ORIGINAL RESEARCH

Spinal Manipulation Vs Sham Manipulation for Nonspecific Low Back Pain: A Systematic Review and Meta-analysis

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

ABSTRACT

Objective: The purpose of this systematic review was to identify and critically evaluate randomized controlled trials of spinal manipulation (SM) vs sham manipulation in the treatment of nonspecific low back pain.

Methods: Four 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 were extracted by 2 independent reviewers, and risk of bias was assessed using the Cochrane Back Review Group Risk of Bias tool.

Results: Nine randomized controlled trials were included in the systematic review, and 4 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 (standardized mean difference = −0.36; 95% confidence interval, −0.59 to −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, and some studies did not conduct intention-to-treat analysis and had a high level of dropouts.

Conclusion: There is some evidence that SM has specific treatment effects and is more effective at reducing nonspecific low back pain when compared with an effective sham intervention. However, given the small number of studies included in this analysis, we should be cautious of making strong inferences based on these results. (J Chiropr Med 2016;15:165-183)

Key Indexing Terms: *Low back pain; Manipulation; Spinal; Placebo; Review of the literature*

INTRODUCTION

Over the course of a lifetime, approximately 80% of people will experience low back pain (LBP). Nonspecific 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]

Nonspecific low back pain 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 nonspecific "placebo effect." Issues around safety of the technique have also been raised. NSLBP is characterized by pain in the posterior lumbar spine, sacral spine, or paraspinal tissues which may be accompanied by decreased range of motion.[5] The etiology is unclear, and a definitive cause remains elusive for researchers.[6] Several different approaches to

[a] Private Practice, Align Body Clinic, Bath, UK.
[b] MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK.
[c] Centre for Academic Mental Health, School of Social and Community Medicine University of Bristol, Bristol, Bristol, UK.
[d] The NIHR Biomedical Research Unit in Nutrition, Diet and Lifestyle at the University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK.
Corresponding author: Rachel E. Perry, BA, MA, MPhil, Bristol Nutrition BRU, Education & Research Centre, Level 3, Upper Maudlin St, Bristol, BS2 8AE.
(e-mail: *jay@alignbodyclinic.co.uk*, *Rachel.Perry@bristol.ac.uk*).
Paper submitted November 24, 2015; in revised form January 19, 2016; accepted January 20, 2016.
1556-3707

http://dx.doi.org/10.1016/j.jcm.2016.04.014

treatment have been identified, with mixed evidence for their success.[7–9] One of the treatments widely used is SM.

SM and Mobilization

Spinal manipulation 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 hypoanalgesic effect, either by structural[10–12] or neurological processes,[13,14] whereas others have postulated that it acts through nonspecific or "placebo" effects.[15,16]

In contrast, spinal mobilization uses low-velocity/low-amplitude cyclical techniques (nonthrust mobilization). It has been argued that this method of action differs from that of HVLA techniques; thus, mobilization and manipulation should be investigated separately.[17] SM can have serious (although very rare) adverse outcomes such as intervertebral disk prolapse and fracture,[18] whereas there are no reported adverse events reported from receiving nonthrust spinal mobilization.[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

To exclude possible placebo effects in trials of SM, the control group must either be screened for previous experience of SM[19] or be exposed to an effective sham intervention.

There is little agreement among experts as to what constitutes an effective sham manipulation.[20] However, there is some evidence as to what may be acceptable as an effective sham manipulation of the lumbar spine. Hancock et al (2006)[20] demonstrated that the most credible sham procedure was Maitland's "log roll."[21] 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)[22] 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)[23] and Machado et al (2008)[24] also identified the importance of equalization of the nonspecific effect of physical touch between participants. The use of an indistinguishable placebo should counteract any subtle differences between groups shown to influence treatment outcomes.[25,26] Other active therapies are not considered a viable control because they can lead to erroneous interpretation due to varied contextual factors which produce a placebo effect or specific treatment effects.[27] Thus, for a sham manipulation to be an effective control, it should physically resemble an 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 articles that did not use an effective sham[7] or permitted techniques that were not solely HVLA.[7,28–30]

The review of Bronfort et al (2010)[7] compared SM to a sham intervention. However, one included study[31] used an inappropriate sham intervention by using gluteal massage. The review of Rubenstein et al (2011)[29] of SM for chronic low back pain included 1 study[32] which used several techniques (HVLA, muscle energy techniques, soft tissue manipulation, fascial manipulation, and craniosacral) in their treatment group. The review of Rubenstein et al (2012)[30] of SM for acute back pain only included 1 study of SM vs a sham intervention. None of these reviews distinguished between SM and mobilization.

The review of Ernst and Harkness (2001)[16] of SM for a range of conditions identified 3 trials[33–35] and recommended that "the specific efficacy of SM for low back pain must await adequately designed sham-controlled trials." The most recent systematic review[36] examined SM, mobilization, and exercise as separate interventions against shams in NSLBP sufferers of various durations. Five studies were analyzed in groups determined by similarity of patients, interventions, comparisons and outcomes. However, no meta-analysis was performed.

This systematic review critically evaluates data from randomized controlled trials (RCTs) using HVLA techniques for people with NSLBP. The purpose of this study was 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 (Appendix A 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/, registration 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 2 reviewers (JR, RP). Disagreements were resolved through discussion (Fig 1). The following inclusion criteria were predefined:

- Type of participant: participants of either sex and >18 years with NSLBP.
- Type of intervention: 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, that is, the physical act of the sham manipulation must be credible; a sham that physically resembles the act of SM to minimize the differences between groups.
- Type of outcome: studies that had a perceived measure of pain as an outcome (eg, visual analogue scale [VAS] pain scores, standardized questionnaires)

We excluded studies that were not randomized and those that used participants with radicular symptoms; had a 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 B.

Only completed RCTs were included (reports of ongoing trials were excluded [eg, protocol articles]). The primary outcome was any measure of pain (both standardized and nonstandardized). The secondary outcome was any adverse event mentioned. Data from included studies were extracted independently by 2 reviewers (JR, RP) using a form with predefined criteria.

The risk of bias (RoB) of all included RCTs was evaluated independently by 2 reviewers (JR, RP) using the Risk of Bias tool of the Cochrane Back Review Group (CBRG)[37] (Appendix C). 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 third reviewer (HS). The manuscript was developed using the Preferred Reporting Items for Systematic Reviews and Meta-analysis checklist.[38]

Meta-analysis

Meta-analyses were performed in Stata 13 (Stata, College Station, TX) using the user-contributed commands metan[39] and metafunnel.[40] Standardized mean differences (SMDs) and corresponding standard errors were calculated for included studies using means, standard deviations, and sample sizes reported in the relevant publications. Because of the small sample sizes of some studies, Hedges' g, an extension of Cohen's d adjusted for small sample bias, was calculated.[41] 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.[42] 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.[43]

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 maximize comparability, the 1-month follow-up data from the study of Senna et al (2011)[44] were 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 posttreatment. A further sensitivity analysis investigated the effect of including studies assessed as being at a high RoB.[33]

Dealing With Missing Data

Where possible, we extracted the number of participants randomized 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.[44] In 1 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, 9 RCTs were identified that met the inclusion criteria for this review. The studies

Table 1. *Characteristics of Selected Studies*

Study Country	Sample Size (Analyzed) Inte 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 wk duration. Patients naive 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/ wk for 2 wk with discrepancies between groups	Pain: 10-cm VAS • BL • Posttreatment at 2 wk
Hadler et al (1987)[45] USA	N = 57 (54) Age = 18-40 SM: n = 26 Male = 18 Sham: n = 28 Male = 13	Pps experiencing NSLBP for 1st time and no longer than 1 mo, groups stratified by NSLBP < 2 wk and 2-4 wk	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 d by phone for 2 wk	Pain: RMLBPDQ • BL • Every 3 (±1) d from treatment (4 questionnaires in total)
Triano et al (1995)[35] USA	N = 209 (145) [a] 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 d 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: 10 cm VAS, OLBPDQ • BL • Posttreatment • 2-wk FU
Hoiiris et al (2004)[46] USA	N = 192 enrolled [a] n =103 in SM or control groups completed (53 in muscle relaxants group) SM: n = 50 (34) Age = 42.2 (9.7) Male: 25 Sham: n = 53 (40) Age = 43.1 (9.8) Male: 32	Pps had subacute NSLBP between 2 and 6 wk 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 wk	Pain: 10-cm VAS, OLBPDQ • BL • Posttreatment, • 2-wk FU
Ghroubi et al (2007)[49] France	N = 64 SM: n = 32 Age = 39.06 (11.05) Male = 5 Sham: n = 32 Age = 37.37 (7.51) Male = 8	Pps presented with first episode of NSLBP of ≤ 6 mo	SM, no levels specified	Side-lying "tensioning of the spine" without thrust	4 treatments for each group	Pain: 10-cm VAS • BL • Posttreatment • 1-mo FU
Kawchuk et al (2009)[48] Canada	N = 6 Age = 36.5 (NR) Male = 4 SM: n = 3 Sham: n = 3	Pps with uncomplicated NSLBP <2 wk duration currently receiving lumbar SM	Anesthetized for 3-5 min and then received a single SMT to lumbar spine	Anesthetized for 3-5 min	1 treatment	Pain: 11-point scale (0-10) Before anesthetic and 30 min after recovery
Senna et al (2011)[44] Egypt	N = 67 SM: n = 26 Age = 40.3(11.67) Male:19, Female:7, Sham: n = 37	Pps between 20 and 60 y suffering NSLBP >6 mo	SM + maintained SM: supine, patients side bent toward	SM techniques, which consisted of manually applied forces of diminished	12 treatments over a 1-mo period Maintained SM group:	Pain: 10-cm VAS, OLBPDQ

Table 1. (*continued*)

Study Country	Sample Size (Analyzed) Inte Range/Mean Age (SD) SE* Sex	Study Setting/ Participants	SM	Sham Control	No. of Treatments	Pain Outcome Measure Assessment Schedule
	Age = 42 (9.66) Male:28, Female:9, SM maintained: n = 25, Age = 41 (11.03) Male:19, Female:6		and rotated away from the lesion, a thrust force applied to the anterior pelvis in a posterior and inferior direction. Followed by posterior pelvic tilt exercises.	magnitude, aimed purposely to avoid treatable areas of the spine and to provide minimal likelihood of therapeutic effects. Followed by posterior pelvic tilt exercises	2× mo for 9 mo	• BL 1 mo (following 12 treatments), • 4-mo FU • 7-mo FU • 10-mo FU
Von Heymann et al (2013)[50] 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) 23 M, 14 F Sham: 25 (14) Median age 39.25(10.23) Male: 13, female: 12	Pps between 18 and 55 y with NSLBP with duration <48 h, 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 wk	Pain: 10-cm VAS, RMLBPDQ • BL • 7-9 d postintervention
Bialosky et al (2014)[46] USA	N = 95 77 F, 33 M 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 (F 19, M 9)	Pps between 18 and 60 y, suffering NSLBP ≥ 4/10 over 24 h 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 wk, 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-wk duration)

BL, baseline; *F*, female; *FU*, follow-up; *M*, male; mins, minutes; *N*, number; *NR*, not reported; *NRS*, numerical rating scale; *NSLBP*, non-specific low back pain; *OLBPDQ*, Oswestry Low back pain disability questionnaire; *Pps*, participants; *SD*, standard deviation; *RMLBPDQ*, Roland Morris Low Back Pain Disability Questionnaire; *SE*, standard error; *SM*, spinal manipulation.

[a] Three groups included in all analyses.

were published from 1986 to 2014. Five originated from the United States,[33,35,45–47] 1 from Canada,[48] 1 from France,[49] 1 from Egypt,[44] and 1 from Germany.[50] Eight were in English and 1 in French.[49] 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 to 65 years. The mean age of participants was incalculable because of incomplete reporting in 1 of the trials.[45] Four types of pain outcome measure were used. To assess pain levels directly, either a VAS or numerical rating scale (NRS) was used. To assess physical function due to pain, either the Oswestry Low Back Pain Disability Questionnaire (OLBPDQ) or the

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 Dropout Rate Described and Acceptable?	Were All Randomized Participants Analyzed 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 RoB)
Waagen 1986[33]	?	?	Yes	No	Yes [a]	No	No	Yes	Yes	No [f]	No (66%) [j]	Yes	5
Hadler 1987[45]	?	?	Yes	No	Yes [a]	Yes [b]	No	Yes	Yes	Yes	Yes (95%) [j]	Yes	8
Triano 1995[35]	Yes	Yes	Yes	No	Yes [a]	No	No	Yes	Yes [h]	Yes	Yes (81%) [j]	Yes	9
Hoiiris 2004[46]	Yes	?	No [i]	No [c]	No[a,i]	Yes	No	Yes	Yes [d]	Yes [e]	Yes (79/82%) [k]	Yes	7
Ghroubi 2007[49]	Yes	?	Yes	No	Yes [a]	Yes	Yes	Yes	Yes	Yes	Yes (100%) [j]	Yes	10
Kawchuk 2009[48]	?	?	Yes	No	Yes [a]	Yes [g]	Yes	Yes	?	Yes	Yes (100%) [j]	Yes	8
Senna 2011[44]	Yes	?	Yes	No	Yes [a]	No (described but unacceptable)	No	Yes	Yes	Yes	Yes (94%)	Yes	8
Von Heymann 2013[50]	Yes	Yes	Yes	No	Yes [a]	No	No	?	Yes	Yes	No (75%)	Yes	7
Bialosky 2014[47]	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes (100%)	Yes	10

[a] Outcome assessor is participant when rating self-report scales like the VAS, RMLBPDQ, or OLBPDQ.
[b] <10%; yes = 1, no = 0, ? = unclear.
[c] Although authors claim the chiropractor was blinded, this would be impossible.
[d] Between the intervention and sham.
[e] Both also received placebo medicine.
[f] Soft tissue performed in sham group only.
[g] No dropouts.
[h] Some analysis of height and weight.
[i] Blinding of participant was tested, and perception of true chiropractic care was significantly higher in chiropractic group ($P < .05$).
[j] The authors did not report compliance directly, so we have inferred compliance from people completing the treatment program.
[k] Based on medication logs or kits, respectively.

Roland-Morris Low Back Pain Questionnaire (RMLBPQ) was used. A summary of the main characteristics is presented in Table 1, and RoB assessment is presented in Table 2.

The number of treatments given varied between studies,[1–11,51] although 2 studies[33,45] did not report the total number. The 2 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.[33,46] 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 2-week follow-up.

Risk of Bias

Overall, only 1 study was assessed as having a high RoB.[33] The remaining 8 studies were rated as having low RoB because they achieved a score greater than 6 (Table 2). For sequence generation, 6[44,46,47,49,50,52] of the 9 trials were assessed as having low RoB; the RoB for the remaining 3 was unclear. For allocation concealment, 3 were rated as low[35,47,50]; the majority were again unclear. With regard to blinding of participants, the majority of studies were assessed as having low RoB because of 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[46] assessed effectiveness of blinding postintervention and found that a higher number correctly guessed group assignment in the SM group. Bialosky et al (2014)[47] assessed believability of the placebo intervention and found that those receiving the placebo were more likely to believe that their intervention was a sham than those in the SM or enhanced placebo group (63% vs 33%, $P < .05$). Kawchuk et al (2009)[48] had anesthetized all participants, so blinding was complete here. Blinding of the practitioner was impossible in all trials because they would be aware which type of manipulation they were performing.

The remaining criteria (selective outcome reporting, intention-to-treat [ITT] analysis, co-interventions reported, compliance levels assessed) were all rated as having low RoB.

Results From the Pain Outcome Measures

Six studies[33,35,44,46,49,50] used a 100-mm VAS pain measure. Five reported an improvement in outcome, with SM groups showing lower levels of pain posttreatment and at follow-up. One study[47] also used an 11-point numerical pain rating scale, but no difference was found between groups.

Four studies[35,44,46,47] used the OLBPDQ to measure physical function due to pain levels. Two studies[44,46] reported an improved outcome with SM.

Senna et al (2011)[44] found differences between nonmaintained SM and sham group at 1 and 4 months using the VAS, but the mean pain score gradually returned to pretreatment levels after the treatment was stopped (1 month). They also found evidence of a difference ($P = .005$) using the OLBPDQ at 1 month follow-up but no other time points. In contrast, the maintained SM group continued improving, indicating that SM needs to be maintained to have a lasting effect.

Von Heymann et al (2013)[50] 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 2 active treatments (SM vs a nonsteroidal 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 2-week posttreatment assessment, there was evidence of reduced pain in the experimental group only.

One study[48] of just 6 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)[35] found evidence of a difference in functioning levels due to pain (OLBPDQ) between the 3 treatment arms immediately posttreatment ($P = .012$), with SM reporting the lowest scores. There was no difference between groups at the 2-week follow-up. The VAS showed a similar pattern of results, although there was no longer evidence of an effect at 2 weeks.

Hoiiris et al (2004)[46] 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 showed the greatest decline in scores. They found weak evidence of a difference between the change for each group ($P = .087$) using the OLBPDQ. Hadler et al (1987)[45] used the RMLBPQ to assess outcomes and reported evidence of an effect of SM among participants who had suffered with NSLBP between 2 and 4 weeks at the 3-day follow-up but not at any other time point.

Dropouts were described and acceptable (<10%) in 5 studies. Four studies had high dropout rates (>10%),[33,35,44,50] 3 of which indicated that the control group had the largest dropout rate.[33,44,50]

Adverse Events

Only 3 trials reported on adverse events. Senna et al (2011)[44] reported that the most common adverse events were local discomfort and tiredness, which had resolved within 24 hours. The other 2 articles just stated that none were reported.[48,50]

The Effect of the Intervention

The effect of SM for NSLBP as measured by the 100-mm VAS is presented in the summary of findings table (Table 3). From 4 studies[35,44,46,49] (287 participants), the SMD is −0.36 (95% confidence interval, −0.59 to −0.12). The quality of evidence is graded as low because of high dropout in 2 studies[35,44] and broken blinding in 1 study,[46] no practitioner could be blinded in any study, and only 1[49] study conducted ITT analysis.

Meta-analysis

Of the 9 included studies, 5[33,35,44,46,49] reported results of the VAS sufficiently for inclusion in a meta-analysis, with 4 included in the main meta-analysis and 5 included in either of the 2 sensitivity analyses.

Each of the following studies recorded information at either 2-week[35,46] or 1-month follow-up.[44,49] These 4 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 to −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 ($I^2 < 0.1\%$, $P = .835$). However, given that there are only 4 studies included in this analysis, caution should be taken making inferences based on these analyses (Fig 2).

Three studies[35,46,49] reported information collected immediately posttreatment, and this was analyzed in a sensitivity analysis (Fig D1). The study of Waagan et al (1986)[33] was excluded from the main meta-analysis because of having a high RoB score (Table 4) but was included in a sensitivity analysis (Fig D2).

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 (Fig D3 in Appendix D).

Sensitivity Analyses

The analysis ran using the posttreatment pain scores shows 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 to −0.08) (Fig D1).

Analysis ran including Waagen et al (1986),[33] assessed to be high RoB, found results consistent with the main analysis (SMD = −0.37; 95% CI, −0.60 to −0.14) (Fig D2).

Forest and funnel plots for the sensitivity analyses can be found in Appendix D (Figs D1-D5).

Discussion

The objective of the present review was to systematically identify and critically evaluate the evidence from RCTs of SM compared with an effective sham placebo on NSLBP. This is the first review to compare SM to an effective control. The review included 9 studies, of which 4 were

Table 3. *Summary of Findings Table: SM Vs an Effective Sham for NSLBP*

	Illustrative Comparative Risks (95% CI)				
	Assumed Risk	Corresponding Risk	No of Participants (Studies)	Quality of the Evidence (GRADE)	
Outcomes	Control Group	Intervention Group			Comments
Pain (as measured by a 100-mm VAS) follow-up 2 wk to 1 mo)		The mean pain symptomology (continuous) in the intervention groups was 0.36 standard deviations lower (0.59 to 0.12 lower)	287 (4)	⊕⊕⊖⊖ [a] low	Low RoB in outcome reporting as participants were blinded effectively [b] Small to moderate SMD = −0.36 (95% CI, −0.59 to −0.12)

Patient or population: Individuals with NSLBP. Settings: clinic. Intervention: SM using high-velocity/low-amplitude thrust. Comparison: effective sham manipulation.

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.

[a] Judgements of low RoB (>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.

[b] The sham manipulation ensured blinding of participants, although one[46] tested blinding and it is possible it may have been broken.

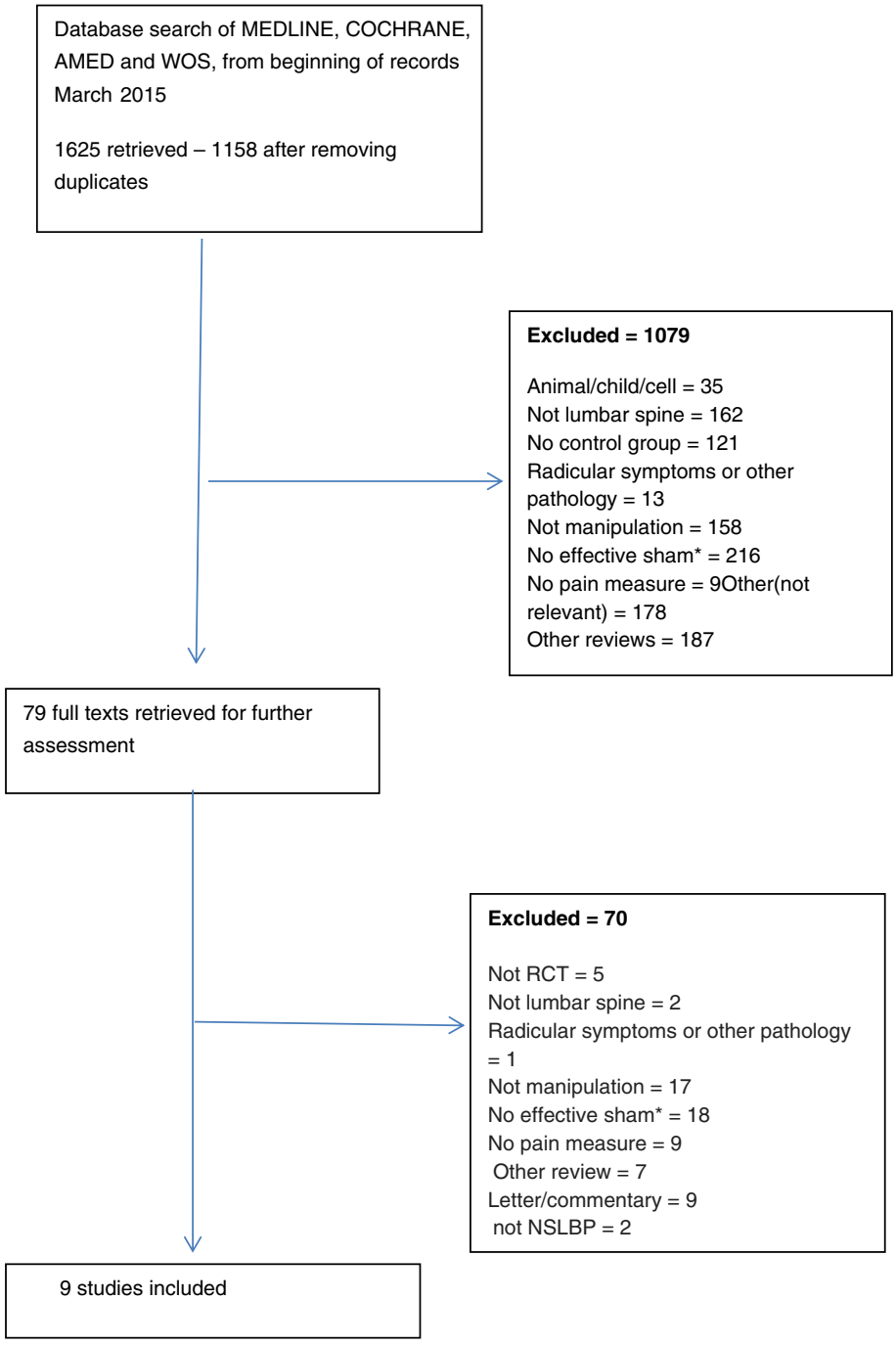

Fig 1. *Effective sham as described in the introduction. NSLBP, nonspecific low back pain; RCT, randomized controoled trial.*

included in the meta-analysis. The majority of trials used either a 100-mm VAS to assess pain levels, or the OLBPDQ or the RMLBPQ to assess physical function due to NSLBP.

The results of the meta-analysis suggest 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 posttreatment and follow-up. The pooled effect estimate of −0.36 (95% CI, −0.59 to −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,[53,54] and the CIs are wide. Caution is needed before drawing conclusions because most studies had some degree of RoB by failing to report on randomization procedure or on allocation concealment.

Table 4. *Results of VAS Pain Scores Included in Meta-analysis*

Study	Posttreatment				Follow-Up			
	Intervention		Control		Intervention		Control	
	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)

SD, standard deviation.

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[33,44,45,47–49] and did not report a priori power calculations. Four studies had high dropout rates (>10%),[33,35,44,50] 3 of which indicated that the control group had the largest dropout rate.[33,44,50] This could indicate dissatisfaction with sham as opposed to SM, which may be an indicator of some treatment effect of SM. Just 2 studies described reasons for dropouts.[46,50] One study[47] 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,[55] 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)[46] used an additional placebo medicine in both groups which may have lessened any relative difference.

Only 4 studies[45–47,50] attempted to standardize the interaction between patient and practitioner to reduce any placebo effect by way of contextual factors.[25,26,56,57] 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

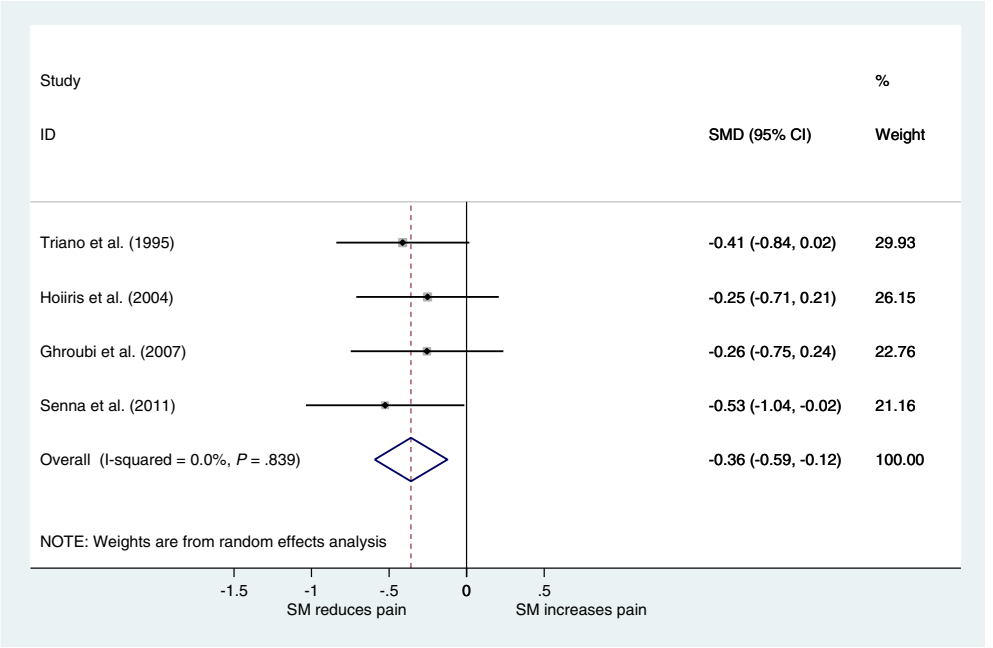

Fig 2. *Forest plot of meta-analysis looking at pain scores of participants receiving SM vs sham SM treatment. CI, confidence interval; SMD, standardized mean differences.*

between studies, making application to practice more difficult to establish. The majority of studies either had no follow-up or just 2 weeks postintervention; 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 makes this unlikely; it is more likely that we did not identify eligible unpublished trials. Publication bias is a problem in all medical research,[58] and it is a particular problem in alternative medicine.[59,60]

Furlan et al (2009)[37] 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 2 post hoc sensitivity analyses which were not planned or originally stated in the protocol but were deemed important once data were extracted. One was to include studies of high RoB,[33] and the other was to see if there was a difference at immediately postintervention compared with last follow-up (using 1-month data rather than the 10-month follow-up data for consistency[44]). There was very little variation in the findings. Functional outcomes (RMLBPDQ and OLBPDQ) due to pain levels were also extracted, as they were deemed to be a further indication of pain levels.

Adverse Events

Poor reporting of adverse events is a frequent criticism of complementary and alternative medicine research.[61] Two previous reviews on complications of SM emphasize its safety[62,63]; however, serious adverse events have been reported.[18] In this review, few studies reported on adverse events at all. In 1 study,[48] the potential for adverse events to occur was higher because they used anesthetic 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.[15,16,64–66] Despite the call of Ernst and Harkness (2001)[16] 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." 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 because they can distract from any benefits of specific treatment effects. The improvements in trial design would reduce the possibility of outcomes being due to nonspecific 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[67,68] has been shown to be reliable. An NRS measuring 0-10 has also been recommended[69] (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

Two reviews have concluded that SM is no more or less effective than other treatments with proven benefits for NSLBP.[29,70] SM may carry a greater risk of adverse events, unlike nonthrust mobilization[5] and massage.[55] Our review, however, found evidence for an effect of SM over effective control. There is currently insufficient evidence to inform practice.

Conclusions

There is some evidence from 4 of the 9 trials (287 participants) that SM has specific treatment effects and is more effective at reducing NSLBP when compared with an effective sham intervention. Although the effect was small-medium in terms of clinical relevance, a similar effect was found both immediately posttreatment and at follow-up. Inconsistency of results across all studies may be due to the use of different interventions, controls, and outcome measures and variable standards of methodology between studies. Currently, the evidence is insufficient to inform practice.

Funding Sources and Conflicts of Interest

No funding sources were reported for this study. Jay Ruddock is a registered osteopath who uses SM in the treatment of patients.

Acknowledgments

The authors thank James Barclay (Librarian) for help with obtaining publications and Lucy Mackay-Tumber and

Robin Kirk for general guidance. We thank the Bristol Nutrition BRU for providing financial contribution for Andy Ness's and Rachel Perry's time.

References

1. CBI. Healthy returns? Absence and Workplace Health Survey 2010. London: CBI; 2010 [Available from: 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;89:1029-1035.
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; 2008.
4. Fawkes C, Leach J, Mathias S, Moore AP. Standardised data collection within osteopathic practice in the UK: development and first use of a tool to profile osteopathic care in 2009. Brighton: National Council for Osteopathic Research; 2010 [Available from: http://www.osteopathy.org.uk/news-and-resources/document-library/research-and-surveys/standardised-data-collection-project/].
5. Seffinger M, Hruby R. Evidence Based Manual Medicine – A Problem Based Approach. Philadelphia: Elsevier; 2007.
6. Magee D, Zachazewski J, Quillen W. Pathology and intervention in musculoskeletal rehabilitation. St Louis: Elsevier; 2009.
7. Bronfort G, Haas M, Evans R, Leininger B, Triano J. Effectiveness of manual therapies; the UK evidence report. *Chiropr Osteopat.* 2010;18:3.
8. Urquhart DM, Hoving JL, Assendelft WJ, Roland M, van Tulder MW. Antidepressants for non-specific low back pain. *Cochrane Database Syst Rev.* 2008(1):CD001703.
9. Furlan AD, van Tulder MW, Cherkin D, et al. Acupuncture and dry-needling for low back pain. *Cochrane Database Syst Rev.* 2005(1):CD001351.
10. Bogduk N, Jull G. The theoretical pathology of acute locked back; a basis for manipulative therapy. *Man Med.* 1985;1:78-82.
11. Shekelle P. Spine update; spinal manipulation. *Spine.* 1994;6:858-861.
12. Triano J. Biomechanics of spinal manipulative therapy. *Spine J.* 2001;1:121-130.
13. Pickar J, McLain R. Responses of mechanosensitive afferents to manipulation of the lumbar facet in the cat. *Spine.* 1995;20:2379-2385.
14. Pickar J. Neurophysiological effects of spinal manipulation. *Spine J.* 2002;2:357-371.
15. Ernst E. Does spinal manipulation have specific treatment effects? *Fam Pract.* 2000;17:554-556.
16. Ernst E, Harkness E. Spinal manipulation: a systematic review of sham-controlled, double-blind, randomized controlled trials. *J Pain Symptom Manag.* 2001;22(4):879-889.
17. Evans D. Mechanisms and effects of spinal high-velocity, low-amplitude thrust manipulation: previous theories. *J Manipulative Physiol Ther.* 2002;25:251-262.
18. Herbert J, Stronski N, Frensh S, Rubenstein S. Serious adverse events and spinal manipulative therapy of the low back region: a systematic review of cases. *J Manipulative Physiol Ther.* 2013http://dx.doi.org/10.1016/j.jmpt.2013.05.009.
19. Koes B. How to evaluate manual therapy; value and pitfalls of randomised controlled trials. *Man Ther.* 2004;9:183-184.
20. Hancock M, Maher C, Latimer J, McAuley J. Selecting an appropriate placebo for a trial of spinal manipulative therapy. *Aust J Physiother.* 2006;52:135-138.
21. 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. *Aust J Physiother.* 2006;52:135-138.
22. Fulda K, Slicho T, Stoll S. Patient expectations for placebo treatments commonly used in osteopathic manipulative treatment (OMT) clinical trials; a pilot study. *Osteopath Med Prim Care.* 2007;1:3.
23. Hawk C, Long C. Use of a pilot to refine the design of a study to develop a manual placebo treatment. *J Neuromuscuolskelet Syst.* 2000;8:39-48.
24. 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. *Eur Spine J.* 2008;17:889-904.
25. Kirsch W, Weixal L. Double-blind versus deceptive administration of a placebo. *Behav Neurosci.* 1988;102:319-323.
26. Turner J, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of the placebo effects in pain treatment and research. *JAMA.* 1994;271:1609-1614.
27. Benedetti F. Placebo effects; understanding the mechanisms in health and disease. New York: Oxford University Press; 2009.
28. Licciardone J, Brimhall A, King L. Osteopathic manipulative treatment for low back pain: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskelet Disord.* 2005;6(4):43-55.
29. Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database Syst Rev.* 2011(2):CD008112.
30. Rubinstein SM, Terwee CB, Assendelft WJJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. *Cochrane Database Syst Rev.* 2012(9):CD008880.
31. Wreje U, Nordgren B, Aberg H. Treatment of pelvic joint dysfunction in primary care—a controlled study. *Scand J Prim Health Care.* 1992;10(4):310-315.
32. Licciardone J, Stoll ST, Fulda KG, et al. Osteopathic manipulative treatment for chronic low back pain: a randomised controlled trial. *Spine.* 2003;28(13):1355-1362.
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. *Man Med.* 1986;2:63-67.
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. *J Manipulative Physiol Ther.* 2009;32(5):330-343.
35. Triano J, McGregor M, Hondras M, Brennan P. Manipulative therapy versus education in chronic low back pain. *Spine.* 1995;20(8):948-953.
36. 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. *J Manipulative Physiol Ther.* 2014;22(2):59-74.
37. Furlan A, Pennick V, Bombardier C, van Tulder M, Editorial Board, Cochrane Back Review Group. Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine.* 2009;34(18):1929-1941.
38. 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.

39. Harris RM, Bradburn MJ, Deeks JJ, Harbord RM, Altman DG, Sterne J. metan: fixed- and random-effects meta-analysis. *Stata J.* 2008;8(1):3-28.
40. Sterne JA, Harbord RM. Funnel plots in meta-analysis. *Stata J.* 2004;4:127-141.
41. 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.
42. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Methods.* 1998;3(4):486-504.
43. Higgins J, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557-560.
44. 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):1427-1437.
45. 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):702-706.
46. Hoiiris K, Pfleger B, McDuffie F, et al. A randomized clinical trial comparing chiropractic adjustments to muscle relaxants for sub acute low back pain. *J Manipulative Physiol Ther.* 2004;27(6):388-398.
47. 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. *J Pain.* 2014;15(2):136-148.
48. Kawchuk G, Haugen R, Fritz J. A true blind for subjects who receive spinal manipulative therapy. *Arch Phys Med Rehabil.* 2009;90:366-368.
49. Ghroubi S, Elleuch H, Baklouti S, Elleuch M. Chronic low back pain and vertebral manipulation. *Ann Readapt Med Phys.* 2007;50(7):570-576.
50. 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):540-548.
51. Ward R C, American Osteopathic Association, editors. Foundations for osteopathic medicine. 2nd ed. Philadelphia: Lippincott, Williams and Wilkins; 2003.
52. 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. *J Manipulative Physiol Ther.* 2009;32(2):118-126.
53. Cohen J. Statistical power analysis for the behavioural sciences. 1st ed. New York, San Francisco, London: Academic Press; 1988 [1-474].
54. 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]].
55. Furlan A, Immamura M, Dryden T, Irvin E. Massage for low back pain. *Cochrane Database Syst Rev.* 2008(4):CD001929.
56. Halpern J. Empathy, using resonance emotions in the service of curiosity. *J Gen Intern Med.* 2007;22(5):696-700.
57. Thomas K. General practice consultations; is there any point being positive? *BMJ.* 1987;294:1200-1202.
58. Easterbrook P, Berlin A, Gopalan R, Matthews R. Publication bias in clinical research. *Lancet.* 1991;337(8746):867-872.
59. Ernst E, Pittler M. Alternative therapy bias. *Nature.* 1997; 385(6616):480.
60. Ernst E. Publication bias in complementary/alternative medicine. *J Clin Epidemiol.* 2007;60(11):1093-1094.
61. Ernst E. 'First, do no harm' with complementary and alternative medicine. *Trends Pharmacol Sci.* 2007;28(2): 48-50.
62. Haldeman S, Kohlbeck F, McGregor M. Risk factors and precipitating neck movements causing vertebral artery dissection after cervical trauma and spinal manipulation. *Spine.* 1999;24:785-794.
63. Powell F, Hanigan W, Olivero W. A risk/benefit analysis of spinal manipulation therapy for relief of lumbar or cervical pain. *Neurosurgery.* 1993;33:73-78.
64. Goldstein M. A challenge to the profession; initiate evidence-based osteopathic medicine now. *J Am Osteopath Assoc.* 1997; 97:448.
65. Howell J. The paradox of osteopathy. *N Engl J Med.* 1999; 341:1465-1468.
66. Ernst E, Canter P. A systematic review of systematic reviews of spinal manipulation. *J R Soc Med.* 2006;99:189-193.
67. Scrimshaw S, Maher C. Responsiveness of visual analogue scale and Mcgill pain scale measures. *J Manipulative Physiol Ther.* 2001;24(8):501-504.
68. Million R, Hall W, Nilsen KH, Baker RD, Jayson MI. Assessment of the progress of the back pain patient. *Spine.* 1982;7(3):204-212.
69. Deyo R, Dworkin S, Amtmann D, et al. Report of the Task Force on Research. Standards for chronic low-back pain. Bethesda: National Institutes of Health; 2013 [Available from: http://painconsortium.nih.gov/NIH_Pain_Programs/Task_Force/cLBP_RTF_FullReport.pdf].
70. 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.
71. 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. *Osteopath Med Prim Care.* 2009; 27(3):7.
72. Cleary C, Fox J. Menopausal symptoms; an osteopathic intervention. *Complement Ther Med.* 1994;2:181-186.
73. 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. *BMC Musculoskelet Disord.* 2006;10(7):11.
74. 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. *J Manipulative Physiol Ther.* 1994;17(6):364-368.
75. 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): 2459-2466.
76. Dishman J, Ball K, Burke J. Central motor excitability changes after spinal manioulation; a transcranial magnetic stimulation study. *J Manipulative Physiol Ther.* 2002;25(1): 1-9.
77. Hancock M, Maher CG, Latimer J, et al. 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):1638-1643.
78. 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. *Eur Spine J.* 2008;17(7):936-943.
79. 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. *J Altern Complement Med.* 2005;11(1):109.
80. Hoehler F, Tobis J, Beurger A. Spinal manipulation for low back pain. *JAMA.* 1981;245(18):1835-1838.
81. 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.
82. Kokjohn K, Schmidt D, Triano J, Brennan P. The effect of spinal manipulation on pain and prostoglandin levels in women with primary dysmenorrhea. *J Manipulative Physiol Ther.* 1992;15(5):279-285.
83. Krekoukias G, Petty N, Cheek L. Comparison of surface electromyographic activity of erector spinae before and after the application of central posteroanterior mobilization on the lumbar spine. *J Manipulative Physiol Ther.* 2009;19(1): 39-45.
84. 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):122-129.
85. Ongley M, Droman T, Klein R, Eek B, Hubert L. A new approach to low back pain. *Lancet.* 1987;2(855):143-146.
86. Mandara A, Fusaro A, Musicco M, Bado F. A randomised controlled trial on the effectiveness of osteopathic manipulative treatment of chronic low back pain. *Int J Osteopath Med.* 2008; 11:156.
87. Perry J, Green A. An investigation into the effects of a unilaterally applied lumbar mobilization technique on peripheral sympathetic nervous system activity in the lower limbs. *Man Ther.* 2008;13(6):492-499.
88. 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;41(1):13-21.
89. Roy R, Boucher J, Comtois A. Heart rate variability modulation after manipulation in pain free patients vs patients with pain. *J Manipulative Physiol Ther.* 2009;32(4):277-286.
90. 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 J.* 2006;6(2):131-137.

Appendix A. 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 B. Studies Excluded

Study	Reason Excluded
Clark et al, 2009[71]	No sham manipulation in control group.
Cleary and Fox, 1994[72]	Fox's "low force osteopathic technique" does not match the inclusion criteria for an HVLA maneuver.
Cleland et al, 2006[73]	No sham control; comparison groups are alternative manipulation or mobilization.
Cote et al, 1994[74]	No sham control; comparison group is a mobilization.
Cramer et al, 2002[75]	No sham control; comparison group is side lying positioning.
Dishman et al, 2002[76]	No sham control; comparison group is side lying positioning.
Hancock et al, 2007[77]	No sham control; placebo is detuned ultrasound therapy.
Hancock et al, 2008[78]	No sham control; placebo is detuned ultrasound therapy.
Hawk et al, 2005[79]	Excluded because of article's own assessment of inadequate blinding.
Hoehler et al, 1981[80]	No sham intervention; massage was used as the control.
Hondras et al, 1999[81]	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[82]	Participants not generalizable.
Krekoukias et al, 2009[83]	No sham control; comparison groups are prone lying and prone lying with touch onto L3 spinal level.
Learman et al, 2009[52]	Although the subjects were extensively screened for pain levels at entry to the study, no follow-up data measuring pain were assessed.
Licciardone et al, 2003[32]	Several different and nonstandardized interventions (muscle energy techniques, soft tissue manipulation, fascial manipulation and craniosacral) were made in the treatment group; the sham group received "fake" treatments in the same modalities.
Licciardone et al, 2013[84]	Several different and nonstandardized interventions (muscle energy techniques, soft tissue manipulation, fascial manipulation, and craniosacral) were made in the treatment group; the sham group received "fake" treatments in the same modalities.
Ongley et al, 1987[85]	As well as an SM, 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[86]	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[87]	No measurement of pain as an outcome.
Puetendura et al, 2010[88]	No measurement of pain as an outcome.
Roy et al, 2009[89]	Although groups were divided into pain and pain free, no measurement of pain was taken.
Santilli et al, 2006[90]	Radicular symptoms present.

HVLA, high velocity/low amplitude; *SM*, spinal manipulation.

Appendix C. 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 dropout rate described and acceptable?	Yes/no/unsure
	7	Were all randomized participants analyzed 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 RoB for RCTs was assessed using the criteria list recommended in the Updated Method Guidelines for Systematic Reviews in the CBRG. Scores of 6 or more were considered low RoB.

Appendix D. Forest and Funnel Plots From Sensitivity Analyses

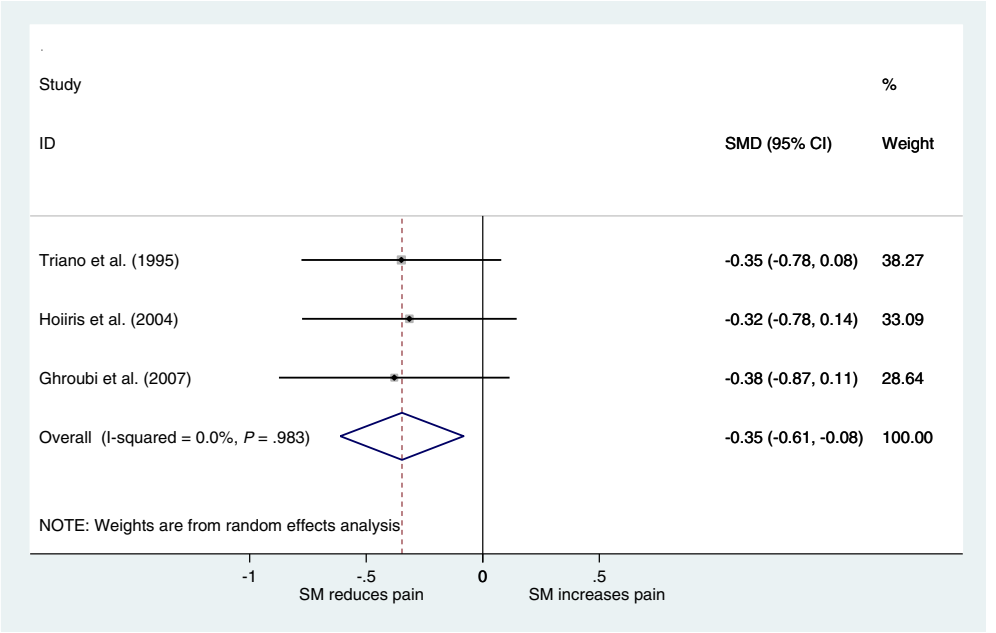

Fig D1. *Forest plot of meta-analysis looking at pain scores of participants receiving SM vs sham SM treatment when assessing pain scores immediately posttreatment. CI, confidence interval; SM, spinal manipulation; SMD, standardized mean differences.*

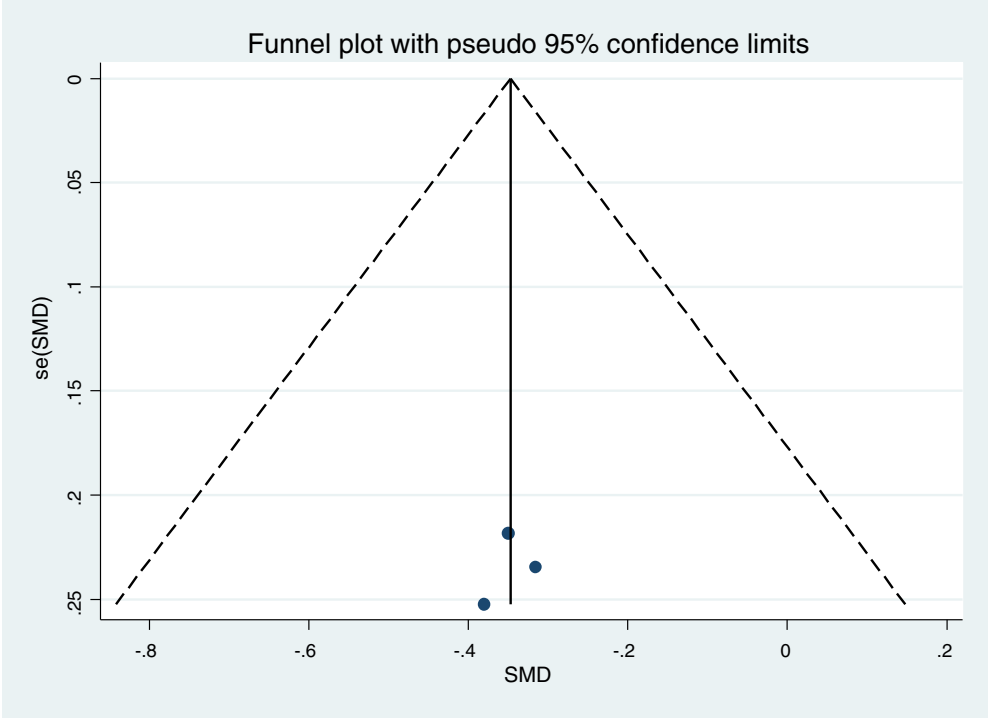

Fig D2. *Funnel plot of studies comparing SM vs sham SM treatment when assessing pain scores immediately posttreatment. SMD, standardized mean differences.*

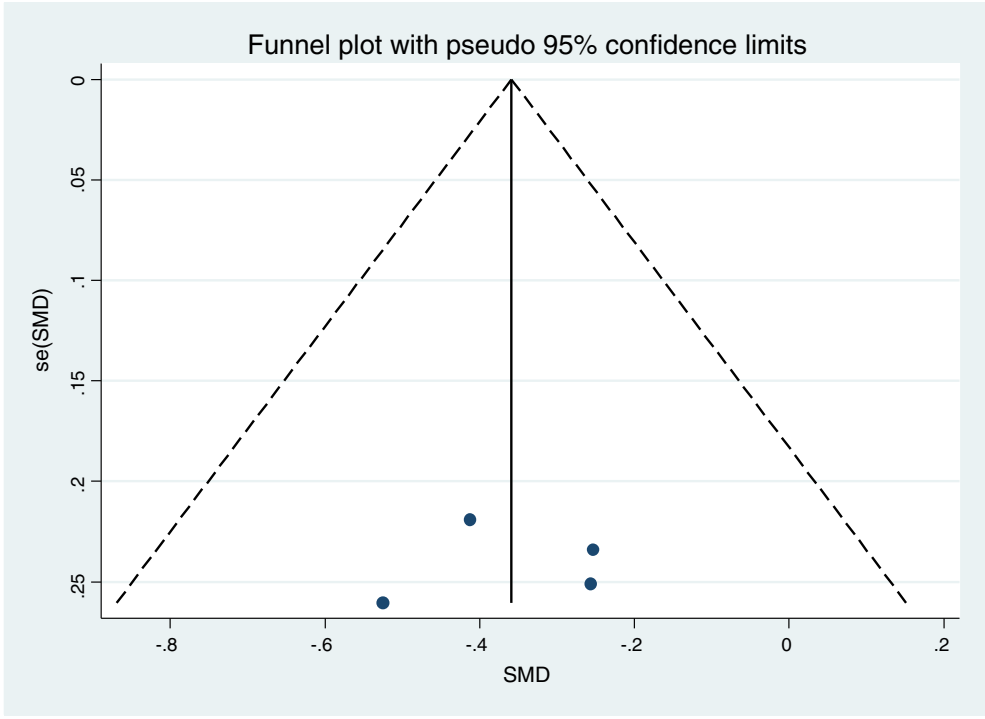

Fig D3. *Funnel plot of studies comparing SM vs sham treatment. SMD, standardized mean differences.*

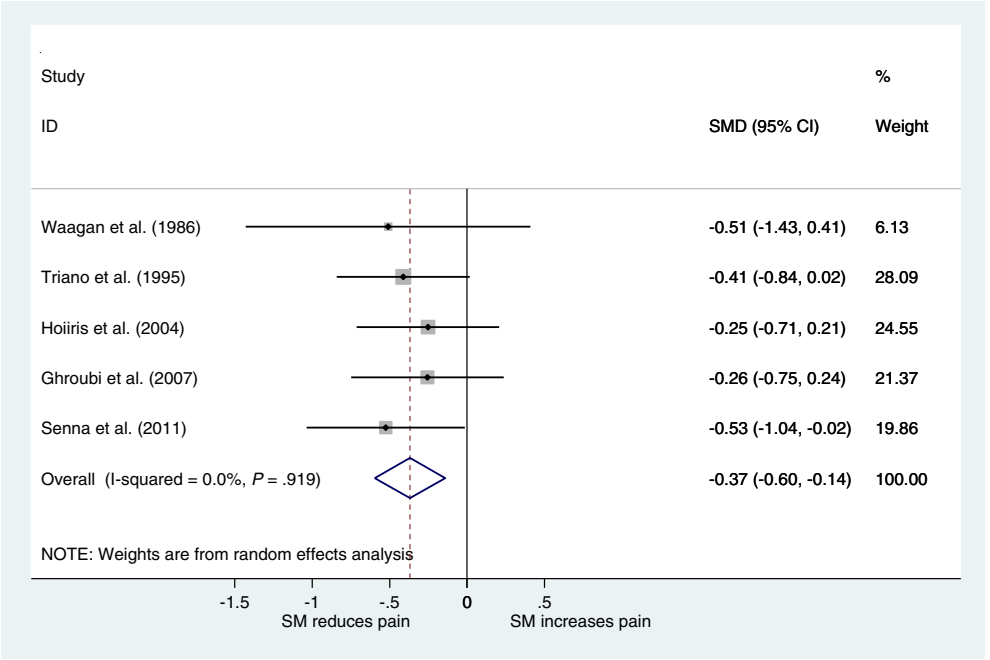

Fig D4. *Forest plot of meta-analysis comparing SM vs sham SM treatment as assessed at follow-up when including Waagen et al (1986). CI, confidence interval; SMD, standardized mean differences.*

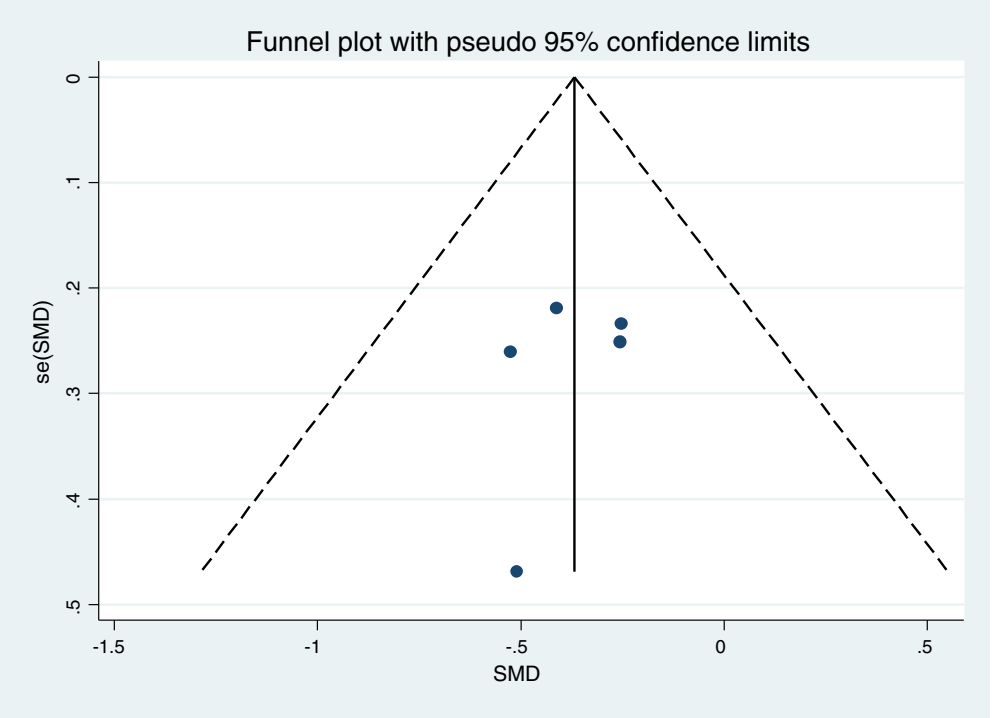

Fig D5. *Funnel plot of studies comparing SM vs sham SM treatment as assessed at follow-up when including Waagen et al (1986). SMD, standardized mean differences.*