low Back Pain

Disability Questionnaire (OLBPDQ) or the Roland-Morris Low Back Pain Questionnaire (RMLBPQ) were used. A summary of the main characteristics is presented in Table 1 and Risk of Bias assessment is presented in Table 2.

The number of treatments given varied between studies (1 to 12), although two studies[33,44] did not report the total number. The two most common methods of SM were side lying, long lever rotational thrust or a supine thrust to the anterior pelvis with the participant rotated away from the lesion. Two studies adjusted anywhere along the spine.[33,45] None of the studies specified which joints were being targeted. Follow up times varied from no follow up to 10 months, with the majority adopting a two week follow-up.

Risk of Bias

Overall, only one study was assessed as having a high risk of bias.[33] The remaining eight studies were rated as low risk of bias as they achieved a score greater than 6 (see table 2). For sequence generation, six[43,45,46,48,49,80] of the nine trials were assessed as having low risk of bias, the risk of bias for the remaining three were unclear. For allocation concealment three were rated as low;[46,49,89] the majority were again unclear. With regards to blinding of participants, the majority of studies were assessed as low risk of bias due to the nature and quality of the sham procedure. Participants also acted as outcome assessors when using self-rating scales thus effective blinding of outcome assessor was possible. One study[45] assessed effectiveness of blinding post-intervention and found a higher number correctly guessed group assignment in the SM group. Bialosky et al (2014)[46] assessed believability of the placebo intervention and found that those receiving the placebo were more likely to believe their intervention was a sham than those in the SM or enhanced placebo group (63% V 33% $P<0.05$).

Kawchuk et al (2009)[47] had anesthetised all participants, so blinding was complete here. Blinding of the practitioner was impossible in all trials as they would be aware which type of manipulation they were performing.

The remaining criteria (selective outcome reporting, intention to treat analysis, co-interventions reported, compliance levels assessed) were all rated as low risk of bias.

Results from the pain outcome measures

Six studies[33,43,45,48,49,89] used a 100 mm VAS pain measure. Five reported an improvement in outcome, with SM groups showing lower levels of pain post-treatment and at follow up. One study[46] also used an 11 point numerical pain rating scale but no difference was found between groups.

Four studies[43,45,46,89] used the OLBPDQ to measure physical function due to pain levels. Two studies[43,45] reported an improved outcome with SM.

Senna et al (2011)[43] found differences between non-maintained SM and sham group at 1 and 4 months using the VAS but the mean pain score gradually returned to pre-treatment levels after the treatment was stopped (1 month). They also found evidence of a difference (p=0.005) using the OLBPDQ at 1 month follow up but no other time points. In contrast, the maintained SM group continued improving, indicating SM needs to be maintained to have a lasting effect.

Von Heymann et al (2013)[49] compared groups receiving SM (plus placebo medication) with a sham (plus placebo medication). However, no formal comparison was made between these arms. An interim analysis found the active treatments to be superior, after which the sham arm was dropped and the trial continued as a 2-arm study

comparing only the two active treatments (SM versus a non-steroidal anti-inflammatory drug “diclofenac”).

Waagen et al (1986)[33] found improvements in pain measured using the VAS in both the experimental group and the control group immediately after the intervention. At the two week post-treatment assessment there was evidence of reduced pain in the experimental group only.

One study[47] of just six participants, used an 11-point scale to measure pain. The authors report a greater proportion of the SM group experiencing less pain, however they do not report any formal analysis.

Triano et al (1995)[89] found evidence of a difference in functioning levels due to pain (OLBPDQ) between the three treatment arms at immediately post-treatment (p=0.012), with SM reporting the lowest scores. There was no difference between groups at the two week follow up. The VAS showed a similar pattern of results, although there was no longer evidence of an effect at two weeks.

Hoiiris et al (2004)[45] reported a decrease in pain and disability scores using the VAS and the OLBPDQ from baseline to 2 week follow up in all treatment arms. The SM group showing the greatest decline in scores. They found weak evidence of a difference between the change for each group (P= 0.087) using the OLBPDQ. Hadler et al (1987)[44] used the RMLBPQ to assess outcomes and reported evidence of an effect of SM among participants who had suffered with NSLBP between 2-4 weeks at the 3 day follow up but not at any other time point.

Drop outs were described and acceptable (<10%) in five studies. Four studies had high dropout rates (>10%)[33,43,49,89] three of which indicated the control group had the largest dropout rate.[33,43,49]

Adverse events

Only three trials reported on adverse events. Senna et al (2011)[43] reported that the most common adverse events were local discomfort and tiredness, which had resolved within 24 hours. The other two papers just stated that none were reported.[47,49]

The effect of the intervention

The effect of SM for NSLBP as measured by the 100mm VAS is presented in the summary of findings table (table 3). From four studies[43,45,48,89] (287 participants) the SMD is -0.36 (95%CI: -0.59,-0.12).The quality of evidence is graded as low due to high drop out in two studies,[43,89] broken blinding in one study,[45] no practitioner could be blinded in any study and only one[48] study conducted intention-to-treat (ITT) analysis.

Meta-analysis

Of the nine included studies, five[33,43,45,48,89] reported results of the VAS sufficiently for inclusion in a meta-analysis, with four included in the main meta-analysis and five included in either of the two sensitivity analyses.

Each of the following studies recorded information at either two week[45,89] or one month follow up.[43,48] These four studies were the only ones with sufficient information for inclusion in the main meta-analysis. After combining effect estimates using a random effects model, we found a pooled SMD of -0.36 (95% CI: -0.59,-0.12), corresponding to a reduction in pain among participants in the SM group at follow up. The I^2 statistic suggests no strong evidence against the assumption of homogeneity between effect estimates (I^2 <0.1%, p=0.835). However, given that there are only four studies included in this analysis, caution should be taken making inferences based on these analyses (see Fig 2).

Three studies[45,48,89] reported information collected immediately post-treatment and this was analysed in a sensitivity analysis (Fig A1). Waagan et al (1986)[33] was excluded from the main meta-analysis due to having high risk of bias score (see table 4), but included in a sensitivity analysis (Fig A3).

Fig 2. Forest plot of meta-analysis looking at pain scores of participants receiving SM vs sham SM treatment

Given the small number of studies included in this analysis, it is difficult to infer too much from the funnel plot, although there is no clear indication of small study effects (see Fig A5 in appendix).

Sensitivity analyses

The analysis run using the post-treatment pain scores show a similar pattern to the follow up scores, with a consistent direction of effect and an attenuated estimate (SMD = -0.35, 95% CI: -0.61, -0.08) (see Fig A1).

Analysis run including Waagen et al. (1986),[33] assessed to be high risk of bias, found results consistent with the main analysis (SMD = -0.37, 95% CI: -0.60, -0.14) (see Fig. A3).

Forest and funnel plots for the sensitivity analyses can be found in Appendix 4 (Figs A1-A5).

DISCUSSION

The objective of the present review was to systematically identify and critically evaluate the evidence from RCTs of spinal manipulation compared to an effective sham placebo on NSLBP. This is the first review to compare SM to an effective control. The review included nine studies of which four were included in the meta-analysis. The majority of trials used either a 100 mm VAS to assess pain levels or the OLBPDQ/RMLBPDQ to assess physical function due to NSLBP.

The results of the meta-analysis suggests a greater reduction in pain scores among participants receiving SM in comparison to those receiving an effective sham placebo. This finding remained consistent when looking at pain recorded at immediately post-treatment and follow up. The pooled effect estimate of -0.36 (95% CI: -0.59, -0.12) indicates that those receiving the SM had less pain (a mean of 0.36 standard deviations lower) than those in the control group. In terms of clinical relevance, this is only a small to moderate effect,[50,65] and the confidence intervals are wide. Caution is needed before drawing conclusions as most studies had some degree of risk of bias by failing to report on randomisation procedure or on allocation concealment.

Several methodological issues need to be considered. Seven trials which reported no evidence of between group differences may have lacked power; as sample sizes were small[33,43,44,46,47,48] and did not report *a priori* power calculations. Four studies had high dropout rates (>10%),[33,43,49,89] three of which indicated the control group had the largest dropout rate.[33,43,49] This could indicate dissatisfaction with sham as opposed to SM which may be an indicator of some treatment effect of SM. Just two studies described reasons for drop outs.[45,49] One study[46] included patients with any duration or type of NSLBP which again may have confounded the results.

Several studies had additional issues with the control group used, which might have contributed to the direction of results. Waagen et al (1986)[33] used massage as part of the control intervention; as massage has specific treatment (and contextual) effects[51] this may have reduced the observable difference between groups. Although the participants were screened for previous experience of SM (therefore justifying its inclusion in the review), this active control needs to be taken into consideration when evaluating the findings. Hoiiris et al (2004)[45] used an additional placebo medicine in both groups which may have lessened any relative difference.

Only four studies[44,45,46,49] attempted to standardise the interaction between patient and practitioner to reduce any placebo effect by way of contextual factors.[26,27,52,53] All other studies did not control for these variables, weakening their findings. There was much variation in number of treatments given and timing of outcome assessments between studies, making application to practice more difficult to establish. The majority of studies either had no follow up or just two weeks post intervention; a longer follow up would be required to ascertain long-term effectiveness of the intervention.

Limitations of the Review

Although the search strategy was comprehensive it is possible that some published clinical trials may not have been identified. However, our systematic and detailed search strategy make this unlikely, it is more likely that we did not identify eligible unpublished trials. Publication bias is a problem in all medical research[54] and it is a particular problem in alternative medicine.[55,56]

Furlan et al (2009)[36] recommend studying NSLBP in groups determined by the duration of symptoms as there are differences in the clinical course depending on the length of

time symptoms have been present. However, this was not possible in this review given the limited number of trials that met the inclusion criteria.

Deviations from the protocol

We conducted two post-hoc sensitivity analyses which were not planned or originally stated in the protocol but were deemed important once data was extracted. One was to include studies of high risk of bias[33] and the other was to see if there was a difference at immediately post–intervention compared to last follow up (using one month data rather than the 10 month follow up data for consistency[43]). There was very little variation in the findings. Functional outcomes (RMLBPDQ and OLBPDQ) due to pain levels were also extracted as it was deemed further indication of pain levels.

Adverse Events

Poor reporting of adverse events is a frequent criticism of complementary and alternative medicine (CAM) research.[57] Several previous reviews on complications of SM emphasise its safety,[58,59] however serious adverse events have been reported.[19] In this review few studies reported on adverse events at all. In one study,[47] the potential for adverse events to occur was higher as they used anaesthetic to ensure adequate blinding of participants; this procedure may be considered an unnecessary risk.

Recommendations for Future Research

Manual therapy practitioners are under pressure to produce evidence for their interventions.[16,17,60,61,62] Despite Ernst and Harkness' (2001)[17] call for more trials to

demonstrate the efficacy of SM for NSLBP, very few have been conducted that would satisfy the criteria of 'adequately designed sham-controlled trials'. In order to respond to this challenge this review suggests several directions of future research.

Treatment and sham interventions should be clearly specified, physically similar and matched for number, duration and interaction between subject and practitioner; these elements should be recorded. Co-interventions should also be avoided as they can distract from any benefits of specific treatment effects. The improvements in trial design would reduce the possibility of outcomes being due to non-specific effects. A standard measure should be used across studies to allow comparison of results and to facilitate formal pooling. A scale such as the VAS[63,64] has been shown to be reliable. A numerical rating scale measuring 0-10 has also been recommended[90] (NIH 2014). All adverse events should be recorded and reported. If no adverse events occurred this should be noted to allow accurate estimation of risks to participants.

Application to Practice

Several reviews have concluded suggesting that SM is no more or less effective than other treatments with proven benefits for NSLBP.[30,66] SM may carry a greater risk of adverse events, unlike non-thrust mobilisation[5] and massage.[51] Our review, however, found evidence for an effect of SM over effective control. There is currently insufficient evidence to inform practice. The decision to receive SM needs to be made by the patient who should be made aware of the current uncertainty.

Conclusions

There is some evidence from four of the nine trials (287 participants) that SM has specific treatment effects and is more effective at reducing NSLBP when compared to

an effective sham intervention. Although the effect was small-medium in terms of clinical relevance, a similar effect was found both immediately post-treatment and at follow up. Inconsistency of results across all studies may be due to the use of different interventions, controls, outcome measures and variable standards of methodology between studies. Currently, the evidence is insufficient to inform practice. Further adequately powered, well designed studies are required.

Funding: No funding was applied for.

Acknowledgements: With thanks to James Barclay (Librarian) for help with obtaining publications, Lucy Mackay-Tumber and Robin Kirk for general guidance. Thanks to the XXXXXX for providing financial contribution for XXXXX and XXXXX time.

Conflicts of interest: XXXXXXXXXXXX is a registered osteopath who uses spinal manipulation in the treatment of patients.

References

1. CBI/Pfizer. Healthy returns? Absence and Workplace Health Survey. 2010; http://www.cbi.org.uk/pdf/20100607-cbi-pfizer-absence-report.pdf,

2. Guo HR, Tanaka S, Halperin WE. Back pain prevalence in US industry and estimates of lost work days. *Am J Public Health* 1999; 89:1029-35.

3. Barnes PM, Bloom B, Nahin R. CDC *National Health Statistics Report #12*. Complementary and Alternative Medicine Use Among Adults and Children: United States, 2007. December 10, 2008.

4. National Council for Osteopathic Research, 2010. The Standardised Data Collection Project http://www.osteopathy.org.uk/resources/research/,

5. Seffinger M, Hruby R. *Evidence Based Manual Medicine – A Problem Based Approach*. 2007. Philadelphia, Elselvier.

6. Magee D. Zachazewski J, Quillen W. *Pathology and Intervention in Musculoskeletal Rehabilitation*, 2009. Missouri, Elselvier.

7. Bronfort G, Haas M, Evans R, Leininger B, Triano J. Effectiveness of manual therapies; the UK evidence report. *Chiropr Osteopat.* 2010 Feb 25;18:3.

8. Urquhart DM, Hoving JL, Assendelft WJ, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. *Cochrane Database of Systematic Reviews* 2008, Issue 1. Art No: CD001703.

9. Furlan AD, van Tulder MW, Cherkin D, Tsukayama H, Lao L, Koes BW, Berman BM. Acupuncture and dry-needling for low back pain. *Cochrane Database of Systematic Reviews* 2005, Issue 1. Art No: CD001351.

10. Ward R. C. (ed) American Osteopathic Association, 2003. *Foundations for Osteopathic Medicine* (2nd edition). Lippincott, Williams and Wilkins, Philadelphia.

11. Bogduk N, Jull G. The theoretical pathology of acute locked back; a basis for manipulative therapy. *Manual Medicine* 1985; 1:78-82.

12. Shekelle P. 1994. Spine Update; Spinal Manipulation. *Spine* 1994; 6:858-61.

13. Triano J. Biomechanics of Spinal Manipulative Therapy. *Spine Journal* 2001; 1: 121-130.

14. Pickar J, McLain R. Responses of mechanosensitive afferents to manipulation of the lumbar facet in the cat. *Spine* 1995; 20: 2379-2385.

15. Pickar, J. Neurophysiological effects of spinal manipulation. *Spine Journal* 2002; 2:357-71

16. Ernst, E., 2000. Does spinal manipulation have specific treatment effects? *Family Practice*, 2000; 17:554-6.

17. Ernst E, Harkness E. Spinal Manipulation: A Systematic Review of Sham-Controlled, Double-Blind, Randomized Controlled Trials. *Journal of Pain and Symptom Management,* 2001; 22 (4):879-89.

18. Evans D. Mechanisms and effects of spinal high-velocity, low-amplitude thrust manipulation: previous theories*. J Manipulative Physiol Ther*. 2002; 25:251–62.

19. Herbert J, Stronski N, Frensh, S, Rubenstein S. 2013. Serious adverse events and spinal manipulative therapy of the low back region: A systematic review of cases. *Journal of Manipulative and Physiological Therapeutics*.2013; doi:10.1016/j.jmpt.2013.05.009.

20. Koes B. How to evaluate manual therapy; value and pitfalls of randomised controlled trials. *Manual Therapy* 2004; 9:183-4.

21. Hancock M, Maher C, Latimer J, McAuley J. Selecting an appropriate placebo for a trial of spinal manipulative therapy. *Australian Journal of Physiotherapy* 2006; 52:135-8.

22. Maitland G. 1986. *Vertebral Manipulation* (5th Edition). London, Butterworths cited in Hancock M, Maher C, Latimer J, McAuley J. Selecting an appropriate placebo for a trial of spinal manipulative therapy. *Australian Journal of Physiotherapy* 2006; 52:135-138.

23. Fulda K, Slicho T, Stoll S. Patient expectations for placebo treatments commonly used in osteopathic manipulative treatment (OMT) clinical trials; a pilot study. *Osteopathic Medicine and Primary Care* 2007*;* 1:3

24. Hawk C, Long C. Use of a pilot to refine the design of a study to develop a manual placebo treatment. *Journal of Neuromuscuolskeletal Systems* 2000; 8:39-48.

25. Machado L, Kamper SJ, Herbert RD, Maher CG, McAuley JH. Imperfect placebos are common in low back pain trials: a systematic review of the literature. *European Spine Journal* 2008; 17: 889-904.

26. Kirsch W, Weixal L. Double-blind versus deceptive administration of a placebo. *Behavioural Neuroscience* 1988; 102:319-23.

27. Turner J, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of the placebo effects in pain treatment and research. *Journal of the American Medical Association* 1994; 271: 1609-14.

28. Benedetti, F., 2009. *Placebo Effects; understanding the mechanisms in health and disease*. New York, Oxford University Press.

29. Licciardone J, Brimhall A, King L. Osteopathic manipulative treatment for low back pain: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskeletal Disorders* 2005; 6(4):43-55.

30. Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews* 2011, Issue 2. Art. No.: CD008112.

31. Rubinstein SM, Terwee CB, Assendelft WJJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. *Cochrane Database of Systematic Reviews* 2012, Issue 9. Art. No.: CD008880.

32. Licciardone J, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W, Swift J Jr. Osteopathic manipulative treatment for chronic low back pain: a randomised controlled trial. *Spine* 2003; 28 (13):1355-62.

33. Waagen, G. Halderman S, Cook G, Lopez D, Deboer K. Short term trial of chiropractic adjustments for the relief of chronic low back pain. *Manual Medicine* 1986; 2:63-7.

34. Hondras M. Long CR, Cao Y, Rowell RM, Meeker WC. A randomised controlled trial comparing two types of spinal manipulation and minimal conservative medical care for adults 55 years and older with subacute or chronic low back pain. *Journal of manipulative Physiological Therapeutics* 2009; 32(5): 330-43

35. Hidalgo B, Detrembleur C, Hall T, Mahaudens P, Nielens H. The efficacy of manual therapy and exercise for different stages of non-specific low back pain: an update of systematic reviews. *Journal of Manual and Manipulative Therapy* 2014; 22(2), 59-74

36. Furlan A, Pennick V, Bombardier C, van Tulder M; Editorial Board, Cochrane Back Review Group. 2009 Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009; 34 (18):1929-41

37. Moher D, Liberati A, Tetzlaff J, Altman DG for the PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *BMJ* 2009; 339:b2535

38. Harris RM, Bradburn MJ, Deeks JJ, Harbord RM, Altman DG, Sterne J. metan: fixed- and random-effects meta-analysis. *Stata Journal* 2008; 8(1): 3-28.

39. Sterne JA, & Harbord RM. Funnel plots in meta-analysis. *Stata Journal* 2004; 4: 127-41.

40. Deeks JJ, Altman DG, Bradburn MJ. Statistical Methods for Examining Heterogeneity and Combining Results from Several Studies in Meta-Analysis. Systematic Reviews in Health Care. *BMJ* 2008; 285-312.

41. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychological Methods* 1998; 3(4): 486-504.

42. Higgins J, Thompson SG, Deeks, JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327:557-60.

43. Senna K, Machaly S. Does Maintained Spinal Manipulation Therapy for Chronic Nonspecific Low Back Pain Result in Better Long-Term Outcome? *Spine* 2011; 36 (18):1427–37.

44. Hadler N, Curtis P, Gillings D, Stinnet S. A benefit of Spinal Manipulation as Adjunctive Therapy for Acute Lower Back Pain: A Stratified Controlled Trial. *Spine 1987;* 12(7):702-6.

45. Hoiiris K, Pfleger B, McDuffie F, Cotsonis G, Elsangak O, Hinson R, Verzosa G. A randomized clinical trial comparing chiropractic adjustments to muscle relaxants for sub acute low back pain. *Journal of Manipulative and Physiological Therapeutics* 2004; 27 (6):388-98.

46. Bialosky J, George SZ, Horn ME, Price DD, Staud R, Robinson ME. Spinal manipulative therapy-specific changes in pain sensitivity in individuals with low back pain. *The Journal of Pain* 2014; 15 (2):136-148.

47. Kawchuk G, Haugen R, Fritz J. A true blind for subjects who receive spinal manipulative therapy. *Arch Phys Med Rehabil* 2009; 90:366-8.

48. Ghroubi S, Elleuch H, Baklouti S,Elleuch M. Chronic Low Back Pain and vertebral Manipulation. *Ann Readapt Med Phys.* 2007; 50(7):570-6.

49. Von Heymann V, Schloemer P, Timm J, Muehlbauer B. Spinal High-Velocity Low Amplitude Manipulation in Acute Nonspecific Low Back Pain. *Spine* 2013; 38 (7): 540-8

50. Cohen J. *Statistical power analysis for the behavioural sciences, 1st edition*. New York, San Francisco, London: Academic Press,1988:1-474

51. Furlan A, Immamura M, Dryden T, Irvin E. Massage for low back pain. *Cochrane Database of Systematic Reviews* 2008 Oct 8;(4):CD001929.

52. Halpern J. Empathy, using resonance emotions in the service of curiosity. *Journal of Gen Intern Med.* 2007; 22(5): 696–700.

53. Thomas K. General practice consultations; is there any point being positive? *BMJ* 1987*;* 294: 1200-2.

54. Easterbrook, P., Berlin, A. Gopalan, R, Matthews, R. Publication bias in clinical research. *Lancet* 1991; 337 (8746):867–72.

55. Ernst E, Pittler M. Alternative therapy bias. *Nature* 1997; 385(6616):480

56. Ernst E. Publication bias in complementary/alternative medicine. *J Clin Epidemiol* 2007a, 60(11):1093-4.

57. Ernst E. 'First, do no harm' with complementary and alternative medicine. *Trends Pharmacol Sci,* 2007b; 28(2):48-50.

58. Haldeman S, Kohlbeck F, McGregor M. Risk factors and precipitating neck movements causing vertebral artery dissection after cervical trauma and spinal manipulation. *Spine* 1999; 24:785-94.

59. Powell F, Hanigan W, Olivero W. A risk/benefit analysis of spinal manipulation therapy for relief of lumbar or cervical pain. *Neurosurgery* 1993; 33: 73-8.

60. Goldstein M. A challenge to the profession; initiate evidence-based osteopathic medicine now. *Journal of the American Osteopathic Association*, 1997; 97: 448.

61. Howell J. The paradox of osteopathy. *The New England Journal of Medicine* 1999; 341:1465-8.

62. Ernst E, Canter P. A Systematic Review of Systematic Reviews of Spinal Manipulation. *Journal of the Royal Society of Medicine 2006;* 99:189-93.

63. Scrimshaw S, Maher C. Responsiveness of Visual Analogue Scale and McGill Pain Scale Measures. *Journal of Manipulative and Physiological Therapeutics* 2001; 24 (8): 501-4.

64. Million R, Hall W, Nilsen KH, Baker RD, Jayson MI. Assessment of the progress of the back pain patient. *Spine* 1982; 7(3):204-12.

65. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org, [accessed 21/4/14]

66. Assendelft W, Morton S, Yu E, Suttorp M, Shekelle P. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med*, 2003; 138:898-906.

67. Clark B, Walkowski S, Conatser RR, Eland DC, Howell JN. Muscle functional magnetic resonance imaging and acute low back pain; a pilot study to characterize lumbar muscle activity asymmetries and examine the effects of osteopathic manipulative treatment. *Osteopathic Medicine and Primary Care*. 2009; 27(3):7.

68. Cleary C, Fox J. Menopausal symptoms; an osteopathic intervention. *Complementary Therapies in Medicine* 1994; 2:181-6.

69. Cleland J, Fritz J, Childs J, Kulig K. Comparison of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule; study protocol of a randomised clinical trial. *MBC Musculoskeletal Disorders* 2006; 10 (7):11.

70. Cote P, Mior S, Vernon, H. The short term effect of a spinal manipulation on pain/pressure threshold in patients with chronic mechanical low back pain. *Journal of Manipulative and Physiological Therapeutics* 1994; 17(6):364-8.

71. Cramer G, Gregerson DM, Knudsen JT, Hubbard BB, Ustas LM, Cantu JA.The effects of side-posture positioning and spinal adjusting on the lumbar Z joints; a randomised controlled trial with sixty-four subjects. *Spine* 2002; 15 (27): 2459-66.

72. Dishman J, Ball K, Burke J. Central Motor Excitability changes after spinal manioulation; a transcranial magnetic stimulation study. *Journal of Manipulative Physiological Therapeutics* 2002; 25 (1):1-9.

73. Hancock M, Maher CG, Latimer J, McLachlan AJ, Cooper CW, Day RO, Spindler MF, McAuley JH. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first line treatment for acute low back pain; a randomised controlled trial. *Lancet*, 2007; 10 (370):1638-43.

74. Hancock M. Maher CG, Latimer J, Herbert RD, McAuley JH. Independent evaluation of a clinical prediction rule for spinal manipulative therapy; a randomised controlled trial. *European Spine Journal* 2008; 17(7):936-43.

75. Hawk C, Long CR, Rowell RM, Gudavalli MR, Jedlicka J. A Randomised Trial Investigating a Chiropractic Manual Placebo, A Novel Design Using Standardised Forces in the delivery of active and control treatments. *Journal of Alternative and Complementary Medicine* 2005;11 (1):109.

76. Hoehler F, Tobis J, Beurger A. Spinal manipulation for low back pain. *Journal of the American Medical Association*, 1981; 245 (18):1835-8.

77. Hondras M, Long C, Brennan P. Spinal manipulative therapy versus a low force mimic manoeuvre for women with primary dysmenorrhoea; a randomised observer blinded trial. *Pain*, 1999; 81(1):105-114.

78. Kokjohn K, Schmidt D, Triano J, Brennan P. The Effect of Spinal Manipulation on Pain and Prostoglandin Levels in Women with Primary Dysmenorrhea. *Journal of Manipulative Physiological Therapeutics,* 1992;15 (5):279-85.

79. Krekoukias G, Petty N, Cheek L. Comparison of surface electromyographic activity of erector spinae before and after the application of central posteroanterior mobilisation on the lumbar spine. *Journal of Manipulative and Physiological Therapeutics* 2009;19 (1):39-45.

80. Learman, K. yers JB, Lephart SM, Sell TC, Kerns GJ, Cook CE. Effects of spinal manipulation on trunk proprioception in subjects with chronic low back pain during symptom remission. *Journal of Manipulative Physiological Therapeutics* 2009; 32 (2):118-26.

81. Licciardone J, Minotti DE, Gatchel RJ, Kearns CM, Singh KP. Osteopathic manual treatment and ultrasound therapy for chronic low back pain: a randomised controlled trial. Ann *Fam Med* 2013; 11(2):122-9.

82. Ongley M, Droman T, Klein R, Eek B, Hubert L. A new approach to low back pain. *Lancet*, 1987;2(855):143-6

83. Mandara A, Fusaro A, Musicco M, Bado F. A Randomised controlled trial on the effectiveness of osteopathic manipulative treatment of chronic low back pain. *International Journal of Osteopathic Medicine* 2008; 11:156.

84. Perry J, Green A. An investigation into the effects of a unilaterally applied lumbar mobilisation technique on peripheral sympathetic nervous system activity in the lower limbs. *Manual Therapy* 2008; 13 (6): 492-9.

85. Puetendura E, Landers MR, Hurt K, Meissner M, Mills J, Young D. Immediate effects of lumbar spine manipulation on the resting and contraction thickness of transverse abdominis in asymptomatic individuals. *J Orthop Sports Phys Ther.* 2011 Jan;41 (1):13-21

86. Roy R, Boucher J, Comtois A. Heart Rate variability modulation after manipulation in pain free patients vs patients with pain. *Journal of Manipulative and Physiological Therapeutics*, 2009; 32 (4): 277-86.

87. Santilli V, Behgi E, Finucci S. Chiropractic manipulation in the treatment of acute low back pain and sciatica with disc protrusion; a randomised double-blind clinical trial of active and simulated spinal manipulations. *Spine Journal*, 2006; 6(2):131-7.

88. Wreje U, Nordgren B, Aberg H. Treatment of pelvic joint dysfunction in primary care – a controlled study. *Scandinavian Journal of Primary Health Care* 1992: 10 (4): 310-15.

89. Triano J, McGregor M, Hondras M, Brennan P. Manipulative therapy Versus Education in Chronic Low back Pain. *Spine*, 1995; 20 (8): 948-53

90 Deyo R, Dworkin S, Amtmann D, Andersson G, Borenstein M, Carragee E. Report of the Task Force on Research. Standards for Chronic Low-Back Pain, 2014. http://painconsortium.nih.gov/NIH_Pain_Programs/Task_Force/cLBP_RTF_FullReport.pdf

TABLE 1: characteristics of selected studies

Study Country	Sample size (analysed) Intervention: Sham Age range/Mean age (SD)SE* Sex	Study setting/ participants	SM	Sham control	No. of treatments	Pain outcome measure Assessment schedule
Waagen *et al* (1986)[33] USA	N = 29 (19) **SM**: N = 11 Age = 25.2 (NR) **Sham**: N=18 Age = 24.3 (NR)	1st time patients at a chiropractic college clinic with pain of > 3 weeks duration. Patients naïve to chiropractic care	High velocity thrust to all levels of spine	Lumbar drop piece on chiropractic table set to mimic thrust, followed by soft tissue manipulation	2-3 treatments /week for 2 weeks with discrepancies between groups	Pain: 10cm VAS •BL •Post- treatment at 2weeks
Hadler *et al* (1987)[44] USA	N = 57 (54) Age = 18-40 **SM**: N = 26 Males = 18 **Sham**: N = 28 Males = 13	Pps experiencing NSLBP for 1st time and no longer than 1 month, groups stratified by NSLBP < 2 weeks and 2-4 weeks	Side-lying long lever rotational thrust to lumbar spine, no levels specified	Side-lying with both knees flexed, light thrust delivered to hips	Not specified FU every 3 days by phone for 2 weeks	Pain: RMLBPDQ •BL •Every 3 (±1) days from treatment (4 questionnaires in total)
Triano *et al* (1995)[89] USA	N= 209 (145)$ Age = 42.3 (14.3) Male= 83: Female = 62 **SM**: N= (47) **Sham**: N= (39) 107 in SM or control groups completed (43 in back education group)	Pps presenting to back clinic who had suffered > 50 days of NSLBP or had a history of > 6 episodes of NSLBP	Side-lying long lever rotational manipulation, no levels specified	Side-lying with both knees bent, thrust delivered to a supported area of the thoracic spine	7 or more, with discrepancies between groups	Pain: 10cm VAS, OLBPDQ •BL •Post-treatment •2 wk FU
Hoiiris *et al* (2004)[45] USA	N=192 enrolled $ N =103 in SM or control groups completed (53 in muscle relaxants group) **SM:** N=50 (34) Age = 42.2 (9.7) Males: 25 **Sham**: N=53 (40) Age = 43.1 (9.8) Males: 32	Pps had sub-acute NSLBP between 2-6 wks duration	Variable adjustments, prone or side-lying for all spine + placebo medicine	Prone or side-lying positioning with practitioner contact and motion with no thrust + placebo medicine	7 treatments for each group, over 2 weeks	Pain: 10 cm VAS, OLBPDQ • BL • post-treatment, • 2 wk FU
Ghroubi *et al* (2007)[48] France	N= 64 **SM**: N=32 Age = 39.06 (11.05) Males = 5 **Sham:** N=32 Age = 37.37 (7.51) Males = 8	Pps presented with first episode of NSLBP of ≤ 6 months	Spinal manipulation, no levels specified	Side lying "tensioning of the spine" without thrust	4 treatments for each group	Pain: 10cm VAS • BL • Post treatment • 1 month FU
Kawchuk *et al* (2009)[47] Canada	N = 6 Age = 36.5 (NR) Males = 4 **SM:** N=3 **Sham:** N=3	Pps with uncomplicated NSLBP <2 weeks duration currently receiving lumbar SM	Anesthetised for 3-5 mins and then received a single SMT to lumbar spine	Anesthetised for 3-5 mins	1 treatment	Pain: 11 point scale (0-10) Before anaesthetic and 30 mins after recovery

Senna *et al* (2011)[43] Egypt	N=67 **SM**: N=26 Age = 40.3(11.67) Male:19, Female:7, **Sham:** N=37 Age= 42 (9.66) Male:28, Female:9, **SM Maintained:** N=25, Age = 41 (11.03) Male:19, Female:6	Pps between 20-60 yrs suffering NSLBP >6 months.	SM + Maintained SM - Supine, patients side bent towards and rotated away from the lesion, a thrust forces applied to the anterior pelvis in a posterior and inferior direction. Followed by posterior pelvic tilt exercises.	SM techniques, which consisted of manually applied forces of diminished magnitude, aimed purposely to avoid treatable areas of the spine and to provide minimal likelihood of therapeutic effects. Followed by posterior pelvic tilt exercises	12 treatments over a 1 month period Maintained SM group– 2x month for 9 months.	Pain: 10cm VAS ,OLBPDQ • BL1 month (following 12 treatments), • 4 month FU • 7 month FU • 10 month FU
Von Heymann *et al* (2013)[49] Germany	N= 100 **SM:**38(33) Median age 34.14 (9.45) Male:24, Female:14 **Diclofenac** 37(33) median age 37.51(10.09) 23M, 14 F **Sham:** 25(14) median age 39.25(10.23) Male:13, Female:12	Pps between 18-55 yrs with NSLBP with duration <48hours, recruited from outpatient practices.	Side lying rotational thrust technique, no levels specified. + Placebo tablets	Patient prone, one leg tractioned, a cephalad impulse is delivered through the sacrum, on the opposite side to the sacrum. + Placebo tablets.	2-3 over 1 week	Pain: 10cm VAS, RMLBPDQ • BL • 7-9 days post intervention
Bialosky *et al* (2014)[46] USA	N=95 77 F, 33M Overall mean age 31.68 **SM:** 28 Male: 7 , Female:21 **Sham:**27 Male:10, Female:17 **Enhanced Sham SM:** 27 Male:7, Female:20) NO ITT – 28 (F19, M9)	Pps between 18-60 yrs, suffering NSLBP ≥ 4/10 over 24hrs on NRS	Supine, side bent towards and rotated away from the lesion, a thrust forces applied to the anterior pelvis in a posterior and inferior direction. No levels specified	Sham - Supine, no side bending, patient rotated away from the lesion then returned pre thrust, the thrust was delivered into the table. Enhanced sham– same physical procedure + suggestion to patient of the benefits of the sham procedure.	6 times over 2 weeks, each visit - SM: 2 each side Sham: 2 each side Enhanced Sham: 2 each side.	Unusual pain NRS (0-10) OLBPDQ • BL • at end of study (2 weeks duration)

SD = standard deviation; SE = standard error; Pps=participants; SM= Spinal manipulation; N = number; mins.= minutes; FU = follow up: NSLBP = non-specific low back pain; M = male; F = female; OLBPDQ – Oswestry Low back pain disability questionnaire; RMLBPDQ = Roland Morris Low Back Pain Disability Questionnaire; VAS = visual analogue scale; NRS= numerical rating scale; NR – not reported, BL – baseline; $ = 3 groups included in all analyses

Table 2: Risk of bias table (Cochrane Back Review Group, 2009)

	Was the method of randomization adequate?	Was the treatment allocation concealment successful?	Was the patient blinded to the intervention?	Was the care provider blinded to the intervention?	Was the outcome assessor blinded to the intervention?	Was the drop-out rate described and acceptable?	Were all randomised participants analysed in the group to which they were allocated?	Are reports of the study free of suggestion of selective outcome reporting?	Were the groups similar at baseline regarding the most important prognostic factors?	Were co-interventions avoided or similar?	Was the compliance acceptable in all groups?	Was the timing of the outcome assessment similar in all groups?	Total score (scores greater than 6 are considered low risk of bias)
Waagen 1986[33]	?	?	YES	NO	YES*	NO	NO	YES	YES	NO[e]	NO (66%)[j]	YES	5
Hadler 1987[44]	?	?	YES	NO	YES*	YES**	NO	YES	YES	YES	YES (95%)[j]	YES	8
Triano 1995[89]	YES	YES	YES	NO	YES*	NO	NO	YES	YES[h]	YES	YES (81%)[j]	YES	9
Hoiiris 2004[45]	YES	?	NO[i]	NO[a]	NO[i]*	YES	NO	YES	YES[b]	YES[c]	YES (79/82%)[k]	YES	7
Ghroubi 2007[48]	YES	?	YES	NO	YES*	YES	YES	YES	YES	YES	YES (100%)[j]	YES	10
Kawchuk 2009[47]	?	?	YES	NO	YES*	YES[f]	YES	YES	?	YES	YES (100%)[j]	YES	8
Senna 2011[43]	YES	?	YES	NO	YES*	NO described but unacceptable	NO	YES	YES	YES	YES (94%)	YES	8
Von Heymann 2013[49]	YES	YES	YES	NO	YES*	NO	NO	?	YES	YES	NO (75%)	YES	7
Bialosky 2014[46]	YES	YES	YES	NO	NO	YES	YES	YES	YES	YES	YES (100%)	YES	10

* Outcome assessor is participant when rating self-report scales like the VAS, RMLBPDQ (Roland Morris Low Back Pain Disability Questionnaire), OLBPDQ (Oswestry Low back pain disability Questionnaire)

** <10%; YES =1, NO = 0, ? = unclear

a = although authors claim the chiropractor was blinded – this would be impossible; b = between the Intervention and sham; c= both also received placebo medicine; d = differences in number of treatments over differing duration periods; e= soft tissue performed in sham group only; f = no dropouts; g= no mention of analysis; h= some analysis of height and weight; i= blinding of participant was tested and perception of true chiropractic care was sig. higher in chiropractic group (P<0.05); j = the authors did not report compliance directly so we have inferred compliance from people completing the treatment programme; k = based on medication logs or kits respectively.

Table 3: Summary of findings table: Spinal manipulation versus an effective sham for non-specific low back pain

Patient or population: Individuals with NSLBP

Settings: Clinic

Intervention: Spinal manipulation using high velocity/low amplitude thrust

Comparison: Effective sham manipulation

Outcomes	Illustrative comparative risks* (95% CI)		No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk			
	Control group	Intervention group			
Pain (as measured by a 100mm VAS) Follow up 2 weeks to 1 months)		The mean pain symptomology (continuous) in the intervention groups was **0.36 standard deviations lower** (0.59 to 0.12 lower)	287 (4)	⨁⨁⊖⊖[1] **low**	Low risk of bias in outcome reporting as participants were blinded effectively[2] Small to moderate SMD = -0.36 (95% CI: -0.59, -0.12)

CI = Confidence Interval

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low quality: We are very uncertain about the estimate

[1] Judgements of low risk of bias (>6) in all studies included in the main meta-analysis, however high level of dropouts in 2 studies, no ITT analysis in 3 studies, all practitioners could not be blinded
[2] The sham manipulation ensured blinding of participants, although one[45] tested blinding and it is possible it may have been broken

Table 4 – Results of VAS pain scores included in meta-analysis

	Post-treatment				**Follow up**			
	Intervention		Control		Intervention		Control	
Study	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
1. Waagen et al. (1986)	-	-	-	-	9	23 (15)	10	31 (15)
2. Triano et al. (1995)	47	13.9 (15.3)	39	19.8 (18.3)	47	13.3 (15.9)	39	21.7 (24.4)
3. Hoiiris et al. (2004)	34	2.44 (2.22)	40	3.18 (2.4)	34	1.71 (1.88)	40	2.21 (2.02)
4. Ghroubi et al. (2007)	32	49.37 (16.78)	32	58.43 (28.8)	32	48.13 (22.78)	32	54.43 (25.76)
5. Senna et al. (2011)	-	-	-	-	26	29.5 (6.03)	37	33.2 (7.53)

Appendix 1 - Search Strategy

Terms recommended by Furlan et al (2009)

1. exp Back Pain/ or exp Low Back Pain/
2. exp Lumbar Vertebrae/
3. Zygapophyseal Joint/
4. (back adj3 pain).ti,ab.
5. (low* adj3 back adj3 pain*).ti,ab.
6. (lumbar adj3 vertebrae*).ti,ab.
7. ((backache or back) adj3 ache).ti,ab.
8. lumbago*.ti,ab.
9. (facet adj3 joint*).ti,ab.
10. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9
11. exp Manipulation, Spinal/ 12. exp Manipulation, Osteopathic/
13. (sham adj3 manipulation*).ti,ab.
14. (spin* adj3 manipulation*).ti,ab.
15. (osteopath* adj manipul*).ti,ab.
16. (high adj3 velocit* thrust).ti,ab.
17. (spin* adj3 adjust*).ti,ab.
18. 11 or 12 or 13 or 14 or 15 or 16 or 17
19. exp Randomized Controlled Trial/
20. exp Controlled Clinical Trial/
21. exp Random Allocation/
22. random$ allocat$.ti,ab.
23. (randomi?ed adj3 controlled adj3 trial).ti,ab.
24. (controlled adj3 clinical adj3 trial).ti,ab.
25. random$.ti,ab.
26. placebo$.ti,ab.
27. exp Placebos/
28. exp Clinical Trial/
29. trial.ti,ab.
30. group$.ti,ab.
31. 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30
32. 10 and 18 and 31

Appendix 2 - Risk of Bias Tool (Furlan et al, 2009)

A	1	Was the method of randomization adequate?	yes/no/unsure
B	2	Was the treatment allocation successful?	yes/no/unsure
C		Was the knowledge of the allocated interventions adequately prevented during the study?	
	3	Was the patient blinded to the intervention?	yes/no/unsure
	4	Was the care provider blinded to the intervention?	yes/no/unsure
	5	Was the outcome assessor blinded to the intervention?	yes/no/unsure
D		Were incomplete outcome data adequately described?	
	6	Was the drop out rate described and acceptable?	yes/no/unsure
	7	Were all randomised participants analysed in the group to which they were allocated?	yes/no/unsure
E	8	Are reports of the study free of suggestion of selective outcome reporting?	yes/no/unsure
F		Other sources of potential bias	
	9	Were the groups similar at baseline regarding the most important prognostic factors?	yes/no/unsure
	10	Were co-interventions avoided or similar?	yes/no/unsure
	11	Was the compliance acceptable in all groups?	yes/no/unsure
	12	Was the timing of the outcome assessment similar in all groups?	yes/no/unsure
		Total score = no of yes answers /12	

The risk of bias for RCTs was assessed using the criteria list recommended in the Updated Method Guidelines for Systematic Reviews in the Cochrane Collaboration Back Review Group.

Scores of 6 or more were considered low risk of bias

Appendix 3 – Studies excluded

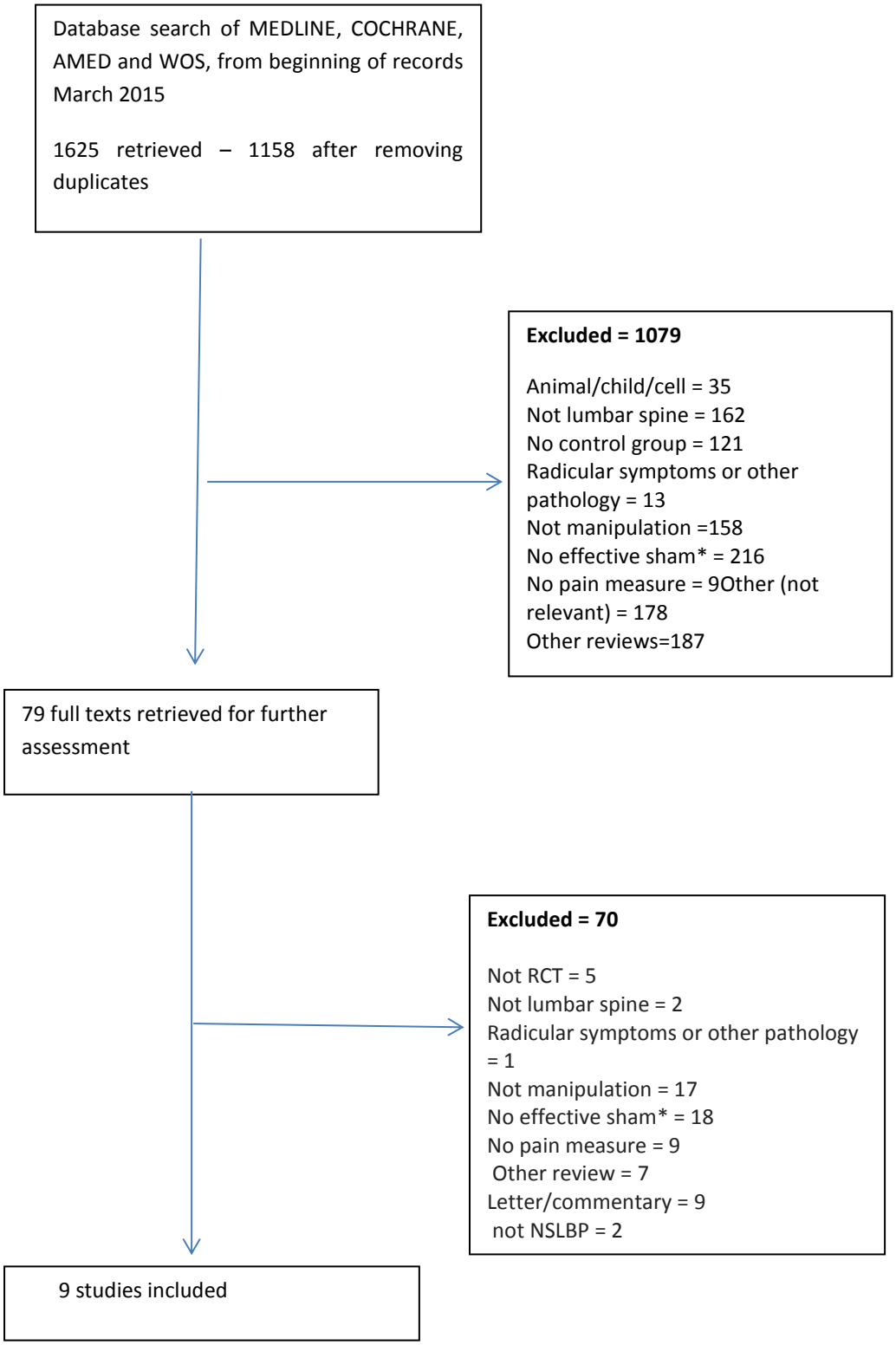
STUDY	REASON EXCLUDED
Clark et al 2009[67]	No sham manipulation in control group.
Cleary and Fox 1994[68]	Fox's 'low force osteopathic technique' does not match the inclusion criteria for an HVLA manoeuvre
Cleland et al 2006[69]	No sham control, comparison groups are alternative manipulation or mobilisation.
Cote et al 1994[70]	No sham control, comparison group is a mobilisation
Cramer et al 2002[71]	No sham control, comparison group is side lying positioning
Dishman et al 2002[72]	No sham control, comparison group is side lying positioning
Hancock et al 2007[73]	No sham control, placebo is detuned ultrasound therapy
Hancock et al 2008[74]	No sham control, placebo is detuned ultrasound therapy
Hawk et al 2005[75]	Excluded due to paper's own assessment of inadequate blinding
Hoehler et al 1981[76]	No sham intervention, massage was used as the control.
Hondras et al 1999[77]	Participants not generalizable
Hondras et al 2009[34]	No sham control, comparison groups are low force manipulations, minimal medical care and exercise therapy
Kokjohn et al 1992[78]	Participants not generalizable
Krekoukias et al 2009[79]	No sham control, comparison groups are prone lying and prone lying with touch onto L3 spinal level

Learman et al 2009[80]	Although the subjects were extensively screened for pain levels at entry to the study, no follow up data measuring pain was assessed.
Licciardone et al 2003[32]	Several different and non-standardised interventions (muscle energy techniques, soft tissue manipulation, fascial manipulation and cranio-sacral) were made both in the treatment group, the sham group received 'fake' treatments in the same modalities.
Licciardone et al 2013[81]	Several different and non-standardised interventions (muscle energy techniques, soft tissue manipulation, fascial manipulation and cranio-sacral) were made both in the treatment group, the sham group received 'fake' treatments in the same modalities.
Ongley et al 1987[82]	As well as a spinal manipulation, painkilling injections were being administered, in the control group the amount of painkilling injection was lowered therefore influencing reported pain levels
Mandara et al 2008[83]	No full data available, abstract is published as a conference presentation, repeated attempts were made to contact the authors with no response.
Perry and Green 2008[84]	No measurement of pain as an outcome
Puetendura et al 2010[85]	No measurement of pain as an outcome
Roy et al 2009[86]	although groups were divided into pain and pain free, no measurement of pain was taken.
Santilli et al 2006[87]	Radicular symptoms present

Appendix 4 – Forest and funnel plots from sensitivity analyses

A1-A5

Fig 1.



Effective sham as described in the introduction.
RCT – randomised controlled trial; NSLBP – non-specific low back pain

Fig 2. Forest plot of meta-analysis looking at pain scores of participants receiving SM vs sham SM treatment

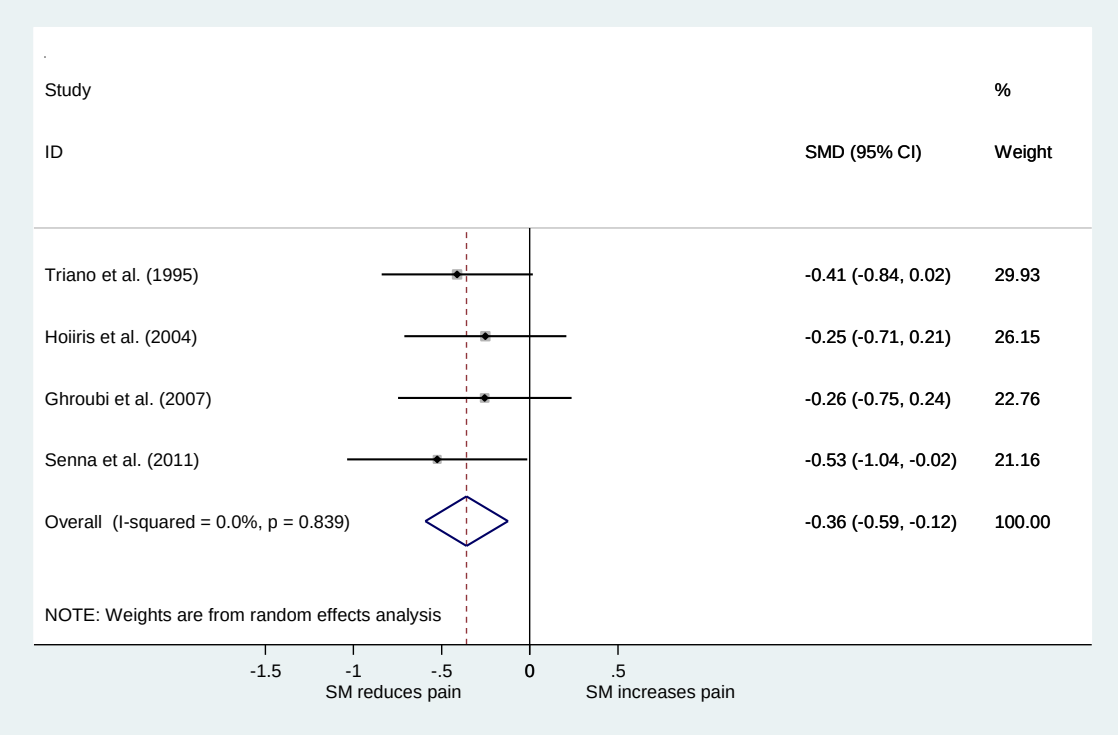


Fig A1. Forest plot of meta-analysis looking at pain scores of participants receiving SM vs sham SM treatment when assessing pain scores immediately post-treatment

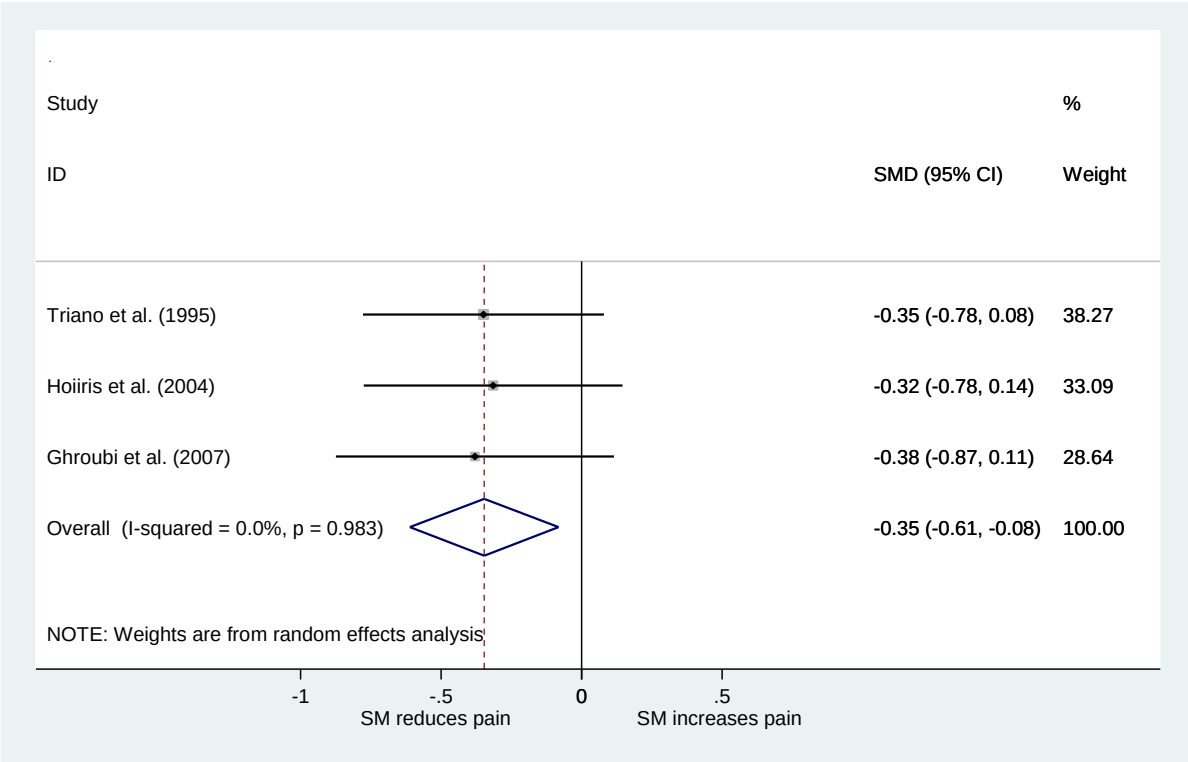


Fig A2. Forest plot of meta-analysis comparing SM vs sham SM treatment as assessed at follow up when including Waagen et al. (1986)

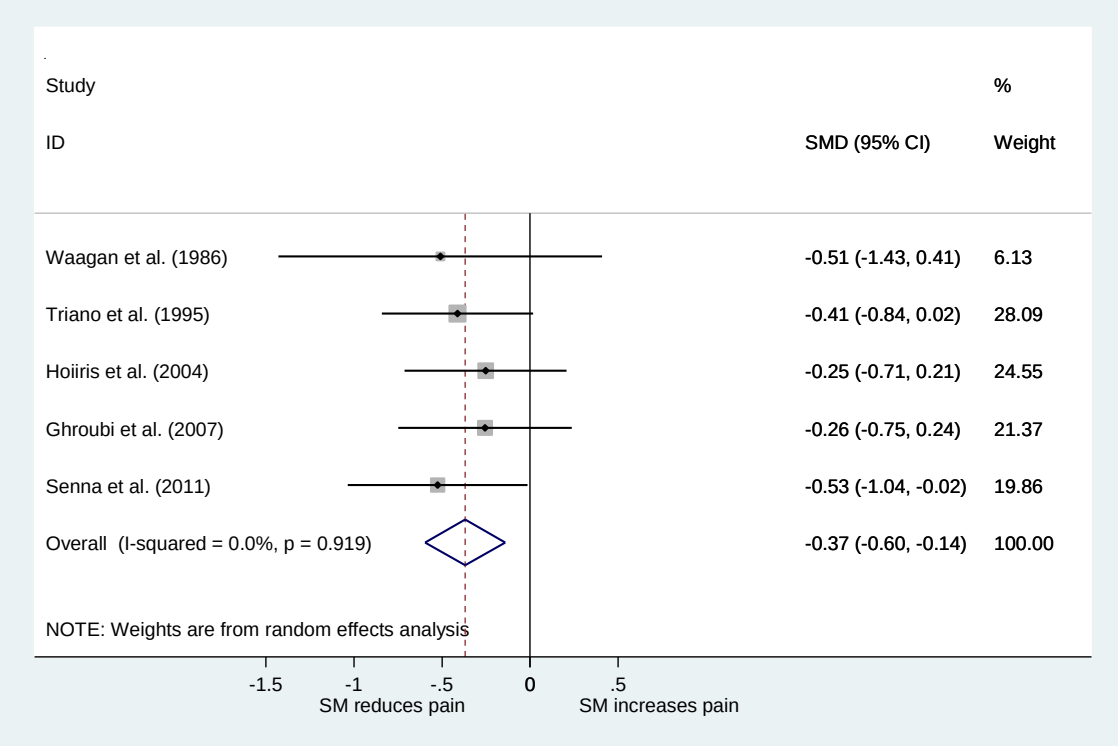


Fig A3. Funnel plot of studies comparing SM vs sham SM treatment when assessing pain scores immediately post-treatment

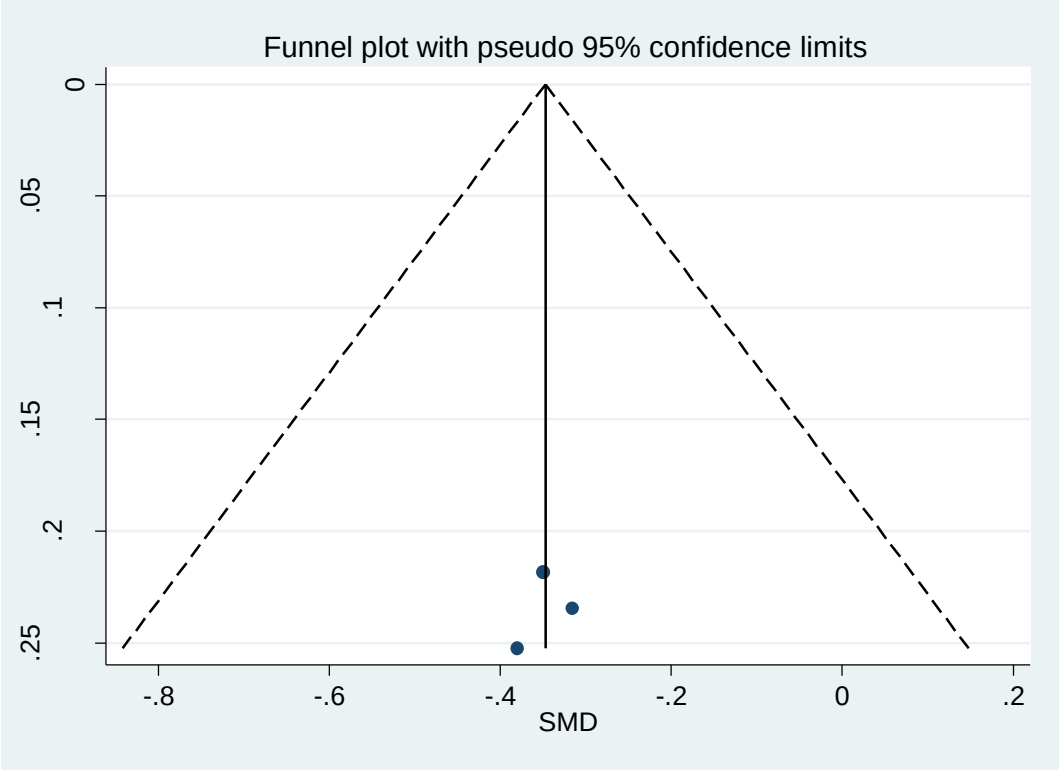


Fig A4. Funnel plot of studies comparing SM vs sham SM treatment as assessed at follow up when including Waagen et al. (1986)

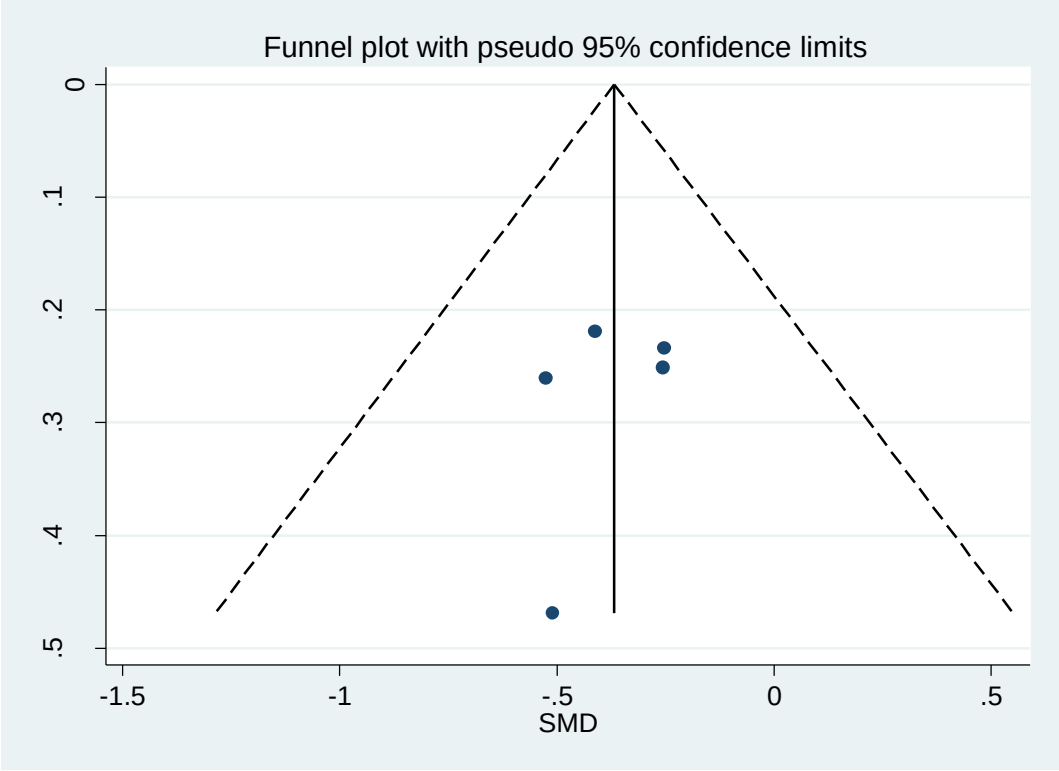


Fig A5. Funnel plot of studies comparing SM vs sham treatment