Spinal Manipulative Therapy for Chronic Low-Back Pain An Update of a Cochrane Review

Rubinstein, S.M.; Middelkoop, M. van; Assendelft, W.J.J.; Boer, M.R. de; Tulder, M.W. van

Spine

SPINE Volume 36, Number 13, pp E825–E846

COCHRANE COLLABORATION

Spinal Manipulative Therapy for Chronic Low-Back Pain

An Update of a Cochrane Review

Sidney M. Rubinstein, DC, PhD,* Marienke van Middelkoop, PhD,† Willem J.J. Assendelft, MD, PhD,‡ Michiel R. de Boer, PhD,§ and Maurits W. van Tulder, PhD ¶

Study Design. Systematic review of interventions.

Objective. To assess the effects of spinal manipulative therapy (SMT) for chronic low-back pain.

Summary of Background Data. SMT is one of the many therapies for the treatment of low-back pain, which is a worldwide, extensively practiced intervention.

Methods. *Search methods*. An experienced librarian searched for randomized controlled trials (RCTs) in multiple databases up to June 2009. *Selection criteria*. RCTs that examined manipulation or mobilization in adults with chronic low-back pain were included. The primary outcomes were pain, functional status, and perceived recovery. Secondary outcomes were return-to-work and quality of life. *Data collection and analysis*. Two authors independently conducted the study selection, risk of bias assessment, and data extraction. GRADE was used to assess the quality of the evidence.

Results. We included 26 RCTs (total participants = 6070), 9 of which had a low risk of bias. Approximately two-thirds of the included studies (N = 18) were not evaluated in the previous review. There is a high-quality evidence that SMT has a small, significant, but not clinically relevant, short-term effect on pain relief (mean difference −4.16, 95% confidence interval −6.97 to −1.36) and functional status (standardized mean difference −0.22, 95% confidence interval −0.36 to −0.07) in comparison with other interventions. There is varying quality of evidence that SMT has a significant short-term effect on pain relief and functional status when added to another intervention. There is a very low-quality evidence that SMT is not more effective than inert interventions or sham SMT for short-term pain relief or functional status. Data were particularly sparse for recovery, return-to-work, quality of life, and costs of care. No serious complications were observed with SMT.

Conclusions. High-quality evidence suggests that there is no clinically relevant difference between SMT and other interventions for reducing pain and improving function in patients with chronic low-back pain. Determining cost-effectiveness of care has high priority.

Key words: low back pain, spinal manipulation, systematic review, meta-analysis, randomized controlled trial, chiropractic. **Spine 2011;36:E825–E846**

Low-back pain is a common and disabling disorder in western society, which represents a great financial burden in the form of direct costs resulting from loss of work and medical expenses, as well as indirect costs.[1] Therefore, adequate treatment of low-back pain is an important issue for patients, treating clinicians, and healthcare policy makers. Spinal manipulative therapy (SMT) is widely used for acute and chronic low-back pain, which has been examined in many randomized controlled trials (RCTs). These trials have been summarized in numerous recent systematic reviews,[2–5] which have formed the basis for recommendations in clinical guidelines.[6–11] Most notably, these guidelines are largely dependent on an earlier version of this Cochrane review.[12] That review concluded that SMT was moderately

From the *Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands; †Department of General Practice, Erasmus Medical Center, Rotterdam, the Netherlands; ‡Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands; §Institute of Health Sciences, Faculty of Earth and Life Sciences, VU University Medical Center, Amsterdam, the Netherlands; and ¶ Department of Health Sciences, Faculty of Earth and Life Sciences, VU University, Amsterdam, the Netherlands.

Acknowledgement date: February 18, 2011. Acceptance date: March 2, 2011.

This work was supported by the Dutch Health Insurance Council (CVZ).

This article is based on Rubinstein SM, van Middelkoop M, Assendelft WJJ, 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. DOI: 10.1002/14651858.CD008112.pub2, a Cochrane Review published in *The Cochrane Library* 2011, Issue 2 (see www.thecochranelibrary.com for information).

It is co-published with the permission of John Wiley & Sons Ltd, on behalf of the Cochrane Collaboration. Cochrane reviews are regularly updated as new evidence emerges and in response to feedback, and *The Cochrane Library* should be consulted for the most recent version of the review."

Contributions of authors: Conception and design—S.M. Rubinstein, M.W. van Tulder, W.J.J. Assendelft; Analysis and interpretation of the data—S.M. Rubinstein, M.R. de Boer, M.W. van Tulder; Drafting of the review—S.M. Rubinstein, M.W. van Tulder; critical revision of the article for important intellectual content: All members. Final approval of the article: All members. Statistical Expertise: M.R. de Boer. Administrative, technical, or logistical support: S.M. Rubinstein, M.R. de Boer. Collection and assembly of data: S.M. Rubinstein, M. van Middelkoop, M.R. de Boer, W.J.J. Assendelft (studies published before 2000).

No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Sidney M. Rubinstein, DC, PhD, Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research VU University Medical Center, PO Box 7057, Room D518, 1007 MB Amsterdam, Netherlands; E-mail: sm.rubinstein@vumc.nl

DOI: 10.1097/BRS.0b013e3182197fe1

superior to sham manipulation and therapies thought to be ineffective or harmful for acute or chronic low-back pain; however, the effect sizes were small and arguably not clinically relevant. Furthermore, SMT was found to be no more effective than other standard therapies (*e.g.*, general practitioner care, analgesics, exercise, or back schools) for short- or long-term pain relief or functional improvement for acute or chronic low-back pain.

Recommendations regarding SMT vary across national guidelines on the management of back pain.[13,14] For example, SMT is considered to be a therapeutic option in the acute phase of low-back pain in many countries, while in other countries, such as the Netherlands, Australia, and Israel, it is not recommended.[13] Similarly, SMT is considered to be a useful option in the subacute or chronic phase in the Danish and Dutch guidelines, but is either not recommended or absent in the other national guidelines.

The purpose of this review is to update the previous Cochrane review, using the most recent guidelines developed by the Cochrane Collaboration in general[15] and by the Cochrane Back Review Group in particular.[16] In contrast to the previous Cochrane review, this review has been split into two parts by duration of the complaint, namely acute[17] and chronic low-back pain. This review reports on chronic low-back pain only, on the basis of the published protocol.[18]

DESCRIPTION OF THE CONDITION

Low-back pain is defined as pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without referred leg pain. Chronic low-back pain is typically defined as pain persisting for more than 12 weeks.[19] *Nonspecific low-back pain* is further defined as low-back pain not attributed to a recognizable, known specific pathology (*e.g.*, infection, tumor, fracture, or radicular syndrome).

DESCRIPTION OF THE INTERVENTION

SMT is considered here as any "hands-on" treatment, including both manipulation and mobilization of the spine.[12] Mobilizations use low-grade velocity, small or large amplitude passive movement techniques within the patient's range of motion and control. Manipulation, on the contrary, uses a high-velocity impulse or thrust applied to a synovial joint over a short amplitude at or near the end of the passive or physiologic range of motion, which is often accompanied by an audible "crack."[20] The cracking sound is caused by cavitation of the joint, which is a term used to describe the formation and activity of bubbles within the fluid.[21,22] Various practitioners, including chiropractors, manual therapists (physiotherapists trained in manipulative techniques), orthomanual therapists (medical doctors trained in manipulative techniques), or osteopaths use this intervention in their practices. However, the diagnostic techniques and philosophy of the various professions differ. The focus of orthomanual medicine is on abnormal positions of the skeleton and symmetry in the spine, whereas manual therapy focuses on functional disorders of the musculoskeletal system, and chiropractic focuses on the musculoskeletal and nervous systems in relation to the general health of the patient.[23]

HOW THE INTERVENTION MIGHT WORK

Many hypotheses exist regarding the mechanism of action for spinal manipulation and mobilization,[24–26] and some have postulated that given their theoretically different mechanisms of action, mobilization and manipulation should be assessed as separate entities.[21] The modes of action might be roughly divided into mechanical and neurophysiologic. The mechanistic approach suggests that SMT acts on a manipulable lesion (often called the functional spinal lesion or subluxation) which proposes that forces to reduce internal mechanical stresses will result in reduced symptoms.[27] However, given the non-nociceptive behavior of chronic low-back pain, a purely mechanistic theory alone cannot explain clinical improvement.[21] Much of the literature focuses on the influence on the neurological system, where it is suggested that spinal manipulation therapy impacts the primary afferent neurons from paraspinal tissues, the motor control system, and pain processing,[26] although the actual mechanism remains debatable.[21,25]

WHY IT IS IMPORTANT TO DO THIS REVIEW

SMT is a worldwide, extensively practiced intervention provided by a variety of professions. However, the efficacy of this therapy for chronic low-back pain is not without dispute. This review, with its comprehensive and rigorous methodology, is thought to provide better insight into this problem. Although numerous systematic reviews have examined the efficacy of SMT for low-back pain,[6,7] very few have conducted a meta-analysis, especially for chronic low-back pain. Also, many of the reviews were narrative rather than systematic and the results were not consistent.[28] The previous version of the Cochrane review was published in 2004 and since then many new trials, including some with large numbers of participants, have been published. In addition, the methodology of systematic reviews has recently been updated,[15] as well as the specific guidelines for reviews of back and neck pain.[16]

OBJECTIVES

The objective of this review was to examine the effectiveness of SMT on pain, functional status, and recovery at the short-, intermediate-, and long-term follow-up measurements in comparison to control treatments (*e.g.*, no treatment, sham, and all other treatments) for adults with chronic low-back pain.

MATERIALS AND METHODS

Criteria for Considering Studies for this Review

Types of studies

Only randomized studies were included. Studies using an inadequate randomization procedure (*e.g.*, alternate allocation, allocation based on birth date) were excluded.

Types of Participants

Inclusion criteria

- Adult participants (18 years of age or older) with low-back pain with a mean duration for the current episode

(for the study population) longer than 12 weeks, meaning more than 50% of the study population had pain that had lasted longer than three months.
- Studies with patients from primary, secondary, or tertiary care.
- Patients with or without radiating pain.

Exclusion Criteria

Subjects with:
- Postpartum low-back pain or pelvic pain due to pregnancy.
- Pain not related to the low-back, *e.g.*, coccydynia.
- Postoperative studies or subjects with "failed-back syndrome" or studies that
  - Examined "maintenance care" or prevention.
  - Were designed to test the immediate postintervention effect of a single treatment only, with no additional follow-up (because we were interested in the effect of SMT beyond one day).
  - Exclusively examined specific pathologies, *e.g.*, sciatica. Note: Studies of sciatica were excluded because it has been identified by many as a prognostic factor associated with a poor outcome,[29,30] especially with SMT.[31,32] Sciatica was defined here as radiating pain following the sciatic distribution and exhibiting signs of a radiculopathy.

Types of Interventions

Experimental intervention

The experimental intervention examined in this review includes both spinal manipulation and mobilization for chronic low-back pain. Unless otherwise indicated, SMT refers to both "hands-on" treatments.

Types of Comparison

Studies were included for consideration if the study design used suggested that the observed differences were due to the unique contribution of SMT. This excludes studies with a multimodal treatment as one of the interventions (*e.g.*, standard physician care + spinal manipulation + exercise therapy) and a different type of intervention or only one intervention from the multimodal therapy as the comparison (*e.g.*, standard physician care alone), thus rendering it impossible to decipher the effect of SMT. However, studies comparing SMT in addition to another intervention to that same intervention alone were included.

Comparison therapies were combined into the following main clusters:

1) SMT *versus* inert interventions
2) SMT *versus* sham SMT
3) SMT *versus* all other interventions
4) SMT in addition to any intervention *versus* that intervention alone

Inert interventions included, for example, detuned diathermy and detuned ultrasound. "All other interventions" included both presumed effective and ineffective interventions for treatment of chronic low-back pain. Determination of what interventions were considered ineffective and effective was based on the literature and our interpretation of those results.[6,7]

Types of Outcome Measures

Only patient-reported outcome measures were evaluated. Physiological measures, such as spinal flexibility or degrees achieved with a straight leg raise test (*i.e.*, Lasègue sign), were not considered clinically relevant outcomes and were not included.

Primary Outcomes

- Pain expressed on a self-reported scale (*e.g.*, visual analog scale [VAS], numerical rating scale [NRS]).
- Functional status expressed on a back-pain specific scale (*e.g.*, Roland-Morris Disability Questionnaire, Oswestry Disability Index).
- Global improvement or perceived recovery (recovered is defined as the number of patients reported to be recovered or nearly recovered).

Secondary Outcomes

- Health-related quality of life (HRQoL) (*e.g.*, SF-36 [as measured by the general health subscale], EuroQol, general health [*e.g.*, as measured on a VAS scale] or similarly validated index).
- Return-to-work.

Search Methods for Identification of Studies

Electronic searches

We identified RCTs and systematic reviews by electronically searching the following databases:
- CENTRAL (The Cochrane Library 2009, issue 2)
- MEDLINE from January 2000 to June 2009)
- EMBASE from January 2000 to June 2009)
- CINAHL from January 2000 to June 2009)
- PEDro up to June 2009
- Index to Chiropractic Literature up to June 2009

The search strategy developed by the Cochrane Back Review Group was followed, using free text words and MeSH headings.[16] A search was not conducted for studies published before 2000 because they were included in the previous Cochrane review.[12] The entire search strategy is available on request from the primary author.

Searching Other Resources

In addition to the aforementioned, we also (1) screened the reference lists of all included studies and systematic reviews pertinent to this topic and (2) searched the main electronic sources of ongoing trials (National Research Register, *meta*-Register of Controlled Trials; Clinical Trials).

DATA COLLECTION AND ANALYSIS

Selection of Studies

Two review authors with a background in chiropractic (S.M.R.) and movement science (M.vM.) independently screened

the titles and abstracts from the search results. Potentially relevant studies were obtained in full text and independently assessed for inclusion. Disagreements were resolved through discussion. A third review author (M.W.vT.) was contacted if an arbiter was necessary. Only full articles were evaluated. Abstracts and proceedings from congresses or any other "gray literature" were excluded. There were no language restrictions.

Data Extraction and Management

A standardized form was used to extract data from the included articles. The following data were extracted: study design (RCT), study characteristics (*e.g.*, country where the study was conducted, recruitment modality, source of funding, risk of bias [RoB]), patient characteristics (*e.g.*, number of participants, age, gender), description of the experimental and control interventions, cointerventions, duration of follow-up, types of outcomes assessed, and the authors' results and conclusions. Data were extracted independently by the same two review authors who conducted the selection of studies. Any disagreements were discussed and an arbiter (M.W.vT.) was consulted when necessary. Key findings were summarized in a narrative format. Data relating to the primary outcomes were assessed for inclusion in the meta-analyses and final value scores (means and standard deviations) were extracted. Change scores were converted to a mean value for the respective follow-up measurement. Outcomes were assessed at 1, 3, 6, and 12 months and data included according to the time closest to these intervals. Only one study examined data beyond 12 months.[33]

Assessment of RoB in Included Studies

The RoB assessment for RCTs was conducted using the 12 criteria recommended by the Cochrane Back Review Group and evaluated independently by same two review authors mentioned earlier (S.M.R., M.vM.). These criteria are standard for evaluating effectiveness of interventions for low-back pain[16] and the operational definitions are available on request. The criteria were scored as "yes," "no," or "unclear" and reported in the *Risk of Bias* table. Any disagreements between the review authors were resolved by discussion, including input from a third independent review author (M.W.vT.). In virtually all cases, an attempt was made to contact authors for clarification of methodological issues if the information was unclear. A study with a low RoB was defined as one fulfilling six or more of the criteria items, which is supported by empirical evidence,[34] and with no fatal flaw, which is defined as those studies with (1) a dropout rate greater than 50% at the first and subsequent follow-up measurements or (2) statistically and clinically relevant important baseline differences for one or more primary outcomes (*i.e.*, pain, functional status) indicating unsuccessful randomization. Quantitative data from studies with a fatal flaw were excluded from the meta-analyses (see RoB in the included studies). Because the review authors were already familiar with the literature, they were not blinded to authors of the individual studies, institution, or journal.

Blinding the patient and practitioner to treatment allocation is nearly impossible in trials of SMT. Given that the primary outcomes assessed in this review are all subjective measures (*i.e.*, pain, functional status, perceived recovery), any attempt to blind the outcome assessor was considered irrelevant because the patient is viewed to be the outcome assessor when evaluating subjective measures. Therefore, if the patient is not blinded, the outcome assessor was also considered not blinded. However, to drop these items from the assessment is to negate the observation that "blinding" of research personnel and participants provides less biased data.

Measures of Treatment Effect

Treatment effect was examined through meta-analyses, but these were conducted only if studies were thought to be clinically homogenous. Clinical homogeneity was defined *a priori* by setting, population, and comparison group. A mean difference (MD) was calculated for pain and when necessary, VAS or NRS scales were converted to a 100-point scale. Other scales were allowed if it was thought that the construct measured was consistent with the outcome being evaluated. For functional status, a standardized mean difference (SMD) was calculated because many different instruments were used (*e.g.*, Roland-Morris Disability Questionnaire, Oswestry Disability Index, disability subscale of the von Korff scale, Disability Rating Index). A negative effect size indicates that SMT is more beneficial than the comparison therapy, meaning that subjects have less pain and better functional status. Quality of life was analyzed by an SMD. Where necessary, scores were transformed, so that a higher score indicates a better outcome, which is how this was typically measured; therefore, a negative effect size indicates that the contrast therapy is more beneficial. For dichotomous outcomes (*i.e.*, recovery, return-to-work), a risk ratio (RR) was calculated and the event defined as the number of subjects recovered or returned-to-work. A positive RR indicates that SMT results in a greater chance of recovery or return-to-work. A random-effects model was used for all analyses because a substantial amount of heterogeneity remained unexplained by the subgroup and sensitivity analyses. Funnel plots were only examined for publication bias for the comparison, SMT *versus* all other interventions, due to the fact that the other comparisons included too few studies. For each treatment comparison, an effect size and a 95% confidence interval (CI) were calculated. All analyses were conducted in Review Manager 5.0.

Assessment of clinical relevance. The determination of clinical relevance was evaluated by one question: "Is the size of the effect clinically relevant?" Levels of clinical relevance were defined as (1) small: MD <10% of the scale (*e.g.*, <10 mm on a 100-mm VAS); SMD or "*d*" scores ≤0.2; relative risk 0.8–1.25; (2) medium: MD 10% to 20% of the scale, SMD or "*d*" scores =0.5; relative risk 1.25–2.0 or 0.5–0.8; and (3) large: MD >20% of the scale, SMD or "*d*" scores ≥0.8, relative risk 0.5–2.0.[15,35]

Unit of Analysis Issues

We attempted to combine data in studies with multiple comparisons where it was thought that similar contrasts were used and the outcomes were thought to be clinically similar. This was conducted for one study,[36] which included two similar forms of exercise as the contrast to SMT, general exercise and motor control exercise. In all other cases, when multiple contrasts were examined in the same comparison (*e.g.*, SMT *vs.* physiotherapy *vs.* standard medical care), the number of subjects in the shared comparison, SMT, was halved. This step corrects for error introduced by "double-counting" of subjects for the "shared comparison" in the meta-analyses. Another study presented data from a crossover trial,[37] in which case, data were presented prior to the crossover of the intervention.

Dealing with Missing Data

In cases where data were reported as a median and interquartile range (IQR), it was assumed that the median was equivalent to the mean and the width of the IQR equivalent to 1.35 times the standard deviation (section 7.7.3.5).[15] In one study,[38] a range was presented along with the median instead of a IQR, in which case, the standard deviation was estimated to be one-quarter of the range, although we recognize that this method is not robust and potentially subject to error (section 7.7.3.6).[15] In another study, data were presented together for neck and low-back pain.[39] A subsequent stratified analysis had been performed for the low-back pain data but was no longer available. However, we were able to extract the results from a recent systematic review,[24] which presented these data as between-group differences. Where data were reported in a graph and not in a table, the means and standard deviations were estimated. When standard deviations were not reported, an attempt was made to contact the author. In the absence of additional information, these were calculated from the confidence intervals, where possible. If the standard deviation for follow-up measurements was missing, its baseline measure was used for the subsequent follow-ups. Finally, if no measure of variation was reported anywhere in the text, the standard deviation was estimated on the basis of other studies with a similar population and RoB.

Assessment of Heterogeneity

Heterogeneity was explored in two manners, informally by vision (eye-ball test) and formally tested by the Q-test (chi-square) and I^2; however, the decision regarding heterogeneity was dependent on I^2.[15] Substantial heterogeneity is defined as $\geq$50%, and where necessary, the effect of the interventions is described if the results are too heterogeneous.

Data Synthesis

The overall quality of the evidence and strength of recommendations were evaluated using GRADE.[40] The quality of the evidence for a specific outcome was based on performance against five principal domains: (1) limitations in design (downgraded when more than 25% of the participants were from studies with a high RoB), (2) inconsistency of results (downgraded in the presence of significant statistical heterogeneity [$I^2 >$50%] and inconsistent findings [in the presence of widely differing estimates of the treatment effect, *i.e.*, individual studies favoring either the intervention or control group]), (3) indirectness (*i.e.*, generalizability of the findings; downgraded when more than50% of the participants were outside the target group, for example, studies that exclusively examined older subjects or included inexperienced treating physicians), (4) imprecision (downgraded when the total number of participants was less than 400 for each continuous outcome and 300 for dichotomous outcomes), and (5) other (*e.g.*, publication bias). Single studies ($N < 400$ for continuous outcomes, $N < 300$ for dichotomous outcomes) were considered inconsistent and imprecise and provide "low-quality evidence," which could be further downgraded to "very low-quality evidence" if there were also limitations in design or indirectness. Summary of Findings tables were generated for the primary analyses and for the primary outcome measures only, regardless of statistical heterogeneity, but when present, this was noted. The quality of the evidence is described as follows:

High quality: Further research is very unlikely to change our confidence in the estimate of effect. There are sufficient data with narrow confidence intervals. There are no known or suspected reporting biases.

Moderate quality: Further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate; one of the domains is not met.

Low quality: Further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change the estimate; two of the domains are not met.

Very low quality: Great uncertainty about the estimate; three of the domains are not met.

No evidence: No evidence from RCTs.

Subgroup Analysis and Investigation of Heterogeneity

Regardless of possible heterogeneity of the included studies, the following stratified analyses were conducted: (1) By control groups as defined in *Types of intervention* (see *Types of comparisons*); and (2) by time, that is, short-term (closest to 1 to 3 months), intermediate (closest to 6 months) and long-term follow-up (closest to 12 months).

Sensitivity Analysis

The following sensitivity analyses were planned *a priori* and conducted to explain possible sources of heterogeneity between studies: (1) for RoB; (2) for studies with an adequate allocation procedure; (3) by duration of the low-back pain (studies that included subacute and chronic *vs.* studies of exclusively chronic low-back pain); (4) by type of technique (high-velocity low-amplitude manipulation); (5) by type of manipulator (chiropractor *vs.* manual therapist or physiotherapist); and (6) by type of comparison therapy (presumed ineffective therapies [*e.g.*, diathermy, ultrasound, single counseling session with advice on back pain] and presumed effective therapies [*e.g.*, exercise, standard medical care, physiotherapy]). In addition, a specific type of contrast (*i.e.*, exercise therapy) was examined *posteriori* because it was thought to be an important contrast, but not earlier defined in the protocol. Summary forest plots were constructed in STATA v.10, which depict these results.

RESULTS

Description of Studies

Characteristics of the included, excluded, and ongoing studies are available upon request.

Results of the Search

Since the publication of the previous review, 18 new trials were identified, which fulfilled the inclusion criteria[33,36,41–56]; thus, this review represents a majority of studies published in the past decade. Eight trials from the previous review are included,[37–39,57–60] one of which recently published long-term results (Figure 1).[61] Multiple publications were identified for many studies and the most prominent publication was used for citation purposes.

The countries in which the studies were conducted varied but were largely limited to North America and Europe. Eight studies were conducted in the United States,[43–47,57,58,60] seven studies in the United Kingdom,[33,37,38,41,48,54,55] five in Finland,[50,51,53,56,61] two in Australia,[36,49] one in Denmark,[52] one in Italy,[59] one in the Netherlands,[39] and one in Tunesia.[42] All trials were published in English except the trial conducted in Tunesia, which was published in French.

Included Studies

In total, 6070 patients were examined in the trials. Study sample sizes ranged from 29 to 1334 (median [IQR] = 149 [86–244]).

Types of studies. In total, four studies were identified, which compared SMT with a placebo in the form of an anti-oedema gel spread over the lumbar region[59] or other inert interventions (*i.e.*, detuned short-wave diathermy[38]; detuned ultrasound[39]; corset and transcutaneous muscle stimulation[58]); three studies that compared SMT with sham SMT[42,47,60]; 21 studies that compared SMT with any other intervention—both presumed effective or ineffective (*i.e.*, acupuncture,[49] back school,[45,59] educational back booklet with or without additional counselling,[33,50] exercise therapy,[33,36,41,43,50,51,54,57,61] myofascial therapy,[45] massage,[58] pain clinic,[55] pharmaceutical/analgesic therapy only,[49,59] short-wave diathermy,[38] standard medical care, consisting of, among other things, analgesic therapy and advice/reassurance,[39,44,46,53] standard physiotherapy,[39,46,56,59,61] and ultrasound[48]); and five studies that compared SMT plus another intervention with the intervention alone (*i.e.*, analgesic therapy,[37] exercise,[52] myofascial therapy,[45] standard medical care and in combination with exercise,[54] and usual care[47]).

Study population. The included studies represent a rather heterogeneous population with regard to duration of pain, presence or absence of radiating pain, and distribution of age (available upon request). Most studies included middle-aged subjects with or without radiating pain. One study included subjects older than 55 years,[44] and two studies included subjects without radiating pain.[42,49] However, in a number of studies it was not clear whether subjects with radiating pain were included or not.[33,38,48,53,60] Relatively few studies examined exclusively chronic low-back pain (*i.e.*, an inclusion criteria that specified that the symptoms must have been present for 3 months or longer) [33,36,41,43,47–49,52,55]; however, most studies indicated that patients had a current episode of low-back pain consisting of months to years.

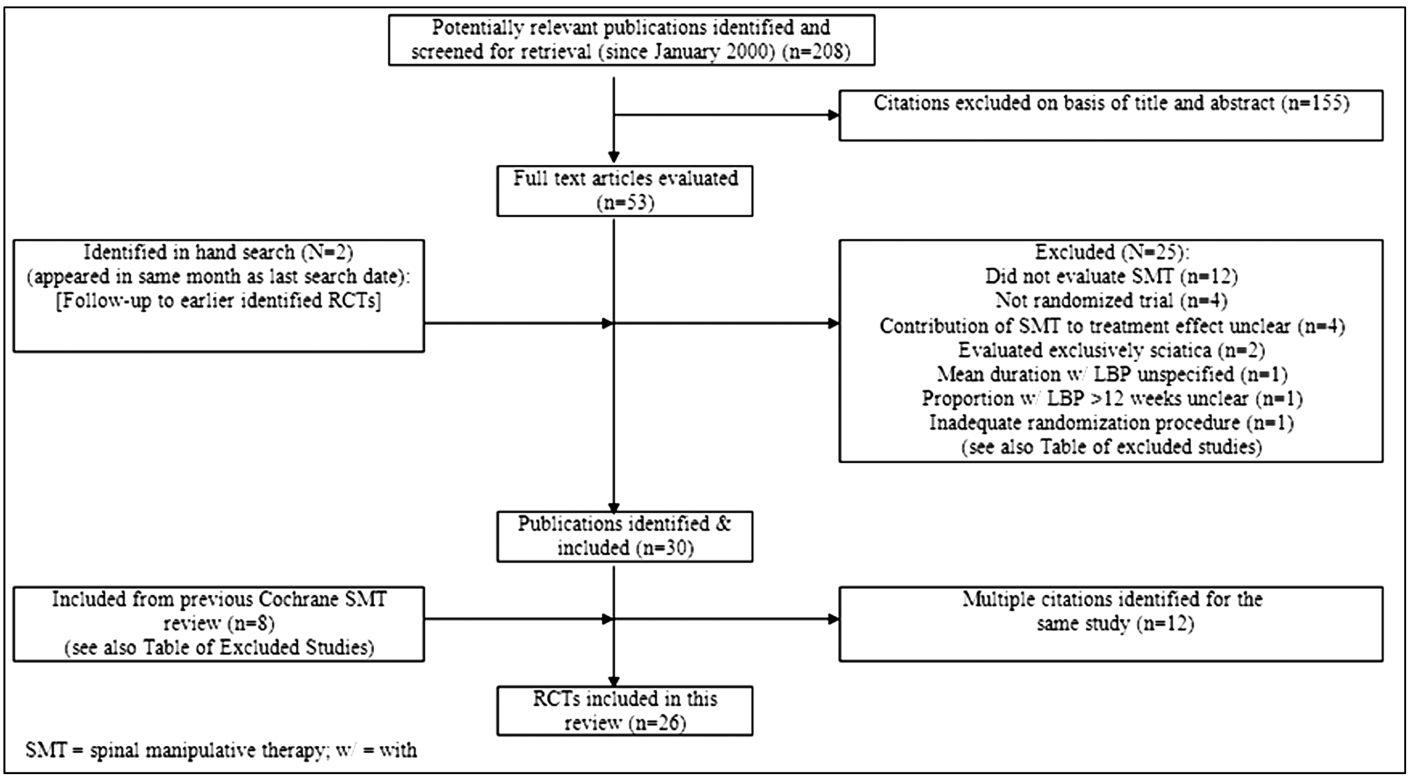

Figure 1. Study flow diagram: Summary of the selection process.

Technique: type, practitioner, number, and duration of treatment. The type of technique, type of treating physician/therapist, and number and duration of the treatments also varied. In 10 studies, treatment was delivered by a chiropractor,[43–46,49,55,57–60] in 5, by a manual or physical therapist,[33,36,39,48,51] in 3, by an osteopath,[38,41,47] in 3, by a medical manipulator or orthomanual therapist,[37,50,52] in 2, by a bone-setter,[56,61] in 1, by a naprapath,[53] and in 1, by a number of different disciplines.[54] In another study, it was unclear what type of SMT treatment was delivered and what the level or skill of the treating physicians was.[42] In virtually all studies, treatment was delivered by a few select experienced physicians/therapists, with the exception of the UK BEAM study,[54] where participants were treated in the manipulative arm of the study in 45 clinics by as many as 84 practitioners of various professions. In another study, treatment was delivered by a few select predoctoral osteopathic manipulative medicine fellows, who could be considered inexperienced in manipulative treatments.[47]

The primary type of (thrust) technique used in the SMT arm of the studies varied highly and was defined as a high-velocity low-amplitude thrust,[41,44–47,49,50,52,54,57,58,60] Maitland mobilization,[36,48] mobilization consisting of flexion-distraction,[43,44] unspecified mobilization,[51,61] unspecified rotational thrust technique,[37,38] and unspecified technique,[33,39,42,53,56,59] or allowed various types of thrust and/or nonthrust techniques to be used within the study.[55]

It is unclear how many treatments the participants received on average because studies did not typically report this. The maximum number of treatments allowed by protocol was, on average, 8 (SD = 4; data from 24 studies). In other studies, this was at the discretion of the therapist/physician and terminated sooner if the patient recovered. Similarly, the treatment period was also quite varied. The duration of the treatment was protocolized for, on average, 7 weeks (SD = 4; data from 23 studies).

Outcome measures: types, timing. All but one study reported on pain.[41] All studies measured this construct *via* a VAS or NRS, with the exception of two,[53,54] which used the pain subscale from the modified von Korff scale. Most studies reported back-pain-specific functional status, consisting of either the Roland-Morris disability questionnaire[36,43–47,50,54,55,57] or Oswestry Disability Index[33,41,48,49,51,56,61]; however, other scales such as the modified von Korff scale[53] (disability data presented separately), Disability Rating Index,[51] and a 4-point nonvalidated scale[59] were also used. Slightly more than one-third of the studies reported on some aspect of perceived recovery[36–39,43,44,46,53,56,57]; however, these data were not always able to be extracted because it was expressed, for example, as a continuous variable[36,39,44] or was not presented separately for the low back.[53] Relatively few studies reported on the secondary outcomes, such as return-to-work or aspects thereof, such as number of sick-leave days,[38,45,47,57,61] costs associated with care,[43,54,61] or HRQoL such as *via* the SF-36,[43–45,47,49,54] EuroQoL,[41,54] HRQoL–15D questionnaire,[56] Nottingham Health Profile,[33] general health status (expressed on a 10-cm VAS),[51] and other (Dartmouth Primary Care Cooperative Information Project chart system [*i.e.*, COOP][57]). In addition, when the SF-36 was measured, data were not always available for the general health subscale, as some studies either reported an overall score[44,45,47] or presented other subscales.[54] One study examined a mixed population (neck and low-back); data are presented for the low-back only.[39]

Timing of the outcome measures ranged from 2 weeks to 2 years postrandomization. The majority reported short- and intermediate-term outcomes, although many reported long-term outcomes as well.

Safety. Slightly more than one-third of the studies reported on adverse events.[37,43–45,49,52–54,57] Adverse events in the SMT group were limited to muscle soreness, stiffness, and/or transient increase in pain. None of the studies registered any serious complications in either the experimental or control group.

Excluded Studies

Many studies were excluded because the proportion of subjects with chronic low-back pain was either unclear or unspecified[62–76]; the mean duration of symptoms for the population was less than 12 weeks (*i.e.*, 50% of the population with less than 12 weeks of low-back pain)[77–81]; the contribution of SMT to the treatment effect could not be discerned[82–85]; the procedure of randomization and allocation was clearly inappropriate[86–90]; the study evaluated exclusively subjects with specific pathology, such as sciatica,[30,64,91] the study included postsurgical patients,[92] or the study did not evaluate SMT as defined here.[93]

Risk of Bias in Included Studies

The results of the RoB for the individual studies are summarized in Figure 2. In total, 9 of the 26 trials met the criteria for a low RoB.[36,39,44–46,53,54,57,61] In total, three studies, all with a high RoB, were identified with a fatal flaw and excluded from the meta-analyses: Two studies had more than 50% dropout at the first follow-up measurement[41,49] and one study was found to have clinically relevant baseline differences between the interventions for one or more primary outcomes suggesting that randomization was not properly conducted.[33]

The followin——g professions were represented in those studies with a low RoB: bone-setters,[61] chiropractors,[44–46,57] manual/physical therapists,[36,39] naprapaths,[53] and combination of various professionals (*i.e.*, chiropractors, physiotherapists, and osteopaths).[54]

Allocation

Slightly less than half of the studies used both an adequate sequence generation and allocation procedure.[36,39,43,44,46,53–57,61] In seven studies, both randomization and allocation were unclear.[37,38,48,52,59,60]

Blinding

In total, three studies attempted to blind patients to the assigned intervention by providing a sham treatment.[42,47,60] Of these, only one evaluated the success of blinding post-treatment.[60] In that study, 52% (n = 15/29) of the participants completed a post-treatment evaluation of the success of the blinding: 17% (n = 1/6) from the experimental group thought they had received sham SMT, while 67% (n = 6/9) from the

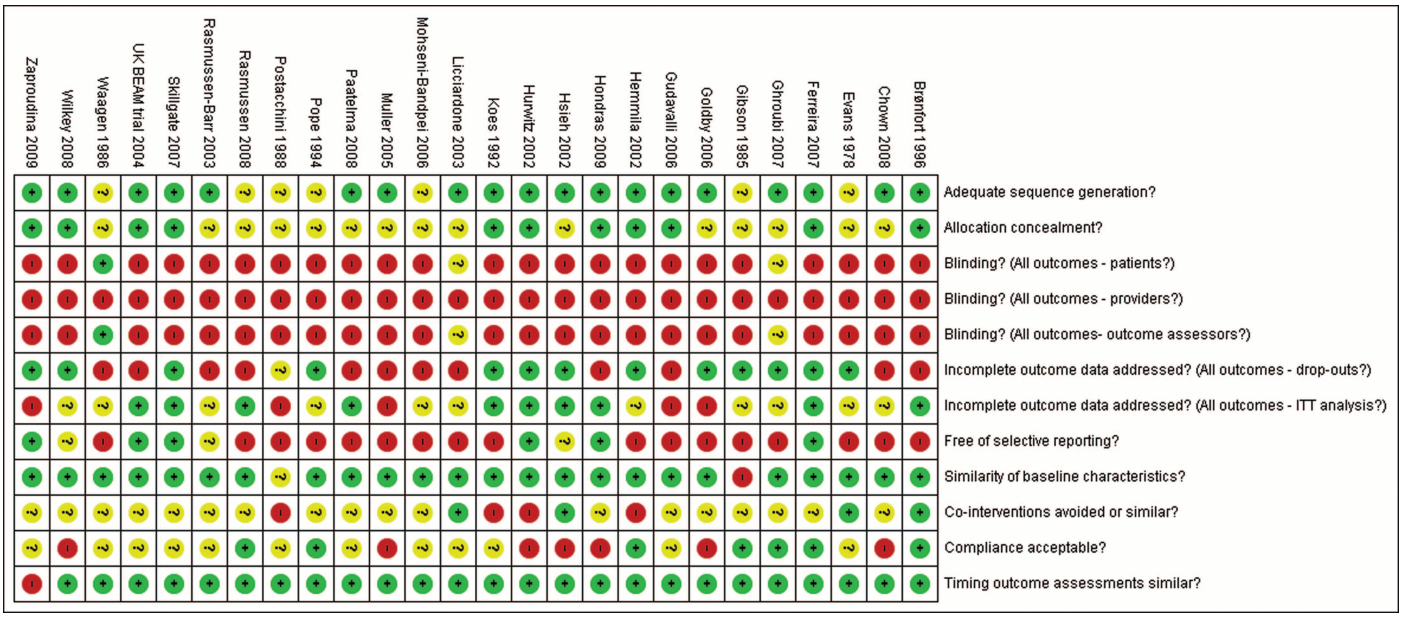

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

sham group thought that they had received SMT, suggesting perhaps that blinding was partially successful.

Incomplete Outcome Data

Half of the studies provided an adequate overview of withdrawals or dropouts and were able to keep these to a minimum for the subsequent follow-up measurements, although not all of these conducted long-term follow-up.[33,36–39,42,45,46,53,55,56,58,61] In another study, there was a difference in the dropout rate between groups.[33]

Selective Reporting

Published or registered protocols were available for relatively few studies,[36,44,53,54,56] despite an extensive and comprehensive search, which included searching for registered clinical trials in www.clinicaltrials.gov, ISRCTN, and other trial registries. In the absence of these, it was difficult for us to determine whether outcomes were measured, but not reported because they were found to be insignificant or unfavorable. Therefore, studies reporting all three primary outcomes (*i.e.*, pain, back-pain specific functional status, and perceived recovery) were considered to have fulfilled this criterion. Only one study was identified with no published protocol or registered in one of the main trial registries but reported all three primary outcomes.[46]

Other Potential Sources of Bias

Publication bias. An examination of publication bias was possible for only one comparison, SMT *versus* any other intervention, because of the paucity of data for the other comparisons. Funnel plots were constructed for the outcomes, pain, and functional status and are available upon request. For the outcome pain, it might appear that small studies favoring SMT are missing. This may indicate publication bias because some studies may have used SMT as a control group in a trial evaluating the effects of another intervention.

Effects of Interventions

Primary analyses

Summary effect estimates are presented when there was no substantial heterogeneity. Findings are summarized in Tables 1–4.

Effect of SMT *Versus* Inert Interventions

In total, four studies were identified,[38,39,58,59] one of which had a low RoB.[39] Based on one study[38] (72 participants), there is very low-quality evidence (high RoB, inconsistency, imprecision) that there is no significant difference between SMT and inert interventions (*i.e.*, detuned short-wave diathermy and detuned ultrasound) for pain relief at 1 and 3 months (MD −6.00, 95% CI −15.82 to 3.82; and MD 7.00, 95% CI −3.58 to 17.58, respectively). For recovery, one study[38] (72 participants) with a high RoB was identified. There is very low-quality evidence (high RoB, inconsistency, imprecision) that there is no significant difference between SMT and inert interventions at 1 and 3 months (RR 1.03, 95% CI 0.49–2.19; and RR 0.96, 95% CI 0.56–1.65, respectively). For return to work, one study[38] with a high RoB was identified. There is also very low-quality evidence (high RoB, inconsistency, imprecision) that there is no significant difference at 1 or 3 months (RR 1.29, 95% CI 1.00–1.65; and RR 1.17, 95% CI 0.97–1.40, respectively). No data were available for functional status or HRQoL.

Three studies were identified for which data for the meta-analyses could not be extracted.[39,58,59] One study (N = 76) demonstrated a significant difference in improvement ($P < 0.05$) between SMT and detuned physiotherapy modalities at 6 weeks, but not 3 months.[39] Another study (N = 127)

TABLE 1. Spinal Manipulative Therapy in Comparison to Inert Interventions for Chronic Low-Back Pain

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	Inert Interventions	SMT				
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 1 mo	The mean pain in the control groups was 27 points	The mean pain in the intervention groups was 6.00 lower (15.82 lower to 3.82 higher)		72 (1 study)	⊕⊖⊖⊖ very low†,‡,§	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 3 mos	The mean pain in the control groups was 6 points	The mean Pain in the intervention groups was 7.00 higher (3.58 lower to 17.58 higher)		70 (1 study)	⊕⊖⊖⊖very low†,‡,§	
Recovery at 1 mo	Study population		RR 1.03 (0.49–2.19)	72 (1 study)	⊕⊖⊖⊖very low*,†,¶	
	273 per 1000	281 per 1000 (134–598)				
	Medium risk population					
Recovery at 3 mo	Study population		RR 0.96 (0.56–1.65)	70 (1 study)	⊕⊖⊖⊖very low*,†,¶	
	438 per 1000	420 per 1000 (245–723)				
	Medium risk population					

Patient or population: *patients with chronic low-back pain;* ***Settings:*** *rather diverse;* ***Intervention:*** *spinal manipulative therapy;* ***Comparison:*** *inert interventions.*

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*

**The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

†High risk of bias.

‡Less than 400 subjects, total.

§Effect includes the possibility of better or worse pain relief with SMT.

¶Effect includes the possibility of better or worse chance of recovery with SMT.

CI indicates confidence interval; RR, risk ratio; SMT, spinal manipulative therapy; VAS, visual analog scale.

TABLE 2. Spinal Manipulative Therapy (SMT) in Comparison with Sham SMT for Chronic Low-Back Pain

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	Sham SMT	SMT				
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 1 mo	The mean pain ranged across control groups from 31 to 58 points	The mean pain in the intervention groups was 3.24 lower (13.62 lower to 7.15 higher)		148 (3 studies)	⊕⊖⊖⊖ very low[†,‡,§,¶,#]	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 3 mo	The mean pain in the control groups was 28.5 points	The mean pain in the intervention groups was 2.50 higher (9.64 lower to 14.64 higher)		55 (1 study)	⊕⊖⊖⊖ very low[†,§,¶,#]	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 6 mo	The mean pain in the control groups was 24.5 points	The mean pain in the intervention groups was 7.10 higher (5.16 lower to 19.36 higher)		51 (1 study)	⊕⊖⊖⊖ very low[†,§,¶,#]	
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 1 mo	The mean functional status in the control groups was 7.7	The mean functional status in the intervention groups was 2.16 lower (4.65 lower to 0.29 higher)		65 (1 study)	⊕⊖⊖⊖ very low[†,§,¶,**]	On the basis of SMD −0.45 (−0.97 to 0.06), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 3 mo	The mean functional status in the control groups was 6.1	The mean functional status in the intervention groups was 0.00 higher (2.3 lower to 2.3 higher)		55 (1 study)	⊕⊖⊖⊖ very low[†,§,¶,**]	On the basis of SMD 0.00 (−0.56 to 0.56), there is no effect.
Functional status Roland-Morris disability questionnaire. Scale from: 0 to 24 (worse function). Follow-up: 6 mo	The mean functional status in the control groups was 5	The mean functional status in the intervention groups was 0.18 higher (2.34 lower to 2.75 higher)		51 (1 study)	⊕⊖⊖⊖ very low[†,§,¶,**]	On the basis of SMD 0.04 (−0.52 to 0.61), strength of the effect is small.

Patient or population: *patients with chronic low-back pain;* ***Settings:*** *Rather diverse;* ***Intervention:*** *SMT;* ***Comparison:*** *sham SMT.*

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.*

**The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

[†]>25% of participants from studies with a high risk of bias

[‡]I^2 = 53%

[§]Licciardone et al included relatively inexperienced osteopathic manipulative physicians.[47]

[¶]Less than 400 subjects, total.

[#] Effect includes the possibility of better or worse pain relief with SMT.

[**]Effect includes the possibility of better or worse function with SMT.

CI indicates confidence interval; VAS, visual analog scale.

TABLE 3. Spinal Manipulative Therapy (SMT) in Comparison to All Other Interventions for Chronic Low-Back Pain

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	All Other Interventions	SMT				
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 1 mo	The mean pain ranged across control groups from 21.3 to 44 points	The mean pain in the intervention groups was 2.76 lower (5.19–0.32 lower)		1405 (6 studies†)	⊕⊖⊖⊖ high	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 3 mo	The mean pain ranged across control groups from 27.5 to 44.7 points	The mean pain in the intervention groups was 4.55 lower (8.68–0.43 lower)		1074 (5 studies†)	⊕⊖⊖⊖ moderate‡	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 6 mo	The mean pain ranged across control groups from 22 to 45.6 points	The mean pain in the intervention groups was 3.07 lower (5.42–0.71 lower)		1105 (4 studies†)	⊕⊖⊖⊖ high	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 12 mo	The mean pain ranged across control groups from 28 to 50.6 points	The mean pain in the intervention groups was 0.76 lower (3.19 lower to 1.66 higher)		1285 (3 studies†)	⊕⊖⊖⊖ high§	
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 1 mo	The mean functional status ranged across control groups from 4 to 20.8	The mean functional status in the intervention groups was 0.9 lower (1.6–0.3 lower)		1402 (6 studies†)	⊕⊖⊖⊖ high	On the basis of SMD −0.17 (−0.29 to −0.06), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 3 mo	The mean functional status ranged across control groups from 6 to 20.9	The mean functional status in the intervention groups was 0.74 lower (1.5 lower to 0.04 higher)		1323 (6 studies†)	⊕⊖⊖⊖ moderate¶	On the basis of SMD −0.18 (−0.37 to 0.01), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 6 mo	The mean functional status ranged across control groups from 3.5 to 9.3	The mean functional status in the intervention groups was 0.58 lower (1.1 lower to 0 higher)		1313 (5 studies†)	⊕⊖⊖⊖ high	On the basis of SMD − 0.12 (−0.23 to 0.00), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 12 mo.	The mean functional status ranged across control groups from 5.7 to 9.2	The mean functional status in the intervention groups was 0.32 lower (0.86 lower to 0.27 higher)		1418 (4 studies†)	⊕⊖⊖⊖ high#	On the basis of SMD −0.06 (−0.16 to 0.05), strength of the effect is small.

(*Continues*)

TABLE 3. (*Continued*)

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	All Other Interventions	SMT				
Recovery at 1 mo	Study population		RR 1.20 (1.04–1.37)	370 (3 studies)	⊕⊖⊖⊖ moderate[**]	
	598 per 1000	718 per 1000 (622–819)				
	Medium risk population					
Recovery at 3 mo	Study population		RR 1.7 (1.2–2.4)	182 (2 studies)	⊕⊖⊖⊖ low[††,‡‡]	
	333 per 1000	566 per 1000 (400–799)				
	Medium risk population					

Patient or population: *patients with chronic low-back pain;* ***Settings:*** *rather diverse;* ***Intervention:*** *spinal manipulative therapy;* ***Comparison:*** *all other interventions.*

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.*

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

[†]Results based on studies with a low risk of bias.

[‡]I^2 = 61%

[§]Effect includes the possibility of better or worse pain relief with SMT.

[¶]I^2 = 52% and widely varying effect estimates in favor of either SMT or the intervention.

[#]Effect includes the possibility of better or worse function with SMT.

[**]>25% of participants from studies with a high risk of bias

[††]High risk of bias

[‡‡]Less than 300 subjects, total.

CI indicates confidence interval; RR, risk ratio; VAS, visual analog scale.

TABLE 4. Spinal Manipulative Therapy (SMT) Plus Any Intervention in Comparison with the Intervention Alone for Chronic Low-Back Pain

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	The Intervention Alone	SMT Plus any Intervention				
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 1 mo	The mean pain ranged across control groups from 27.8 to 46.5 points	The mean pain in the intervention groups was 5.88 lower (10.85–0.9 lower)		228 (3 studies)	⊕⊖⊖⊖ low†,‡	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 3 mo	The mean pain ranged across control groups from 45.2 to 49.6 points	The mean pain in the intervention groups was 7.23 lower (11.72–2.74 lower)		1016 (2 studies)	⊕⊕⊕⊕ high	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 6 mo	The mean pain ranged across control groups from 29.9 to 36.5 points	The mean pain in the intervention groups was 6.77 lower (14.07 lower to 0.53 higher)		143 (2 studies)	⊕⊖⊖⊖ low†,‡,§	
Pain VAS. Scale from 0 to 100 (worse pain). Follow-up: 12 mo	The mean pain ranged across control groups from 20 to 47.6 points	The mean pain in the intervention groups was 3.31 lower (6.6–0.02 lower)		1000 (2 studies)	⊕⊕⊕⊕ high	
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 1 mo	The mean functional status ranged across control groups from 5.8 to 6.9	The mean functional status in the intervention groups was 2.05 lower (3.73–0.36 lower)		156 (2 studies)	⊕⊖⊖⊖ low†,‡	On the basis of SMD −0.40 (−0.73 to −0.07), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 3 mo	The mean functional status ranged across control groups from 5.5 to 6.7	The mean functional status in the intervention groups was 1.06 lower (1.82–0.29 lower)		1078 (2 studies)	⊕⊕⊕⊕ high	On the basis of SMD −0.22 (−0.38 to −0.06), strength of the effect is small.
Functional status Roland-Morris disability questionnaire. Scale from 0 to 24 (worse function). Follow-up: 6 mo	The mean functional status ranged across control groups from 5.1 to 6.2	The mean functional status in the intervention groups was 1.44 lower (3.07 lower to 0.14 higher)		142 (2 studies)	⊕⊖⊖⊖ low†,‡,¶	On the basis of SMD −0.30 (−0.64 to 0.03), strength of the effect is small.

(*Continues*)

TABLE 4. (*Continued*)

Outcomes	Illustrative Comparative Risks* (95% CI)		Relative Effect (95% CI)	No. of Participants (Studies)	Quality of the Evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
	The Intervention Alone	SMT Plus any Intervention				
Functional status Roland-Morris disability questionnaire. Scale from: 0 to 24 (worse function). Follow-up: 12 mo	The mean functional status ranged across control groups from 5.7 to 6.2	The mean functional status in the intervention groups was 0.97 lower (1.56–0.41 lower)		994 (1 study)	⊕⊖⊖⊖ high	On the basis of SMD −0.21 (−0.34 to −0.09), strength of the effect is small.
Recovery at one month	Study population		RR 3.40 (1.12–10.28)	32 (1 study)	⊕⊖⊖⊖ very low[†,#]	
	176 per 1000	598 per 1000 (197–1000)				
	Medium risk population					

Patient or population: *patients with chronic low-back pain;* ***Settings:*** *rather diverse;* ***Intervention:*** *SMT plus any intervention;* ***Comparison:*** *the intervention alone.*

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.*

**The basis for the assumed risk (e.g., the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

[†]>25% of participants from studies with a high risk of bias.

[‡]Fewer than 400 subjects, total.

[§]Effect includes the possibility of better or worse pain with SMT.

[¶]Effect includes the possibility of better or worse function with SMT.

[#]Fewer than 300 subjects, total.

CI indicates confidence interval; RR, risk ratio.

demonstrated no statistically significant difference in pain ($P < 0.05$) between SMT and use of a corset or transcutaneous muscle stimulation.[58] Because of poor reporting, it is unclear from the study (N = 95) whether there was a statistically significant difference in improvement between SMT and a placebo group (*i.e.*, anti-oedema gel) at 3 weeks or 6 months.[59]

Effect of SMT *Versus* Sham SMT

In total, three studies were identified, all with a high RoB.[42,47,60] There was substantial heterogeneity for pain at 1 month; thus, the results are described here. Two studies demonstrated a nonsignificant effect in favor of SMT,[42,60] while another study demonstrated a nonsignificant effect in favor of sham SMT.[47] All examined different forms of SMT, that is, unspecified SMT, osteopathic SMT, and chiropractic SMT, respectively, and all were relatively small studies. For pain relief, based on one study (55 participants), there is very low-quality evidence (high RoB, inconsistency, indirectness, imprecision) that there is no significant difference between SMT and sham SMT at 3 and 6 months (MD 2.50, 95% CI −9.64 to 14.64; and MD 7.10, 95% CI −5.16 to 19.36, respectively).[47] For functional status, based on the aforementioned study, there is also very low-quality evidence (high RoB, inconsistency, indirectness, imprecision) that there is no significant difference at 1, 3, or 6 months (SMD −0.45, 95% CI −0.97 to 0.06; SMD 0.00, 95% CI −0.56 to 0.56; SMD 0.04, 95% CI −0.52 to 0.61).[47] No data were available from any study on recovery, return to work, or HRQoL.

Effect of SMT *Versus* all Other Interventions

In total, 15 studies were examined in the meta-analyses, 8 with a low RoB.[36,38,43–46,48,50,51,53–57,61] Data from three studies were not included because these data could not be extracted,[39,58,59] and data from the one study with a low RoB are described later, where relevant.[39]

For pain and to a lesser extent, functional status, there was substantial heterogeneity for the short-term and intermediate follow-ups; therefore, results are reported separately for these outcomes for only studies with a low RoB. This step was taken because heterogeneity across studies was much less when accounting for RoB and far more studies were available for this comparison than any of the other comparisons. Furthermore, there was, at most, a 2-point difference in pain (100-point scale, range: 0.13–2.01) and at most a 0.13-point difference for functional status (SMD, range 0–0.13) for any of the particular time measurements between studies with a low RoB only and all studies; therefore, we feel confident in presenting these stratified results here. In general, the effect was not systematically greater when including all studies than including studies with a low RoB only. In total, eight studies with a low RoB were examined.[36,44–46,53,54,57,61]

For pain, there is high-quality evidence that SMT provides statistically significantly better pain relief than other interventions at 1 and 6 months (MD −2.76, 95% CI −5.19 to −0.32; and MD −3.07, 95% CI −5.42 to −0.71, respectively) (Figure 3); however, there is also high-quality evidence from three studies (1285 participants) that SMT is not statistically more effective for pain relief at 12 months (MD −0.76, 95% CI −3.19 to 1.66).[36,46,54] At 3 months, despite substantial heterogeneity from five studies (1047 participants), SMT provides significantly better pain relief than the control interventions (MD −4.55, 95% CI −8.68 to −0.43; $I^2 = 61\%$).[36,53,54,57,61] It is noteworthy that only one of the effect estimates (N = 56) favors the control group in this particular comparison.[61]

For functional status, there is high-quality evidence that SMT provides statistically significantly better functional improvement at 1 month than other interventions (SMD −0.17, 95% CI −0.29 to −0.06). There is moderate-quality evidence (inconsistency) of no statistically significant effect at 3 months (SMD −0.18, 95% CI −0.37 to 0.01) and high-quality evidence of no statistically significant effect at 6 and 12 months (SMD −0.12, 95% CI −0.23 to 0.00; and SMD −0.06, 95% CI −0.16 to 0.05, respectively) (Figure 4).

Four studies examined perceived recovery,[38,43,56,61] one with a low RoB.[61] There is moderate-quality evidence (high RoB) from three studies at 1 month[38,43,61] (370 participants) and low-quality evidence (high RoB, imprecision) from two studies[38,56] (182 participants) at 3 months that SMT provides a significantly better chance of recovery than the contrast interventions (RR 1.20, 95% CI 1.04–1.37; and RR 1.70, 95% CI 1.20–2.40, respectively). There is also low-quality evidence (inconsistency, imprecision) from one study demonstrating no statistically significant difference in effect on recovery at 6 or 12 months (RR 1.05, 95% CI 0.81–1.38; and RR 1.17, 95% CI 0.87–1.55, respectively).[61] One study reported significantly ($P < 0.05$) greater improvement for SMT *versus* standard medical care, but not physiotherapy at 6 weeks, and no significant difference between either at 3 months.[39]

Four studies (596 participants),[38,43,57,61] two of which had a low RoB,[57,61] examined return to work. There is low-quality evidence (high RoB, imprecision) that there is no statistically significant effect of SMT on return to work at any short- or long-term interval. Four studies examined HRQoL[43,51,56,57] (478 participants), one of which had a low RoB. Based on these three studies, there is moderate-quality evidence (high RoB) at 1 month demonstrating no statistically significant difference in effect on HRQoL (RR −0.08, 95% CI −0.29 to 0.13) and very low-quality evidence (high RoB, inconsistency, imprecision) of no significant difference in effect at 3 months (RR 0.21, 95% CI −0.27 to 0.70).

Effect of SMT Plus Another Intervention *Versus* the Intervention Alone

In total, five studies were identified,[37,45,47,52,54] two of which had a low RoB.[45,54] There is low-quality evidence (high RoB, imprecision) from three studies[45,47,52] (228 participants) that SMT has a statistically significant effect on pain relief at 1 month (MD −5.88, 95% CI −10.85 to −0.90) and high-quality evidence from two studies[47,54] (1016 participants) that SMT has a statistically significant effect on pain relief at 3 months (MD −7.23, 95% CI −11.72 to −2.74). There is also high-quality evidence from two studies[52,54] (1000 participants) that SMT has a statistically significant effect on pain relief at 12 months (MD −3.31, 95% CI −6.60 to −0.02). However,

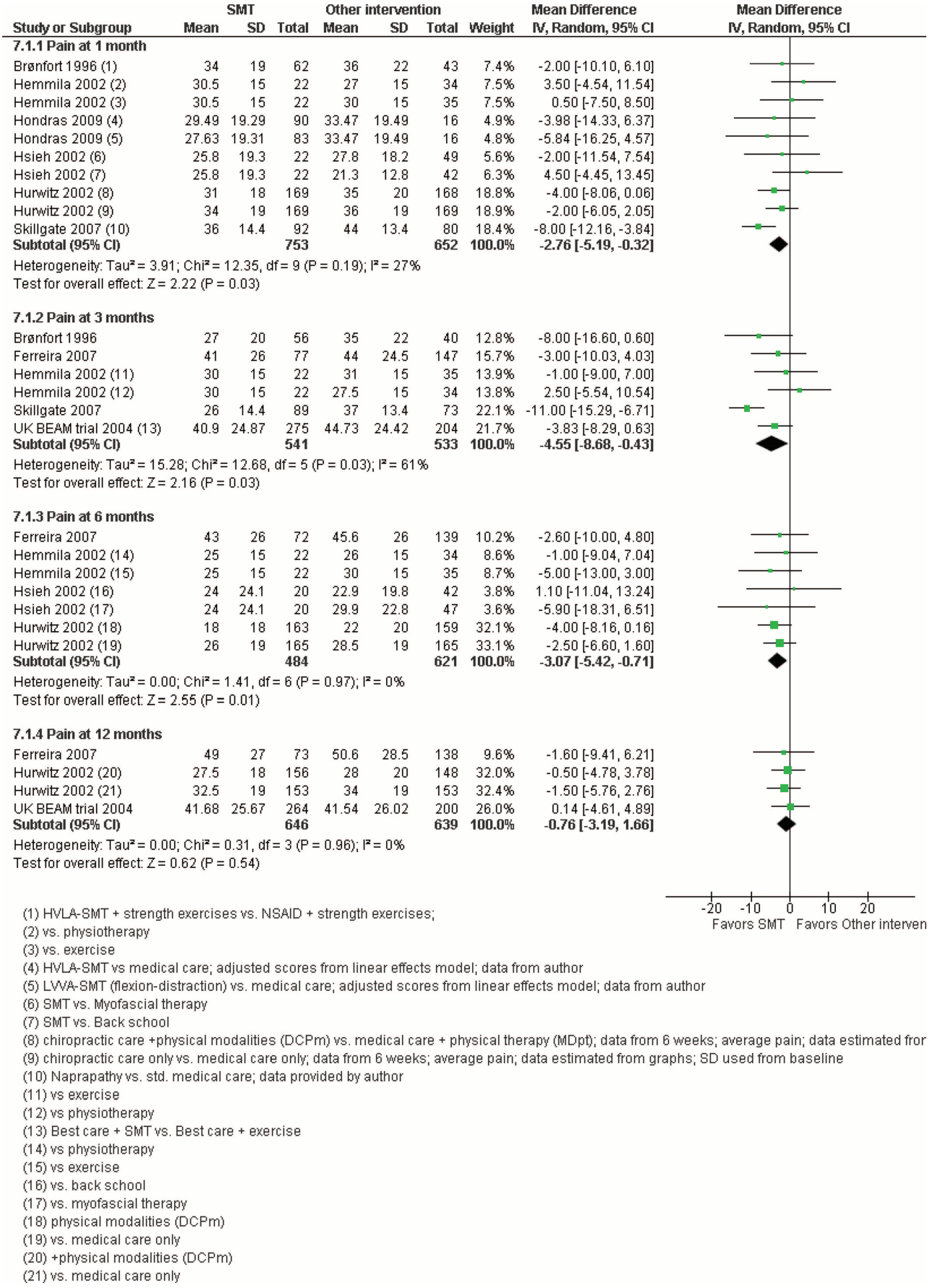

Figure 3. Forest plot of comparison: 3 Spinal manipulative therapy vs. all other therapies, outcome: 3.1 Pain.

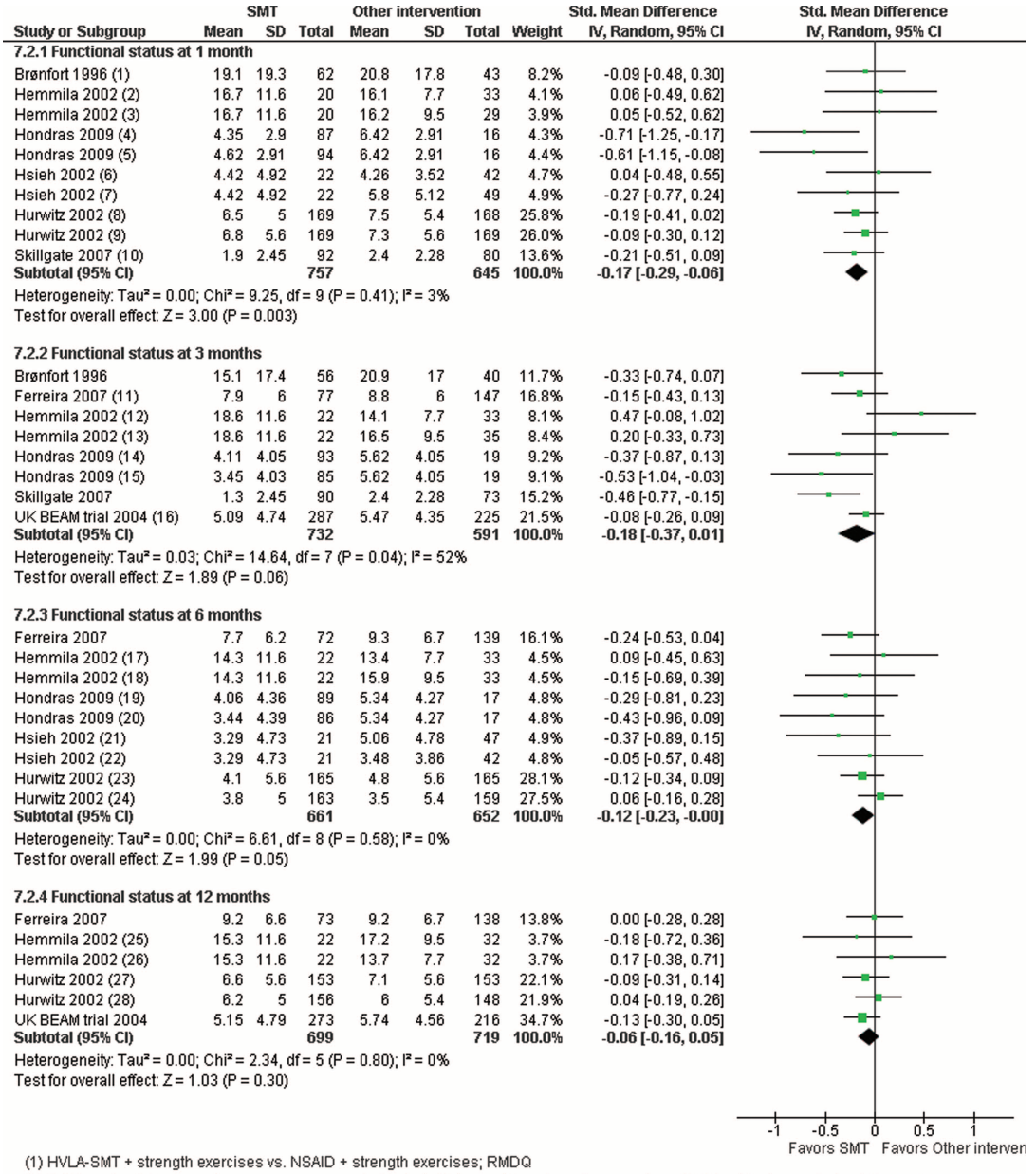

Figure 4. Forest plot of comparison: 3 Spinal manipulative therapy vs. all other therapies, outcome: 3.2 Functional status.

there is low-quality evidence (high RoB, imprecision), which demonstrates no statistically significant difference in effect on pain relief at 6 months (MD −6.77, 95% CI −14.07 to 0.53).

Three studies examined functional status, two of which had a low RoB.[45,47,54] There is low-quality evidence (high RoB, imprecision) from two studies (156 participants) that SMT has a statistically significant effect on functional status at 1 month (SMD −0.40, 95% CI −0.73 to −0.07) and high-quality evidence from two studies[47,54] at 3 months (1078 participants) that SMT has a statistically significant effect on functional status (SMD −0.22, 95% CI −0.38 to −0.06) and a statistically significant effect at 12 months (SMD −0.21, 95% CI −0.34 to −0.09). However, there is low-quality evidence (high RoB, imprecision) that SMT has no statistically significant effect at 6 months (SMD −0.30, 95% CI −0.64 to 0.03).

One study with a high RoB examined perceived recovery.[37] There is very low-quality evidence (high RoB, inconsistency, imprecision) that SMT demonstrates significantly greater recovery at 1 month than the comparison group (RR 3.40, 95% CI 1.12–10.28). No data were available on return to work or HRQoL.

SENSITIVITY ANALYSES

Sensitivity analyses were conducted for the comparison SMT *versus* all other interventions only. Only two outcomes were examined, pain and functional status. The sparseness of data for the other comparisons rendered further subanalyses meaningless. These analyses were conducted to determine the robustness of our original analyses and determine whether other factors might have influenced the overall pooled effect.

On the basis of these sensitivity analyses, results appear more prominently for those studies with a low RoB because heterogeneity across studies was much less than when all studies were pooled; however, the overall pooled effect between all studies and those with a low RoB were only marginally different for pain and functional status at all time measurements. It is noteworthy that a small difference in effect was observed for SMT *versus* interventions thought to be ineffective as opposed to SMT *versus* interventions thought to be effective; however, this amounted to a difference of at most, 5 points on a 100-point scale (for pain at 1 month) or 0.3 points in SMD (for functional status at 1 month). However, none of these analyses suggested a clinically relevant effect on pain or functional status at any time interval not observed in the primary analyses. Furthermore, with the exception of two studies,[37,55] both with a high RoB, no other study demonstrated a clinically relevant effect for any comparison or time interval for the primary outcomes, pain, functional status, or perceived recovery. The sensitivity analyses were less remarkable at the remaining time intervals and all are available upon request.

We wanted to examine the effect of SMT in subjects with radiating pain; however, most studies included subjects with or without radiating pain and did not present separate analyses, so this sensitivity analysis was not performed. Finally, while it was not part of the original sensitivity analysis, lowering the threshold value for I^2 to 40% would not have had any bearing on the presentation of these results.

DISCUSSION

Summary of Main Results

In general, there is high-quality evidence that SMT has a statistically significant short-term effect on pain relief and functional status in comparison with other interventions as well as varying quality of the evidence that SMT has a statistically significant short-term effect on pain relief and functional status when SMT is added to another intervention. However, the size of the effects was small and not apparently clinically relevant. In addition, there is very low-quality evidence that SMT is no more effective than inert interventions or sham SMT for short-term pain relief or functional status. Seemingly, these results are conflicting. This might be explained by the fact that relatively few, small studies, quite typically with a high RoB, evaluated the latter comparisons; thus, these studies have a high likelihood of a type II error stemming from the low power of the study to detect a statistically significant and clinically relevant effect. However, studies with a high RoB typically overestimate the effect in comparison with studies with a low RoB,[34] so it is unclear to what extent, if any, various forms of bias have on those results. Furthermore, it is questionable to what extent studies investigating sham SMT were able to successfully blind their subjects, as only one study evaluated this posttreatment, suggesting that the investigators were not entirely successful; so it is debatable whether these data can be considered representative for this comparison. Nevertheless, improper blinding is likely to lead to an overestimation of the effect, not underestimation; thus, it is also difficult to interpret the essence of these findings in relation to our more robust comparison, SMT *versus* other interventions. Data were particularly sparse for recovery, return to work, and quality of life, in addition to costs of care; therefore, no firm conclusions can be drawn regarding these outcomes.

Recently, there has been much discussion regarding the clinical importance of small effects identified in continuous outcomes, such as those examined in this review. Formerly, it was thought that the effect of a treatment was trivial if the mean difference between the treatment and a control group was appreciably less than the smallest change thought to be clinically important. This might not necessarily be so.[94] The addition of the number needed to treat (NNT) may aid interpretation of trials with continuous outcomes, especially when expressed as an SMD.[95] For example, the largest benefit demonstrated from any of the treatments in the UK BEAM (2004) trial was 1.87 points on the Roland-Morris disability questionnaire, which translates to a between-group difference that is not clinically important.[96] A recent reanalysis of these data suggests that despite the small mean differences between interventions, NNTs were small, on average, four to five for manipulation plus exercise or manipulation alone, respectively in comparison to "best care" at 3 months' follow-up.[95] This means that referring four to five patients for manipulation, would, on average, yield one additional case of improvement. Even a conservative estimate with these data suggests a potentially attractive NNT ratio. However, it should be noted that this represents a *post hoc* analysis and there are some general

limitations to the use of NNT analyses.[97] Furthermore, calculation of an NNT is based on determination of a threshold value of improvement, which is also open for discussion. Finally, statistical power is lost when converting scales to binary outcomes; therefore, this technique might be attractive only when sample sizes are sufficiently large.[94]

Despite the methodological rigor maintained in this review, there are likely to be objections. One objection typically raised by clinicians is the lack of respect to the type of manipulative therapy delivered (*e.g.*, high-velocity low-amplitude manipulation *vs.* mobilization) or profession of the therapist (*e.g.*, chiropractor *vs.* manual therapist or physiotherapist). Sensitivity analyses were conducted to distinguish whether this resulted in a different effect; however, those results suggest that neither the technique nor the profession of the therapist had a profound influence on the overall pooled effect.

Another objection might lie with the lack of examining a more homogenous group of subjects with low-back pain. Nonspecific low-back pain, even when examined by duration, can probably be viewed as a rather heterogeneous group. Even within this review, a number of studies included subjects with and without radiating pain; therefore, defining a homogenous population and identifying subgroups seem important. Recent work suggests that clinically important effects are observed when treatment is matched to the patient's signs and symptoms rather than provided to all patients with low-back pain.[98] Furthermore, recent recommendations from a UK consensus, which included senior researchers experienced in clinical trials for musculoskeletal conditions, include examining subgroups.[99]

None of the included studies that examined adverse events reported serious complications. Serious complications following SMT for low-back pain are extremely rare and have been documented in case reports only, which include cauda equina syndrome , paraplegia, and death.[5] Risk estimates vary widely for cauda equina syndrome, ranging from less than 1 case per million treatments[100] to 1 case per 100 million manipulations.[101] Given the extremely low incidence of serious complications, a review of RCTs provides limited information; however, estimates based on case reports are likely to underestimate risk, while large prospective cohorts are lacking. To our knowledge, only one systematic review has examined the safety of SMT to the low-back based on case reports and surveys, which concluded that the risk of SMT causing a clinically worsened disc herniation or cauda equina syndrome in a patient presenting with lumbar disc herniation to be estimated at 1 in 3.7 million treatments.[102]

Limitations and Strengths

There are a number of limitations to this review. The primary limitation, which is common to many systematic reviews, is the lack of studies with a low RoB. Despite the fact that the majority of the studies included in this review were published in the last decade, methodologically well-conducted studies remain scarce.

A second limitation is the possibility of publication bias, which we attempted to minimize through an extensive database search. We did not actively seek unpublished studies; however, it could be argued that this is unlikely to have had an important impact on the overall results. Surprisingly, many of the studies published in the last decade did not have a published protocol and, to our knowledge, had not registered their study in one of the many trial registries, indicating that many trials conducted in the 21st century still do not conform to international procedure. In the absence of 100% conformity, it remains difficult to ascertain to what extent studies do not publish their findings because the results prove less than favorable. In addition, we uncovered a couple of irregularities, for example, a study that began recruitment 10 years ago but has not yet been published (ISRCTN61808774) or another study that was terminated without further explanation (NCT00269503).

Finally, we would have liked to have conducted a meta-regression for the purpose of exploring heterogeneity between studies; however, there were too few studies per outcome to allow for a meaningful analysis and the distribution of data for the outcomes, pain and functional status, appeared to be clustered; that is, the data did not follow a normal distribution. Furthermore, results from the sensitivity analyses did not suggest any important directions of effect for the confounders and effect modifiers examined.

Strengths of this review include the methodological rigor applied, including a published protocol and the meta-analyses, which allowed us to conduct meaningful sensitivity analyses.

Agreements and Disagreements with Other Studies or Reviews

Ostensibly, these results are consistent with the previous review, which concluded that there is evidence that SMT is neither superior nor inferior to other effective treatments for patients with chronic low-back pain. In comparison to the previous review,[12] approximately two-thirds of the studies included are new and many more studies have been included with a low RoB; therefore, our findings are thought to be much more robust. These results are also consistent with other recent systematic reviews, which conducted principally narrative analyses[7,24,103]; however, the findings from our review are more optimistic than another review,[104] which conducted meta-analyses. Another systematic review was identified, which pooled data from six trials of osteopathic manipulative therapy and concluded that osteopathic manipulative therapy significantly reduces low-back pain[105]; however, that review did not limit the results to trials of chronic low-back pain. A recent review of systematic reviews, including the earlier version of this review, concluded that SMT produces small clinical benefits that are equivalent to those of other commonly used therapies.[5]

CONCLUSIONS

Implications for Practice

High-quality evidence suggests that there is no clinically relevant difference between SMT and other interventions for reducing pain and improving function in patients with chronic low-back pain. Therefore, the decision to refer for SMT

should be based on costs, preferences of the patient and providers, and relative safety of the treatment options.

Implications for Research

Future studies should:

1. Evaluate the effects of SMT as an additional or adjunct therapy, for example, in the case of SMT in multimodal treatment packages.
2. There is a dire need for cost-effectiveness studies. If SMT is equal to other presumed effective interventions for chronic low-back pain, SMT may be more cost-effective because the therapy is typically provided in a limited number of treatment sessions (as compared to, *e.g.*, exercise therapy or behavioral treatment).

➢ Key Points

- ❑ A systematic review was conducted to assess the effects of SMT for chronic low-back pain.
- ❑ 26 RCTs were identified (N = 6070), nine of which had a low risk of bias.
- ❑ There is high-quality evidence that SMT has a small, statistically significant but not clinically relevant, short-term effect on pain relief and functional status in comparison with other interventions.
- ❑ No serious complications were observed with SMT.
- ❑ The decision to refer for SMT should be based on costs, preferences of the patient and providers, and relative safety of the treatment options.

Acknowledgments

The authors thank the members of the editorial board of the Cochrane Back Review Group for their constructive comments on the protocol and draft version of this review and Ms. Rachel Couban for her assistance with the development of the search strategies. They also thank Dr. Sally Morton, Mrs. Emily Yu, Ms. Marika Suttorp, and Dr. Paul Shekelle for their work on the original review, which laid the groundwork for this update. In addition, they thank Veronica Morton, Nina Zaproudina, Cynthia Long, John Licciardone, and Peter McCarthy for providing additional data not found in their original publications.

References

1. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 2008;8:8–20.
2. Brown A, Angus D, Chen S, et al. *Costs and Outcomes of Chiropractic Treatment for Low Back Pain [technology report no. 56]*. Ottawa, Ontario, Canada: Canadian Coordinating Office for Health Technology Assessment; 2007.
3. Brox JI, Hagen KB, Juel NG, et al. Is exercise therapy and manipulation effective in low back pain? *Tidsskr Nor Laegeforen* 1999;119:2042–50.
4. Brønfort G, Haas M, Evans RL, et al. Efficacy of spinal manipulation and mobilization for low back pain and neck pain: a systematic review and best evidence synthesis. *Spine J* 2004;4: 335–56.
5. Cherkin DC, Sherman KJ, Deyo RA, et al. A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Intern Med* 2003;138:898–906.
6. Airaksinen O, Brox JI, Cedraschi C, et al. European guidelines for the management of chronic nonspecific low back pain, chapter 4. *Eur Spine J* 2006;15(suppl 2):S192–298.
7. Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med* 2007;147:492–504.
8. Manchikanti L, Staats PS, Singh V, et al. Evidence-based practice guidelines for interventional techniques in the management of chronic spinal pain. *Pain Physician* 2003;6:3–81.
9. Staal JB, Hlobil H, van Tulder MW, et al. Occupational health guidelines for the management of low back pain: an international comparison. *Occup Environ Med* 2003;60:618–26.
10. van Tulder MW, Becker A, Bekkering T, et al. European guidelines for the management of acute nonspecific low back pain in primary care [2004], chapter 3. *Eur Spine J* 2006;15(suppl 2):S169–91.
11. Waddell G, Hutchinson A, Feder G, et al. *Clinical Guidelines for the Management of Acute Low Back Pain*. Royal College of General Practitioners; 2001.
12. Assendelft WJJ, Morton SC, Yu EI, et al. Spinal manipulative therapy for low-back pain. *Cochrane Database of Syst Rev* 2004;(1):CD000447. DOI: 10.1002/14651858.CD000447.pub2.
13. Koes BW, van Tulder MW, Ostelo R, et al. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine* 2001;26:2504–13.
14. van Tulder MW, Tuut M, Pennick V, et al. Quality of primary care guidelines for acute low back pain. *Spine* 2004;29:E357–62.
15. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration Version 5.0.0. www.cochrane.org/resources/handbook/index.htm. Updated February 2008.
16. Furlan AD, Pennick V, Bombardier C, et al. 2009. Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;34:1929–41.
17. Rubinstein SM, Terwee C, Assendelft WJJ, et al. Spinal manipulation for acute low-back pain. *Cochrane Database Syst Rev* 2010;12:CD008880. DOI: 10.1002/14651858.CD008880.
18. Rubinstein SM, van Middelkoop M, Assendelft WJJ, et al. Spinal manipulative therapy for chronic low back pain. *Cochrane Database Syst Rev* 2009;2(4):CD008112. DOI:10.1002/14651858.CD008112.
19. Spitzer WO, LeBlanc FE, DuPuis M. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. [Report of the Quebec Task Force on Spinal Disorders]. *Spine* 1987;12(7):S1–S59.
20. Sandoz R. The significance of the manipulative crack and of other articular noises. *Ann Swiss Chiro Assoc* 1969;4:47–68.
21. Evans DW. Mechanisms and effects of spinal high-velocity, low-amplitude thrust manipulation: previous theories. *J Manipulative Physiol Ther* 2002;25:251–62.
22. Unsworth A, Dowson D, Wright V. "Cracking joints." A bioengineering study of cavitation in the metacarpophalangeal joint. *Ann Rheum Dis* 1971;30:348–58.
23. van de Veen EA, de Vet HC, Pool JJ, et al. Variance in manual treatment of nonspecific low back pain between orthomanual physicians, manual therapists, and chiropractors. *J Manipulative Physiol Ther* 2005;28:108–16.
24. Brønfort G, Evans R, Kawchuk G, et al. Evidence-informed management of chronic low back pain with spinal manipulation and mobilization. *Spine J* 2008;8:213–25.
25. Khalsa PS, Eberhart A, Cotler A, et al. The 2005 conference on the biology of manual therapies. *J Manipulative Physiol Ther* 2006;29:341–6.
26. Pickar JG. Neurophysiological effects of spinal manipulation. *Spine J* 2002;2:357–71.
27. Triano JJ. Biomechanics of spinal manipulative therapy. *Spine J* 2001;1:121–30.
28. Assendelft WJ, Lankhorst GJ. Effectiveness of manipulative therapy for low-back pain: systematic literature reviews and guidelines are inconclusive. *Ned Tijdschr Geneeskd* 1998;142:684–7.
29. Bouter LM, van Tulder MW, Koes BW. Methodological issues in low back pain research in primary care. *Spine* 1998;23: 2014–20.

30. Brønfort G, Evans RL, Maiers M, et al. Spinal manipulation, epidural injections, and self-care for sciatica: a pilot study for a randomized clinical trial. *J Manipulative Physiol Ther* 2004;27:503–8.
31. Axen I, Jones JJ, Rosenbaum A, et al. The Nordic Back Pain Subpopulation Program: validation and improvement of a predictive model for treatment outcome in patients with low back pain receiving chiropractic treatment. *J Manipulative Physiol Ther* 2005;28: 381–5.
32. Malmqvist S, Leboeuf-Yde C, Ahola T, et al. The Nordic back pain subpopulation program: predicting outcome among chiropractic patients in Finland. *Chiropr Osteopat* 2008;16:13.
33. Goldby LJ, Moore AP, Doust J, et al. A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. *Spine* 2006;31:1083–93.
34. van Tulder MW, Suttorp M, Morton S, et al. Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine* 2009;34:1685–92.
35. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 1st ed. New York, San Francisco, London: Academic Press; 1988:1–474.
36. Ferreira ML, Ferreira PH, Latimer J, et al. Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. *Pain* 2007;131: 31–7.
37. Evans DP, Burke MS, Lloyd KN, et al. Lumbar spinal manipulation on trial. Part 1: clinical assessment. *Rheumatol Rehabil* 1978;17:46–53.
38. Gibson T, Grahame R, Harkness J, et al. Controlled comparison of short-wave diathermy treatment with osteopathic treatment in non-specific low back pain. *Lancet* 1985;8440:1258–61.
39. Koes BW, Bouter LM, Mameren H, et al. The effectiveness of manual therapy, physiotherapy, and treatment by the general practitioner for nonspecific back and neck complaints. A randomized clinical trial. *Spine* 1992;17:28–35.
40. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
41. Chown M, Whittamore L, Rush M, et al. A prospective study of patients with chronic back pain randomised to group exercise, physiotherapy or osteopathy. *Physiotherapy* 2008;94(1):21–8.
42. Ghroubi S, Elleuch H, Baklouti S, et al. Chronic low back pain and vertebral manipulation [in French]. *Ann Readapt Med Phys* 2007;50:570–6.
43. Gudavalli MR, Cambron JA, McGregor M, et al. A randomized clinical trial and subgroup analysis to compare flexion-distraction with active exercise for chronic low back pain. *Eur Spine J* 2006;15:1070–82.
44. Hondras MA, Long CR, Cao Y, et al. A randomized controlled trial comparing 2 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;32:330–43.
45. Hsieh CY, Adams AH, Tobis J, et al. Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. *Spine* 2002;27:1142–8.
46. Hurwitz EL, Morgenstern H, Harber P, et al. A randomized trial of medical care with and without physical therapy and chiropractic care with and without physical modalities for patients with low back pain: 6-month follow-up outcomes from the UCLA low back pain study. *Spine* 2002;27:2193–204.
47. Licciardone JC, Stoll ST, Fulda KG, et al. Osteopathic manipulative treatment for chronic low back pain: a randomized controlled trial. *Spine* 2003;28:1355–62.
48. Mohseni-Bandpei MA, Critchley J, Staunton T, et al. A prospective randomised controlled trial of spinal manipulation and ultrasound in the treatment of chronic low back pain. *Physiotherapy* 2006;92:34–42.
49. Muller R, Giles LG. Long-term follow-up of a randomized clinical trial assessing the efficacy of medication, acupuncture, and spinal manipulation for chronic mechanical spinal pain syndromes. *J Manipulative Physiol Ther* 2005;28:3–11.
50. Paatelma M, Kilpilkoski S, Simonen R, et al. Orthopaedic manual therapy, McKenzie method or advice-only for low back pain in working adults: a randomized controlled trial with one year follow-up. *J Rehabil Med* 2008;40:858–63.
51. Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. *Manual Ther* 2003;8:233–41.
52. Rasmussen J, Laetgaard J, Lindecrona A, et al. Manipulation does not add to the effect of extension exercises in chronic low-back pain (LBP). A randomized, controlled, double-blind study. *Joint Bone Spine: Revue du Rhumatisme* 2008;75:708–13.
53. Skillgate E, Vingard E, Alfredsson L. Naprapathic manual therapy or evidence-based care for back and neck pain: a randomized, controlled trial. *Clin J Pain* 2007;23:431–9.
54. UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: effectiveness of physical treatments for back pain in primary care. *BMJ* 2004;329: 1377–81.
55. Wilkey A, Gregory M, Byfield D, et al. A comparison between chiropractic management and pain clinic management for chronic low-back pain in a national health service outpatient clinic. *J Altern Complement Med* 2008;14:465–73.
56. Zaproudina N, Hietikko T, Hanninen OOP, et al. Effectiveness of traditional bone setting in treating chronic low back pain: a randomized pilot trial. *Compl Therap Med* 2009;17:23–8.
57. Brønfort G, Goldsmith CH, Nelson CF, et al. Trunk exercises combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. *J Manip Physiol Ther* 1996;19:570–82.
58. Pope MH, Phillips RB, Haugh LD, et al. A prospective randomized three-week trial of spinal manipulation, transcutaneous muscle stimulation, massage and corset in the treatment of subacute low back pain. *Spine* 1994;22:2571–7.
59. Postacchini F, Facchini M, Palieri P. Efficacy of various forms of conservative treatment in low back pain: a comparative study. *Neuro-Orthop* 1988;6:28–35.
60. Waagen GN, Haldeman S, Cook G, et al. Short term trial of chiropractic adjustments for the relief of chronic low back pain. *Manual Med* 1986;2:63–7.
61. Hemmila HM, Keinanen-Kiukaanniemi SM, Levoska S, et al. Long-term effectiveness of bone-setting, light exercise therapy, and physiotherapy for prolonged back pain: a randomized controlled trial. *J Manipulative Physiol Ther* 2002;25:99–104.
62. Andersson GB, Lucente T, Davis AM, et al. A comparison of osteopathic spinal manipulation with standard care for patients with low back pain. *N Engl J Med* 1999;341:1426–31.
63. Beyerman KL, Palmerino MB, Zohn LE, et al. Efficacy of treating low back pain and dysfunction secondary to osteoarthritis: chiropractic care compared with moist heat alone. *J Manipulative Physiol Ther* 2006;29:107–14.
64. Coxhead CE, Inskip H, Meade TW, et al. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet* 1981;1:1065–8.
65. Doran DML, Newell DJ. Manipulation in treatment of low back pain: a multicentre study. *BMJ* 1975;2:161–14.
66. Glover JR, Morris JG, Khosla T. Back pain: a randomized clinical trial of rotational manipulation of the trunk. *Br J Ind Med* 1974;31:59–64.
67. Herzog W, Conway PJW, Wilcox BJ. Effects of different treatment modalities on gait symmetry and clinical measures for sacroiliac joint patients. *J Manipulative Physiol Ther* 1991;14:104–9.
68. Kinalski R, Kuwik W, Pietrzak D. The comparison of the results of manual therapy versus physiotherapy methods used in treatment of patients with low back pain syndromes. *J Manu Med* 1989;4:44–6.
69. MacDonald RS, Bell CJM. An open controlled assessment of osteopathic manipulation in non-specific low-back pain. *Spine* 1990;15:364–70.
70. Meade TW, Dyer S, Browne W, et al. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ* 1990;300:1431–7.
71. Rupert RL, Wagnon R, Thompson P, et al. Chiropractic adjustments: results of a controlled clinical trial in Egypt. *ICA Int Rev Chirop* Winter 1985:58–60.

72. Shearar KA, Colloca CJ, White HL. A randomized clinical trial of manual versus mechanical force manipulation in the treatment of sacroiliac joint syndrome. *J Manipulative Physiol Ther* 2005;28:493–501.
73. Sims-Williams H, Jayson MI, Young SM, et al. Controlled trial of mobilisation and manipulation for patients with low back pain in general practice. *Br Med J* 1978;2:1338–40.
74. Sims-Williams H, Jayson MI, Young SM, et al. Controlled trial of mobilisation and manipulation for low back pain: hospital patients. *Br Med J* 1979;2:1318–20.
75. Triano JJ, McGregor M, Hondras MA, et al. Manipulative therapy versus education in chronic low back pain. *Spine* 1995;20: 948–55.
76. Zylbergold RS, Piper MC. Lumbar disc disease: comparative analysis of physical therapy treatments. *Arch Phys Med Rehabil* 1981;62:176–9.
77. Brønfort G. Chiropractic versus general medical treatment of low back pain: a small scale controlled clinical trial. *Am J Chirop* 1989;2:145–50.
78. Cherkin DC, Deyo RA, Battie M, et al. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. *N Engl J Med* 1998;339:1021–9.
79. Hoehler FK, Tobis JS, Buerger AA. Spinal manipulation for low back pain. *JAMA* 1981;245:1835–8.
80. Mathew JA, Mills SB, Jenkins VM, et al. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol* 1987;26:416–23.
81. Skargren EI, Oberg BE, Carlsson PG, et al. Cost and effectiveness analysis of chiropractic and physiotherapy treatment for low back and neck pain. Six-month follow-up. *Spine* 1997;22:2167–77.
82. Aure OF, Nilsen JH, Vasseljen O. Manual therapy and exercise therapy in patients with chronic low back pain: a randomized, controlled trial with 1-year follow-up. *Spine* 2003;28:525–31.
83. Haas M, Groupp E, Kraemer DF. Dose-response for chiropractic care of chronic low back pain. *Spine J* 2004;4:574–83.
84. Niemisto L, Lahtinen-Suopanki T, Rissanen P, et al. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. *Spine* 2003;28:2185–91.
85. Ongley MJ, Klein RG, Dorman TA, et al. A new approach to the treatment of chronic low back pain. *Lancet* 1987;2:143–6.
86. Arkuszewski Z. The efficacy of manual treatment in low back pain: a clinical trial. *Manu Med* 1986;2:68–71.
87. Coyer AB, Curwin I. Low back pain treated by manipulation. *Br Med J* 1955;1:705–7.
88. Hough E, Stephenson R, Swift L. A comparison of manual therapy and active rehabilitation in the treatment of non specific low back pain with particular reference to a patient's Linton & Hallden psychological screening score: a pilot study. *BMC Musculoskeletal Disorders* 2007;8:106–16.
89. Nwuga VCB. Relative therapeutic efficacy of vertebral manipulation and conventional treatment in back pain management. *Am J Phys Med* 1982;1:160–4.
90. Petty NJ. The effect of posteroanterior mobilisation on sagittal mobility of the lumbar spine. *Manu Ther* 1995;1:25–9.
91. Burton AK, Tillotson KM, Cleary J. Single-blind randomised controlled trial of chemonucleolysis and manipulation in the treatment of symptomatic lumbar disc herniation. *Eur Spine J* 2000;9:202–7.
92. Timm KE. A randomized-control study of active and passive treatments for chronic low back pain following L5 laminectomy. *J Orthop Sports Phys Ther* 1994;20:276–86.
93. Geisser ME, Wiggert EA, Haig AJ, et al. A randomized, controlled trial of manual therapy and specific adjuvant exercise for chronic low back pain. *Clin J Pain* 2005;21:463–70.
94. Guyatt GH, Juniper EF, Walter SD, et al. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690–3.
95. Froud R, Eldridge S, Lall R, et al. Estimating the number needed to treat from continuous outcomes in randomised controlled trials: methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. *BMC Med Res Methodol* 2009;9:35.
96. Tveito TH, Eriksen HR. United Kingdom back pain exercise and manipulation (UK BEAM) trial: is manipulation the most cost-effective addition to "best care"? *BMJ* 2005;330:674.
97. Wu LA, Kottke TE. Number needed to treat: caveat emptor. *J Clin Epidemiol* 2001;54:111–6.
98. Brennan GP, Fritz JM, Hunter SJ, et al. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain. *Spine* 2006;31:623–31.
99. Foster NE, Dziedzic KS, Van Der Windt DA, et al. Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations. *BMC Musculoskeletal Disorders* 2009;10:3.
100. Assendelft WJ, Bouter LM, Knipschild PG. Complications of spinal manipulation: a comprehensive review of the literature. *J Fam Pract* 1996;42:475–80.
101. Shekelle PG, Adams AH, Chassin MR, et al. Spinal manipulation for low-back pain. *Ann Intern Med* 1992;117:590–8.
102. Oliphant D. Safety of spinal manipulation in the treatment of lumbar disc herniations: a systematic review and risk assessment. *J Manipulative Physiol Ther* 2004;27:197–210.
103. Lawrence DJ, Meeker W, Branson R, et al. Chiropractic management of low back pain and low back-related leg complaints: a literature synthesis. *J Manipulative Physiol Ther* 2008;31:659–74.
104. Ferreira ML, Ferreira PH, Latimer J, et al. Does spinal manipulative therapy help people with chronic low back pain? *Aust J Physiother* 2002;48:277–84.
105. Licciardone JC, Brimhall AK, King LN. Osteopathic manipulative treatment for low back pain: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskeletal Disorders* 2005;6:43.