QUERI

The Effectiveness and Harms of Spinal Manipulative Therapy for the Treatment of Acute Neck and Lower Back Pain: A Systematic Review

April 2017

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP)
West Los Angeles VA Medical Center
Los Angeles, CA
Paul G. Shekelle, MD, PhD, Director

Investigators:
Principal Investigator:
Paul G. Shekelle, MD, PhD

Co-investigators:
Neil M. Paige, MD, MSHS

Research Associates:
Isomi M. Miake-Lye, BA
Jessica M. Beroes, BS
Marika Suttorp Booth, MS
Roberta Shanman, MS

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

PREFACE

Quality Enhancement Research Initiative's (QUERI) Evidence-based Synthesis Program (ESP) was established to provide timely and accurate syntheses of targeted healthcare topics of particular importance to Veterans Affairs (VA) clinicians, managers and policymakers as they work to improve the health and healthcare of Veterans. The ESP disseminates these reports throughout the VA, and some evidence syntheses inform the clinical guidelines of large professional organizations.

QUERI provides funding for 4 ESP Centers and each Center has an active university affiliation. The ESP Centers generate evidence syntheses on important clinical practice topics, and these reports help:

- develop clinical policies informed by evidence;
- guide the implementation of effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- set the direction for future research to address gaps in clinical knowledge.

In 2009, the ESP Coordinating Center was created to expand the capacity of HSR&D Central Office and the 4 ESP sites by developing and maintaining program processes. In addition, the Center established a Steering Committee comprised of QUERI field-based investigators, VA Patient Care Services, Office of Quality and Performance, and Veterans Integrated Service Networks (VISN) Clinical Management Officers. The Steering Committee provides program oversight, guides strategic planning, coordinates dissemination activities, and develops collaborations with VA leadership to identify new ESP topics of importance to Veterans and the VA healthcare system.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP Coordinating Center Program Manager, at Nicole.Floyd@va.gov.

Recommended citation: Shekelle PG, Paige NM, Miake-Lye IM, Beroes JM, Booth MS, Shanman R. The Effectiveness and Harms of Chiropractic Care for the Treatment of Acute Neck and Lower Back Pain: A Systematic Review. VA ESP Project #05-226; 2017.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the **West Los Angeles VA Medical Center, Los Angeles, CA**, funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

INTRODUCTION

Back pain and neck pain are among the most common symptoms prompting patients to seek care. Many treatments are used for back pain. Spinal manipulative therapy (SMT) is a treatment option available in VA. In order to better understand the potential role of SMT in treating acute back or neck pain, VA requested an up-to-date synthesis of the evidence.

The Key Questions are:

Key Question 1: What are the benefits and harms of spinal manipulation/chiropractic services for acute lower back pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Key Question 1A: What is the relationship between the use of spinal manipulation/chiropractic services for lower back pain and the use of opiate medication?

Key Question 2: What are the benefits and harms of spinal manipulation/chiropractic services for acute neck pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Key Question 2A: What is the relationship between the use of spinal manipulation/chiropractic services for acute neck pain and the use of opiate medication?

METHODS

Data Sources and Searches

Spinal manipulation is a topic that has been the subject of numerous prior systematic reviews, including 3 reviews by members of the ESP review team. Therefore, instead of searching for original evidence in databases such as PubMed, we instead began with reference mining existing systematic reviews, and then performing an update search to identify new studies published since the end date of the searches of the most recent reviews. Then we consulted our technical experts for any additional studies we might have overlooked.

Study Selection

Participants: Adults with acute (defined as 6 weeks or less) neck or lower back pain. Patients with sciatica were included. Studies of patients with chronic back pain were excluded, as were studies where we could not determine the duration of pain. If studies included patients with longer durations of pain, we included them if they presented stratified results or if the majority of patients had pain for less than 6 weeks duration. Studies of children were excluded.

Intervention: Spinal manipulation by any provider type. Studies where spinal manipulation was given alone or as part of a “package” of therapies were included. “Chiropractic care” was considered as including SMT for the great majority of patients.

Comparator (study design): Other forms of management for acute pain, such as analgesics, exercises, physical therapy, etcetera. Sham-controlled studies were included.

Outcome: Pain management, functional status, quality of life, opiate use, disability claims, return to work, health care utilization.

Timing: Studies had to report at least one outcome within 6 weeks to be eligible.

Setting: Ambulatory/outpatient settings. Studies in hospital settings were excluded.

Study design: Only randomized controlled trials (RCTs) were eligible for assessing benefits. Both RCTs plus observational studies were used for assessing harms.

Data Abstraction and Quality Assessment

Data were extracted by 2 reviewers, and discrepancies were reconciled after discussion. Articles had data abstracted on the anatomical location of the pain, authors' description of the SMT provided, type of professional performing the treatment, co-interventions, whether that treatment was provided alone or as part of a package of other treatments, whether patients were selected as more likely to respond to SMT or unselected, data on any of the outcomes listed above (*eg*, pain, functional status, *etc*), as well as data needed to complete the Cochrane Back Group Risk of Bias assessment.

We assessed the quality of studies using the Cochrane Back Group Risk (CBG) of Bias Tool (ROB). This tool has 11 items in the following domains: randomization; concealment; baseline differences; blinding – patient; blinding – care provider; blinding – outcome; co-interventions; compliance; dropouts; timing; and intent to treat.

Data Synthesis and Analysis

We constructed evidence tables showing the study characteristics and results for all included studies. Random effects meta-analyses were conducted using the Hartung-Knapp Method.

RESULTS

Results of Literature Search

We identified 181 potentially relevant titles from our systematic review search and identified one additional title from the references of one of our included articles, for a total of 182 titles for screening. From the 49 systematic reviews we mined for references, we identified 136 potentially relevant titles. To this we added 15 titles recommended by experts and 1,639 titles identified in an update search for a total of 1,790 titles for screening.

After excluding 1,564 titles as clearly not relevant, we reviewed 226 abstracts. Of these, we excluded 28 abstracts and included 198 abstracts for full-text review. After full-text review, we excluded 150 articles: 77 articles rejected as studying patients with pain longer than 6 weeks or unspecified; 38 articles rejected for study design (*ie*, not a randomized controlled trial); 10 articles rejected as duplicate articles of already-screened articles; 9 articles rejected as providing relevant background information but not otherwise included; 7 articles rejected as not reporting

on SMT; 3 articles rejected for having no relevant outcome; 2 articles rejected for studying patients in hospital; 3 articles rejected for other reasons; and 1 article we were unable to retrieve.

Of the 48 included articles, we identified 40 articles relevant to effectiveness of SMT and 8 articles relevant to adverse events. Of the 40 effectiveness articles, 26 were included in the analyses. Of the 14 not included in the analyses, 3 publications were focused on the subpopulation of patients with sciatica, 2 publications were only relevant to clinical prediction discussions, 2 publications did not have the necessary outcome data, and one publication had a unique patient population judged by our TEP as clinically dissimilar to the other studies.

Quality Assessment

In the low back pain analysis, one study scored a high of 9 out of 11 possible points, 6 studies scored 7 points, 4 studies scored 6 points, 2 studies scored 4 points, 7 studies scored 3, and 6 studies scored 2 points (see Table 1).

Of the 26 studies, 25 studies met the timing criteria and 17 met the randomization criteria. None of the studies met the blinding of providers criteria, and only 4 met the criteria for blinding of patients. Using a threshold of 6, 12 studies were classified as high quality and 14 studies were classified as low quality.

Acute Low Back Pain without Sciatica

Twenty studies reported results that we could use for meta-analytic pooling.

Immediate-term Pain (less than 2 weeks)

There were 11 studies reporting immediate-term pain outcomes using a VAS or numeric rating scale, 2 comparing SMT to sham, and 9 comparing SMT to another therapy (Figure 3). The overall random effects pooled estimate was -8.49 mm (95% CI: -16.46, -0.52) favoring treatment with SMT. There was heterogeneity, with an $I^2 = 76.1\%$. There was no evidence of publication bias in the overall pooled result, with Begg's rank correlation = 0.15 and Egger's test p-value of 0.58. Two studies comparing SMT to sham reported non-statistically significant benefits.

Immediate-term Function (less than 2 weeks)

There were 10 studies reporting immediate-term function measured with the RMDQ or ODI, 3 comparing SMT to sham, and 7 comparing SMT to another therapy (Figure 4). The overall random effects pooled estimate was an effect size of -0.24 (95% CI: -0.55, 0.08) favoring treatment with SMT. There was heterogeneity, with an $I^2 = 52.1\%$. There was no evidence of publication bias in the overall pooled result, with Begg's rank correlation = 0.17 and Egger's test p-value of 0.14. Three studies compared SMT to sham and the overall random effects pooled estimate was an effect size of -0.14 (95% CI: -0.26, -0.11).

Short-term Pain (3-6 weeks)

There were 12 studies reporting short-term pain using VAS or numeric rating scale, 2 comparing SMT to sham, and 10 comparing SMT to another therapy (Figure 4). The overall random effects pooled estimate across all studies was an effect size of -9.95 mm (95% CI: -15.6, -4.3) favoring treatments with SMT. There was heterogeneity, with an $I^2 = 67.2\%$. There was no evidence of

publication bias in the overall pooled result, with Begg's rank correlation of 0.92 and Egger's test p-value of 0.58.

Short-term Function (3-6 weeks)

There were 8 studies reporting short-term function outcomes measured with the RMDQ or ODI, 2 comparing SMT to sham and, 6 comparing SMT to another therapy (Figure 5). The overall random effects pooled estimate was an effect size of -0.39 (95% CI: -0.71, -0.07). There was heterogeneity, with an $I^2 = 72.1\%$. There was no evidence of publication bias, with Begg's rank correlation = 0.85 and Egger's test p-value = 0.10. Two studies comparing SMT to sham reported non-statistically significant benefits.

Exploring Sources of Heterogeneity

There was significant heterogeneity in almost all the pooled analyses of SMT, suggesting that there are other factors influencing the outcome. In addition to the comparison group, we investigated 5 possible sources of heterogeneity: outcome, timing of the outcome, intervention, patients, and study quality.

There were no statistically significant differences in the effect of SMT based on any of these variables, although there was a suggestion that SMT's benefit was greater for thrust (as compared to non-thrust) SMT, and in studies of better methodological quality. The 3 studies that reported the largest beneficial effects for SMT all selected patients based on specific criteria.

Other Outcomes

Too few studies included outcomes other than pain and function to allow us to draw conclusions. Four studies reported return-to-work or duration of sick leave (2 of which reported no differences between groups and one each reported shorter and longer sick leave for the SMT group), one study reported no differences in SF-12 outcomes, and 2 studies reported utilization data.

Acute Low Back Pain with Sciatica

We found 3 randomized controlled clinical trials using SMT in patients with back pain and sciatica. This was too few to draw conclusions.

Adverse Events

In the 26 RCTs of SMT for acute low back pain included in our pooled analyses, 18 publications made no mention of any assessment of adverse events, 3 publications made general comments about adverse events ("no adverse effects were documented…"), and 5 publications reported on specific adverse events, none of which were judged to be related to the treatment except for "the treatment hurts" being statistically more common in the group of patients receiving SMT (as part of a package of therapies) compared to those receiving conventional medical care.

SMT in General

We identified 8 studies that prospectively assessed adverse events in patients receiving SMT, generally by asking consecutive patients receiving SMT from a sample of manual therapy clinicians to complete a survey. The results of these studies, which ranged from 68 patients to

1,058 patients, are broadly consistent. Mild, transient adverse events are reported by 50%-60% of patients, with the most common reported events being local discomfort or an increase in pain.

Serious Adverse Events

There have been numerous case reports, collections of case reports, and systematic and non-systematic reviews of serious adverse events of SMT, of SMT for low back pain, and of SMT for neck pain. The limitations of not being able to assess causality and not being able to calculate frequency have not been overcome.

Summary of Results for Key Questions and Strength of Evidence

Key Question 1: What are the benefits and harms of spinal manipulation/chiropractic services for acute lower back pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Twenty-six studies of SMT treatments for acute low back pain found overall statistically significant evidence of a clinical benefit that was, on average, modest. However, there was substantial heterogeneity in results, with some studies reporting much larger effects and some studies reporting no effect at all. We explored 6 potential sources of heterogeneity, and although type of manipulation, patient selection, and study quality may explain some of the heterogeneity, most of the differences in outcome between studies remain unexplained.

We judged the quality of evidence as moderate that treatment with SMT improved the outcomes of pain and function in patients with acute low back pain, due to heterogeneity of results.

We judged the quality of evidence as high that transient minor musculoskeletal adverse events are common following SMT, although they may be equally common following non-SMT manual therapy.

We judged the quality of evidence as insufficient regarding SMT and outcomes for patients with low back pain and sciatica.

Key Question 1A: What is the relationship between the use of spinal manipulation/chiropractic services for lower back pain and the use of opiate medication?

Among the 26 studies included in our pooled analysis only one specifically reported on the use of opiate medications.

With only a single study reporting this outcome and that one not reporting the actual use by treatment group, we classified the quality of evidence as insufficient for this outcome.

A number of studies have reported on the association of chiropractic care and opioid use using claims data. While these studies have reported lower use of opioids in patients also or first receiving chiropractic care because of their observational design the studies are not able to control for selection bias and therefore were not considered as evidence for this report.

Key Question 2: What are the benefits and harms of spinal manipulation/chiropractic services for acute neck pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Only 5 studies were identified of SMT compared to a non-SMT treatment group for patients with acute neck pain. Although each study reported favorable results on at least one outcome, in total only 198 patients have been studied in total.

We rated the evidence as low that SMT improves outcomes in patients with acute neck pain due to study quality concerns and imprecision of results (too few studies).

Key Question 2A: What is the relationship between the use of spinal manipulation/chiropractic services for acute neck pain and the use of opiate medication?

None of the included studies reported on the use of analgesic medications or opiate medication as an outcome.

DISCUSSION

Limitations

In general, we did not find evidence of publication bias, although no evidence of bias is not the same as evidence of no publication bias.

Study Quality

Study quality was highly variable and in our pooled analysis is split about equally between studies considered "high" and studies considered "low" quality. Our analysis found no evidence to support a hypothesis that our results are due to low-quality studies with inflated effect sizes.

Heterogeneity

Heterogeneity in the results is the primary limitation of this analysis. The statistical evidence of heterogeneity was significant and visual inspection of the forest plots illustrated this: some studies of SMT found positive results, while others, for the same outcome, found essentially no benefit (ES = 0, ES = 0.06, *etc*). Our investigation of multiple potential sources of heterogeneity yielded no results that were statistically significant, although visually there were suggestions that the type of SMT may be important. Nevertheless, the majority of heterogeneity remains unexplained and the large degree of heterogeneity may limit the enthusiasm of some clinicians and policymakers for advocating more widespread use of SMT.

Applicability of Findings to the VA Population

We identified no studies specific to VA population. Nevertheless, acute back pain in primary care is probably quite similar within VA to outside VA, and these results have to be considered at least moderately applicable to VA populations.

Research Gaps/Future Research

There continues to be a great deal of unexplained heterogeneity in results of SMT for acute low back pain, so a research gap is better understanding what contributes to patient selection and intervention to improve the consistency of the result. This could include an attempt at replication of the clinical prediction rule RCT or new RCTs with more detailed data collection on the patient clinical characteristics and details of the SMT intervention. For neck pain, there are simply too few studies to draw firm conclusions. Additional RCTs are warranted. Attention should be paid to collecting clinical variables and details of the intervention to use in the exploration of possible heterogeneity of treatment effects.

EVIDENCE REPORT

INTRODUCTION

Back pain and neck pain are among the most common symptoms prompting patients to seek care. While data specific to Veterans are not available, in the general population lifetime prevalence estimates of low back pain are as high as 80% in the US population. Most persons can expect to have an episode of acute back pain, acute neck pain, or both at some point.

Many treatments are used for back pain without having been established as so clearly superior as to extinguish the use of others. Treatments include analgesics, muscle relaxants, bed rest, exercises, physical therapy modalities, heat, ice, spinal manipulation, acupuncture, and others.

Spinal manipulative therapy (SMT) is a treatment option available in VA, provided mostly but not entirely by Doctors of Chiropractic. In practice, most patients referred to VA chiropractors have chronic pain. In order to better understand the potential role of SMT in treating acute back or neck pain, VA requested an up-to-date synthesis of the evidence.

There have been several prior reviews on spinal manipulation, including one by a member of the ESP team that concluded SMT is superior to a sham, but not clearly superior to other effective treatments for acute low back pain.[1] The most recent Cochrane Review on the subject, however, concluded that SMT is not more effective than any other intervention or sham.[2] Thus, one goal of this review is to help resolve disagreements in the results from prior reviews.

METHODS

TOPIC DEVELOPMENT

The VA has had a significant increase in requests for chiropractic care since these services became covered by the VHA. With an increased focus on interdisciplinary care within the VHA, findings from an evidence synthesis about the effectiveness of spinal manipulative therapy (SMT) or chiropractic care will help the VA identify approaches for treating acute neck and lower back pain and ensure the VA is providing Veterans with optimal healthcare services.

This report was developed based on a nomination from operational partners Lucille Beck, PhD, Deputy Chief Patient Care Services Officer, Rehabilitation and Prosthetic Services (10P4R), Anthony Lisi, DC, Director, VHA Chiropractic Service Rehabilitation and Prosthetic Services; Section Chief, Chiropractic Service, VA Connecticut Healthcare System; and David Chandler, PhD, Deputy Chief Consultant, Rehabilitation and Prosthetic Services (10P4R).

The proposed Evidence-based Synthesis Program (ESP) evidence synthesis will be used by the Office of Rehabilitation and Prosthetic Services (10P4R), Chiropractic Service, to inform VA clinical practice and national policy as the VA continues to implement chiropractic services across the country.

The Key Questions are:

Key Question 1: What are the benefits and harms of spinal manipulation/chiropractic services for acute lower back pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Key Question 1A: What is the relationship between the use of spinal manipulation/chiropractic services for lower back pain and the use of opiate medication?

Key Question 2: What are the benefits and harms of spinal manipulation/chiropractic services for acute neck pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Key Question 2A: What is the relationship between the use of spinal manipulation/chiropractic services for acute neck pain and the use of opiate medication?

The PROSPERO registration number is CRD42015017916.

SEARCH STRATEGY

Spinal manipulation is a topic that has been the subject of numerous prior systematic reviews, including 3 reviews by members of the ESP review team. Therefore, instead of searching for original evidence in databases such as PubMed, we instead began by reference mining existing systematic reviews, and then performing an update search to identify new studies published since the end date of the searches of the most recent reviews. Then we consulted our technical experts for any additional studies we might have overlooked. See Appendix A for full search strategy.

STUDY SELECTION

All reference titles and abstracts were screened independently by 2 reviewers. If either reviewer selected a title or abstract, it was included for further review. Full-text articles were then reviewed in duplicate, with all discrepancies discussed with the group. References were selected based on the following inclusion criteria:

*Participant*s: Adults (ages 18 and older) with acute (defined as 6 weeks or less) neck or lower back pain. Patients with sciatica were included. Studies of patients with chronic back pain were excluded, as were studies where we could not determine the duration of pain. If studies included patients with longer durations of pain, we included them if they presented stratified results or if the majority of patients had pain for less than 6 weeks duration. Studies of children were excluded; this included pediatric populations or patients under the age of 18.

Intervention: Spinal manipulation by any provider type. Studies where spinal manipulation was given alone or as part of a "package" of therapies were included. "Chiropractic care" was considered as including SMT for the great majority of patients. The definitions of SMT types were refined during the data abstraction process, and a more detailed description of the intervention is given in the following Data Abstraction section.

Comparator (study design): Other forms of management for acute pain, such as analgesics, exercises, physical therapy, etcetera. Sham-controlled studies were included.

Outcome: Pain management, functional status, quality of life, opiate use, disability claims, return to work, health care utilization.

Timing: Studies had to report at least one outcome within 6 weeks to be eligible.

Setting: Ambulatory/outpatient settings. Studies in hospital settings were excluded.

Study design: Only randomized controlled trials (RCTs) were eligible for assessing benefits. Both RCTs plus observational studies were used for assessing harms.

DATA ABSTRACTION

Data were extracted by 2 reviewers, and discrepancies were reconciled after discussion. Articles had data abstracted on the anatomical location of the pain, authors' description of the SMT provided, type of professional performing the treatment, co-interventions, whether that treatment was provided alone or as part of a package of other treatments, whether patients were selected as more likely to respond to SMT or unselected, data on any of the outcomes listed above (*eg*, pain, functional status, *etc*), as well as data needed to complete the Cochrane Back Group Risk of Bias assessment. For studies included in our own prior reviews of SMTs we used data abstracted from those studies at that time.

Based on the authors' description of the SMT provided, we categorized the study as using thrust or non-thrust technique. Thrust was defined as high velocity low amplitude (HVLA), such as "a short-lever, high-velocity thrust directed specifically at a 'manipulable lesion.'"[3] Non-thrust was defined as low velocity high amplitude (LVHA), such as one study where "most participants had several low-velocity mobilization techniques."[4] Any studies where the research team had

questions were brought to the Technical Expert Panel (TEP) for their input. In one case, we contacted the original author to clarify this (John Triano, personal communication, October 28, 2015).

Statistical data were extracted by the project statistician. We focused on data from follow-ups less than 6 weeks. For continuous outcomes, the sample size, mean, and standard deviation were extracted for each SMT group and comparator group within each trial. For count data, odd-ratios (OR) were extracted if means were not reported. Since pain and function scales often differed across studies, a standardized mean difference (SMD) was calculated between the SMT group and each comparator group. A negative SMD indicates that the SMT group is doing better at follow-up than the comparator group. The few studies that reported ORs were converted into SMDs and combined with the continuous trials.

QUALITY ASSESSMENT

We assessed the quality of studies using the Cochrane Back Group Risk (CBG) of Bias Tool (ROB) (see Appendix B for full tool). This tool has 11 items in the following domains: randomization, concealment, baseline differences, blinding – patient, blinding – care provider, blinding – outcome, co-interventions, compliance, dropouts, timing, and intent to treat. Prior research has shown the CBG ROB Tool to identify studies at an increased risk of bias using a threshold of 5 or 6 as a summary score.[5]

DATA SYNTHESIS

We constructed evidence tables showing the study characteristics and results for all included studies.

Studies were pooled within outcome measures and 95% confidence intervals were constructed. Studies using a 100 mm Visual Analog Scale or 11-point Numeric Rating Scale or other numeric pain scale were pooled together by converting all outcomes to a 0-100 measure (using the appropriate multiplier); studies reporting the Roland Morris Disability Questionnaire (RMDQ, scored on a 0-24 scale) and studies reporting the Oswestry Disability Index (ODI, scored as 0-100) were pooled as a functional outcome using an effect size approach. Studies reporting none of these were not pooled, but discussed narratively.

Random effects meta-analyses were conducted using the Hartung-Knapp Method.[6,7] Tests of heterogeneity were performed using the I^2 statistic.[8] All meta-analyses were conducted with Stata statistical software, version 12.0[9] and R3.2.2. The Begg's rank correlation[10] and Egger regression asymmetry test[11] were used to examine publication bias. To further explore possible sources of heterogeneity, such as timing, outcome, type of practitioner, and type of manipulation, bivariate meta-regressions were conducted.

The meta-analyses were organized based on 2 follow-up times and the 2 outcomes. Two studies[12,13] were in the gap between immediate and short-term outcomes; they were closest to immediate-term so they were classified in the immediate-term group. Within these 4 groupings the intervention was assessed in comparison to control interventions classified as either sham SMT or all other therapies. [14] Studies comparing SMT to sham-SMT were not pooled with

studies comparing SMT to other therapies. Studies were included in each pooled analysis only once.

An *a priori* analysis considered 3 potential sources of heterogeneity: the comparison group, the outcome, and the timing of the outcome. In addition, 3 post-hoc hypotheses were developed to test possible explanations for observed heterogeneity: by type of manipulation, comparing thrust techniques to non-thrust techniques; by the types of patients enrolled (selected or not selected); and by study quality, comparing higher-quality trials to lower-quality trials.

The Minimum Clinically Important Difference (MCID) for either pain or function in acute low back pain is not well-established empirically. The MCID for the Roland Morris Disability Questionnaire has been proposed as low as 1.5 points and as high as 5 points. A recent systematic review of studies of the minimum clinically important difference for pain scales in acute pain concluded that no single value could be supported.[15] Therefore, we have not chosen a MCID value for pain or function, but frame our results as a range or in comparison to other treatments for acute low back pain.

RATING THE BODY OF EVIDENCE

The evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria, which uses the domains of study design limitations, inconsistency, indirectness, and imprecision in results.[16] The GRADE Working Group classified the quality of evidence across outcomes according to the following criteria:

High: We are very confident that the estimate of effect lies close to the true effect for this outcome.

Moderate: We are moderately confident that the estimate of effect lies close to the true effect for this outcome.

Low: We have limited confidence that the estimate of effect lies close to the true effect for this outcome.

Insufficient: We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome.

PEER REVIEW

A draft version of this report was reviewed by 3 technical experts and 2 members of VA operations. Reviewer comments were addressed and out responses were incorporated into the final report. The complete set of comments and responses can be found in Appendix C.

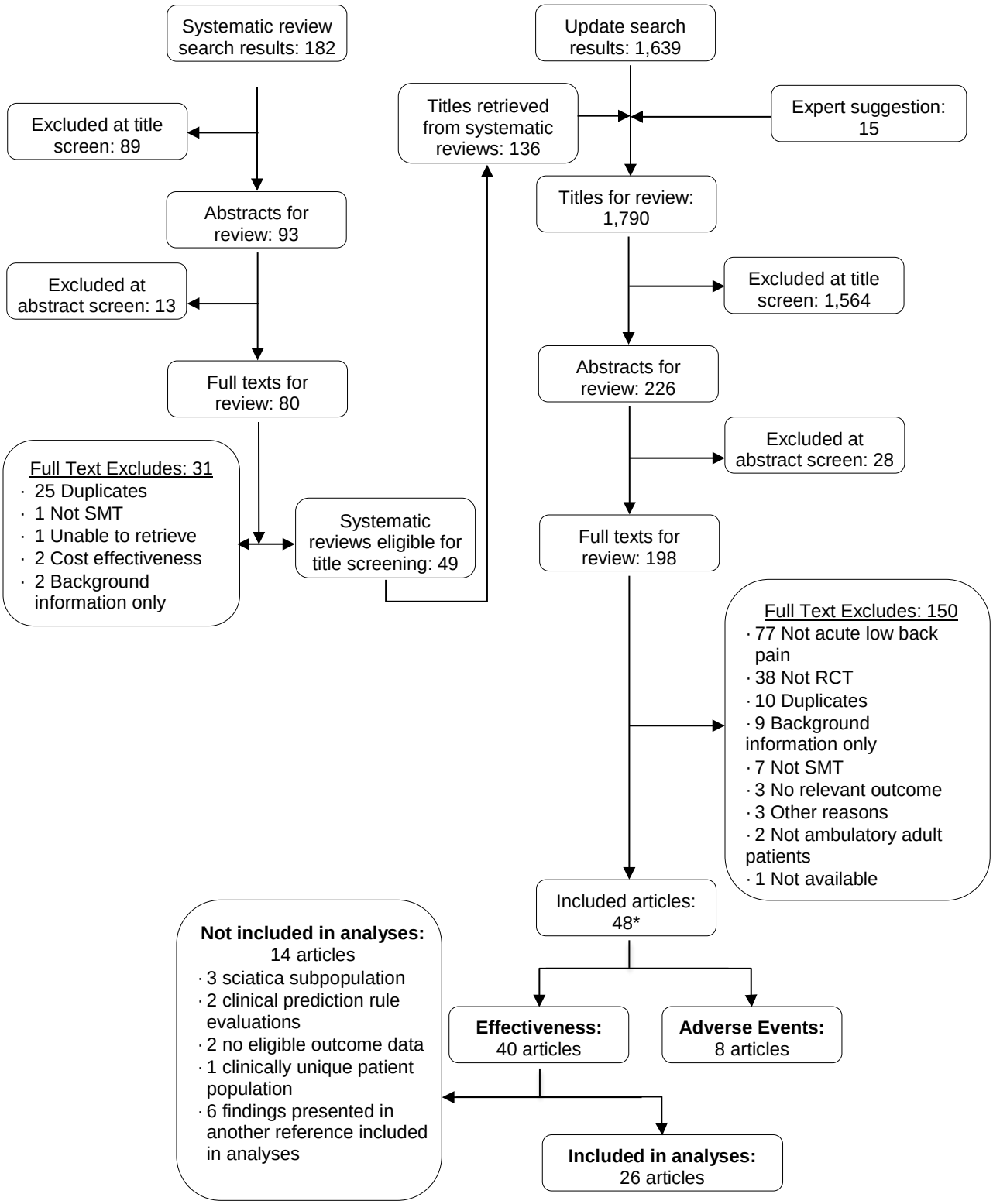
RESULTS

LITERATURE FLOW

We identified 181 potentially relevant titles from our systematic review search and identified one additional title from the references of one of our included articles for a total of 182 titles for screening. After reviewing titles, we excluded 89 titles, leaving 93 references for abstract screening. Of these 80 were included to be reviewed at the full-text level, excluding 13 abstracts. From the 80 full texts, we identified 49 systematic reviews that we mined for references, 8 references relevant to adverse events, and 31 excluded full texts. The excludes were comprised of 25 duplicates of already screened articles, 2 articles containing background information only, 2 articles focusing on cost effectiveness, 1 article not reporting on SMT, and 1 article we were unable to retrieve (see Figure 1 for literature flow details).

From the 49 systematic reviews we mined for references, we identified 136 potentially relevant titles. To this we added 15 titles recommended by experts and 1,639 titles identified in an update search for a total of 1,790 titles for screening. After excluding 1,564 titles as clearly not relevant, we reviewed 226 abstracts. Of these, we excluded 28 abstracts and included 198 abstracts for full-text review. After full-text review, we excluded 150 articles: 77 articles rejected as studying patients with pain longer than 6 weeks or unspecified; 38 articles rejected for study design (*ie*, not a randomized controlled trial); 10 articles rejected as duplicate articles of already-screened articles; 9 articles rejected as providing relevant background information but were not otherwise included; 7 articles rejected as not reporting on SMT; 3 articles rejected for having no relevant outcome; 2 articles rejected for studying patients in hospital; 3 articles rejected for other reasons; and 1 article that we were unable to retrieve.

Figure 1. Literature Flow Chart



Description of the Evidence

Of the 48 included articles, we identified 40 articles relevant to effectiveness of SMT and 8 articles relevant to adverse events. Of the 40 effectiveness articles, 26 were included in the analyses. Of the 14 not included in the analyses, 3 publications were focused on the subpopulation of patients with sciatica, 2 publications were only relevant to clinical prediction discussions, 2 publications did not have the necessary outcome data, and one publication had a unique patient population judged by our TEP as clinically dissimilar to the other studies. All of these studies are discussed in the narrative synthesis. The final 6 studies were excluded from pooled analyses because they presented duplicate data. Five articles by Blomberg and colleagues related to one randomized controlled trial,[17-21] and were grouped for analyses as one study, as were 3 articles by Grunnesjö and colleagues relating to another randomized controlled trial.[22-24]

Within the group of 26 studies included in the analyses, there were 13 studies where physical therapists provided the therapy, 7 studies where SMT was provided by chiropractors (DCs), 5 studies where SMT was provided by medical doctors (MDs), and 3 studies where SMT was provided by osteopaths (DOs). These were not mutually exclusive, as some studies employed multiple types of professionals.

Of the 26 studies, 17 studies utilized a thrust technique. 6 studies used a non-thrust technique, and 3 studies used a mix of both. If all patients received both thrust and non-thrust techniques, we classified it as thrust SMT. “Mixed” studies were ones where not all of patients, or most patients, clearly received thrust-type SMT.

Quality Assessment

In the low back pain analysis, one study scored a high of 9 out of 11 possible points, 6 studies scored 7 points, 4 studies scored 6 points, 2 studies scored 4 points, 7 studies scored 3, and 6 studies scored 2 points (Table 1).

Of the 26 studies, 25 studies met the timing criteria and 17 met the randomization criteria. None of the studies met the blinding of providers criteria, and only 4 met the criteria for blinding of patients using a threshold of 6. Twelve studies were classified as high quality and 14 studies were classified as low quality.

Table 1. Quality Scores of RCTs of SMT for Acute Low Back Pain

Article	Randomization	Concealment	Blinding, provider	Blinding, patient	Blinding, outcome	Dropouts	Timing	ITT	Baseline Differences	Co-interventions	Compliance	Total
Bergquist-Ullman 1977[25]	+	?	-	-	-	-	+	-	?	?	-	2
Blomberg[17-21]	+	+	-	-	-	+	-	+	-	+	+	6
Cherkin 1998[3]	+	+	-	-	-	+	+	+	+	-	?	6
Childs 2004[26]	+	+	-	-	-	+	+	+	+	?	+	7
Cramer 1993[27]	?	?	-	-	-	+	+	+	?	?	?	3
Cruser 2012[28]	?	+	-	-	-	+	+	+	+	+	+	7
Delitto 1993[29]	+	-	-	-	?	?	+	+	+	?	?	4
Erhard 1994[30]	+	?	-	-	?	-	+	-	+	?	?	3
Farrell 1982[31]	?	?	-	-	-	+	+	?	+	?	?	3
Fritz 2015[32]	+	+	-	-	-	+	+	+	+	-	+	7
Glover 1974[33]	+	?	-	-	-	?	+	?	?	?	?	2
Godfrey 1984[12]	+	?	-	+	+	-	+	-	?	+	+	6
Goertz 2013[34]	+	+	-	-	-	-	+	+	+	+	+	7
Grunnesjö[22-24]	+	+	-	-	-	+	+	+	+	?	+	7
Hadler 1987[35]	?	?	-	+	-	+	+	-	-	?	?	3
Hallegraeff 2009[13]	+	+	-	-	-	+	+	+	-	?	+	6
Hancock 2007[4]	+	+	-	+	-	+	+	+	+	+	+	9
Heymann 2013[36]	?	+	-	+	+	-	+	+	+	?	?	6
Hoiriis 2004[37]	+	?	-	-	-	-	+	+	?	?	?	3
Juni 2009[38]	+	+	-	-	-	+	+	+	?	+	+	7
MacDonald 1990[39]	?	?	-	-	-	+	+	-	+	?	+	4
Morton 1999[40]	+	-	-	-	-	?	+	+	?	?	?	3
Postacchini 1988[41]	?	?	-	-	-	+	+	-	?	-	?	2
Rasmussen 1979[42]	?	?	-	-	-	+	+	-	?	?	?	2
Skargren 1998[43]	?	?	-	-	-	-	+	-	+	-	?	2
Waterworth 1985[44]	+	?	-	-	-	+	+	-	?	?	?	3

* + = yes, - = no, ? = unsure/don't know; full criteria specified in Appendix B.

KEY QUESTION 1: What are the benefits and harms of spinal manipulation/chiropractic services for acute lower back pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Acute Low Back Pain without Sciatica

Twenty studies reported results that we could use for meta-analysis pooling. Figures 2 through 6 present the results, stratified by outcome: immediate-term (≤ 2 weeks) pain, immediate-term function, short-term (3-6 weeks) pain, and short-term function. Within each outcome, results are presented stratified by comparison group.

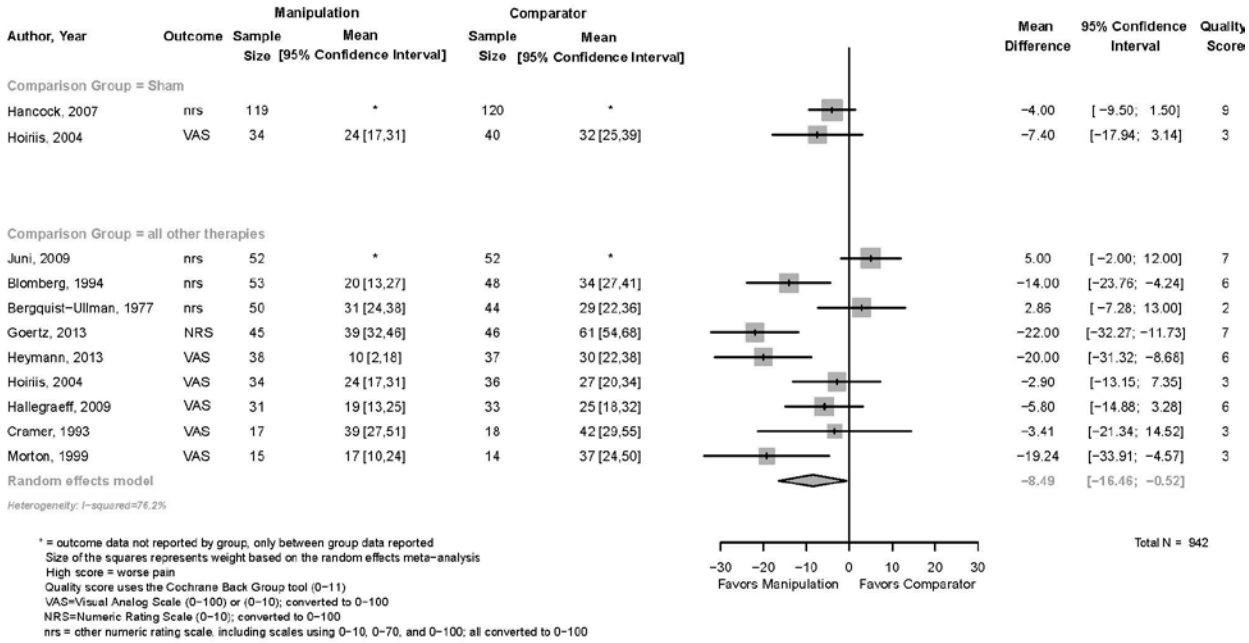
Immediate-term Pain (≤ 2 weeks)

Fourteen studies reported 17 comparisons of SMT to other treatments and reported short term outcomes. We initially compared the effects of SMT to relatively narrow categories of alternatives: physical therapy, diatherapy (usually de-tuned diatherapy), sham, analgesics, back school, bed rest, etcetera.

However, this resulted in many comparisons with only one or 2 studies, limiting our analytic power. Furthermore, visual and statistical inspection of the forest plots did not support a need for all these different categories, since with few exceptions there were not visual or statistically significant differences in effectiveness across all these different comparison groups. We then grouped comparison groups into the following, using physiology and the authors' intent (as to whether the comparison was thought to be an active treatment or a "placebo" treatment) as our guide: manual therapies intended to be active; manual therapies intended to be inactive; true sham SMT; conventional or usual medical care; analgesics/muscle relaxants as an isolated pharmaceutical intervention; and bed rest (since bed rest is a potentially harmful intervention). This still resulted in most categories having relatively few studies, limiting power. Since visual and statistical analysis of the forest plots did not support any statistically significant differences between effectiveness and comparison group category, we therefore elected to pool across all comparison groups (except sham-controlled studies, which were kept separate). This classification was justified because many of the comparison interventions were intended to be inactive (detuned diathermy, light massage, etc.) or of uncertain effectiveness ("usual medical care"); and for those comparisons where the other treatment was expected to be effective the existing RCTs and systematic reviews indicate the benefit is small, at best.[45-47]

Figure 2 presents the results for immediate-term pain. There were 11 studies reporting immediate-term pain outcomes using a VAS or numeric rating scale, 2 comparing SMT to sham, and 9 comparing SMT to another therapy (Figure 3). The overall random effects pooled estimate was -8.49 mm (95% CI: -16.46, -0.52) favoring treatment with SMT. There was heterogeneity, with an $I^2 = 76.1\%$. There was no evidence of publication bias in the overall pooled result, with Begg's rank correlation = 0.15 and Egger's test p-value = 0.58. Two studies comparing SMT to sham reported non-statistically significant benefits.

Figure 2. Effect of SMT on Immediate term Pain

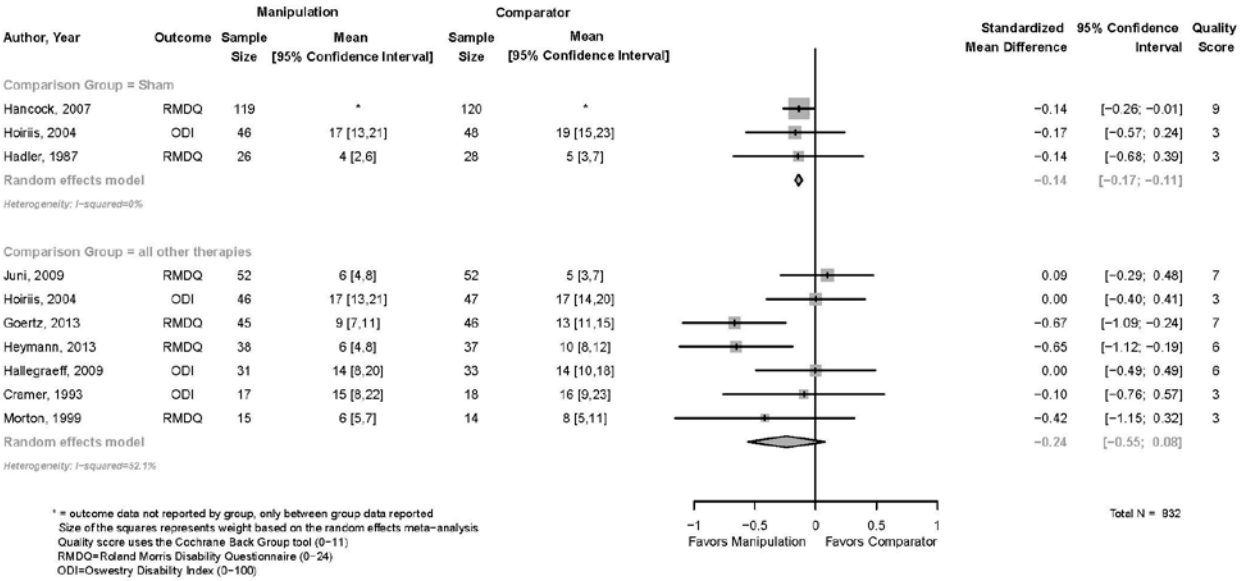


* = outcome data not reported by group, only between group data reported
Size of the squares represents weight based on the random effects meta-analysis
High score = worse pain
Quality score uses the Cochrane Back Group tool (0-11)
VAS=Visual Analog Scale (0-100) or (0-10); converted to 0-100
NRS=Numeric Rating Scale (0-10); converted to 0-100
nrs = other numeric rating scale, including scales using 0-10, 0-70, and 0-100; all converted to 0-100

Immediate-term Function (≤ 2 weeks)

There were 10 studies reporting immediate-term function measured with the RMDQ or ODI, 3 comparing SMT to sham, and 7 comparing SMT to another therapy (Figure 4). The overall random effects pooled estimate was an effect size of -0.24 (95% CI: -0.55, 0.08) favoring treatment with SMT. There was heterogeneity, with an $I^2 = 52.1\%$. There was no evidence of publication bias in the overall pooled result, with Begg's rank correlation = 0.17 and Egger's test p-value = 0.14. Three studies compared SMT to sham and the overall random effects pooled estimate was an effect size of -0.14 (95% CI: -0.26, -0.11).

Figure 3. Effect of SMT on Immediate term Function



* = outcome data not reported by group, only between group data reported
Size of the squares represents weight based on the random effects meta−analysis
Quality score uses the Cochrane Back Group tool (0−11)
RMDQ=Roland Morris Disability Questionnaire (0−24)
ODI=Oswestry Disability Index (0−100)

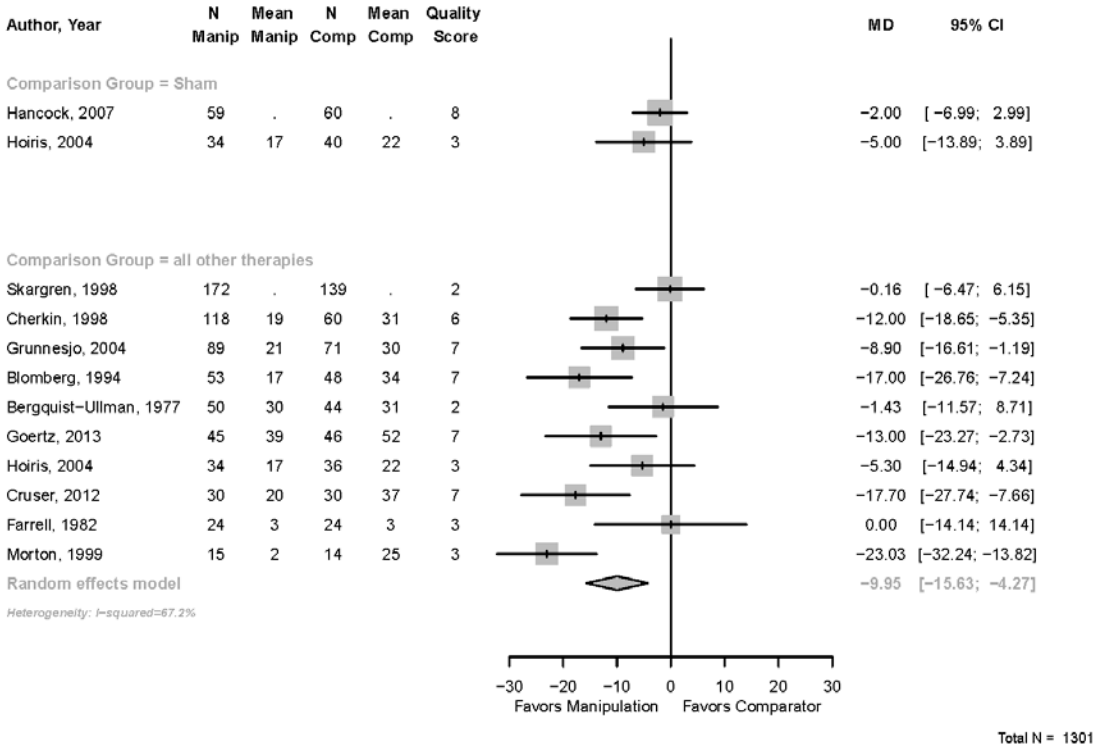
Short-term Pain (3-6 weeks)

There were 12 studies reporting short-term pain using VAS or numeric rating scale, 2 comparing SMT to sham, and 10 comparing SMT to another therapy (Figure 4). The overall random effects pooled estimate across all studies was an effect size of -9.95 mm (95% CI: -15.6, -4.3) favoring treatments with SMT. There was heterogeneity, with an $I^2 = 67.2\%$. There was no evidence of publication bias in the overall pooled result, with Begg's rank correlation of 0.92 and Egger's test p-value of 0.58.

Studies Not Included in the Pooled Analysis for Pain

Five studies reported pain outcomes that were not measured with a 100 mm VAS or numeric pain scale.[12,33,41,42,44] All were old studies (30-40 years ago), and all but one were judged as low quality. Two of the 5 studies concluded SMT had an effect [41,42] and 3 studies concluded it did not.[12,33,44]

Figure 4. Effect of SMT on Short term Pain



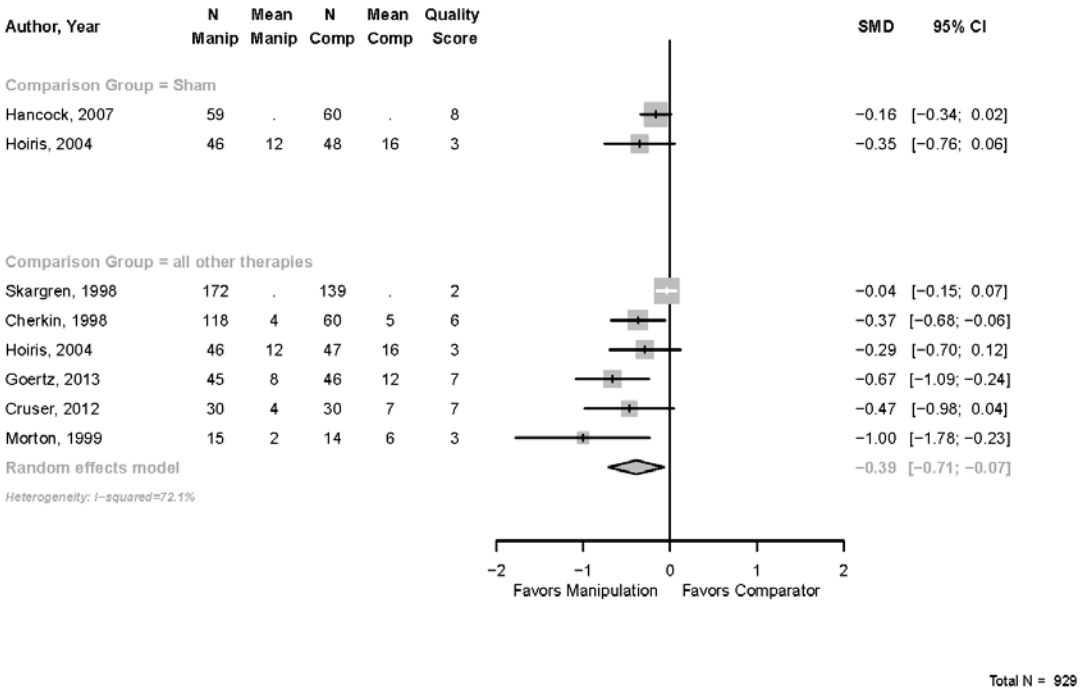
Short-term Function (3-6 weeks)

There were 8 studies reporting short-term function outcomes measured with the RMDQ or ODI, 2 comparing SMT to sham, and 6 comparing SMT to another therapy (Figure 5). The overall random effects pooled estimate was an effect size of -0.39 (95% CI: -0.71, -0.07). There was heterogeneity, with an $I^2 = 72.1\%$. There was no evidence of publication bias, with Begg's rank correlation = 0.85 and Egger's test p-value = 0.10. Two studies comparing SMT to sham reported non-statistically significant benefits.

Studies Not Included in the Pooled Analysis for Function

Five studies did not report function outcomes using the RMDQ or ODI.[12,17,23,39,44] With one exception, all the studies were performed more than 20 years ago. Three studies were judged as high quality and 2 studies were low quality. Three studies concluded SMT had an effect compared to usual medical care, advice to stay active, or advice on posture, exercises and avoidance of occupational stress [17,23,39] and 2 studies concluded it did not.[12,44]

Figure 5. Effect of SMT on Short term Function



Exploring Sources of Heterogeneity

As noted above, there was significant heterogeneity in almost all the pooled analyses of SMT, suggesting that there are other factors influencing the outcome. In addition to the comparison group, we investigated 5 possible sources of heterogeneity:

The outcome. Pain and function are the 2 most commonly reported outcomes, and it is possible that SMT affects these 2 outcomes differently.

The timing of the outcome. The natural history of acute low back pain is that the great majority of patients recover within a few weeks. Therefore, if measured at 4-6 weeks from enrollment there may be no differences between groups because even most untreated patients will have recovered. However, this could miss differences in the pace of recovery. SMT may speed the pace of recovery, in which case measurement of outcomes at shorter time points – 1 week or 2 weeks – could detect differences between treatments that may be lessened or absent at 4-6 weeks.

The intervention. "Spinal manipulative therapy" is a term that encompasses a large variation in the type of manual therapy. Anecdotally almost all manual therapists believe that different kinds of manipulation have differential effectiveness, particularly when matched to certain patient clinical characteristics. However, direct evidence that this is the case has been lacking. There is experimental evidence that, at least among patients meeting a clinical prediction rule for SMT, thrust-type manipulation is more effective than non-thrust-type manipulation.[48] Clinicians on our technical expert panel all agreed that they believed thrust-type SMT was more effective, in general, than non-thrust SMT. Therefore we classified each study's intervention as either thrust-type SMT or non-thrust SMT. Since we hypothesized that thrust-type SMT is the more active of the two, studies in which all patients received both thrust and non-thrust SMT were classified as thrust SMT. Studies where therapists could choose from a range of SMT, some of which were thrust and some of which were non-thrust, and for which no additional data were presented to indicate the relative frequency of these actually delivered, were classified as "mixed" and not included in the analysis. This category included 3 studies.[20,22,28] In general, studies had to use the specific word "thrust" when describing their manipulation, or use the descriptor "high velocity low amplitude," to be classified as thrust manipulation. Studies with unclear descriptions of their interventions were presented to our technical experts for their interpretation.

The patients. Analogous to the discussion above about difference in types of SMT, almost all manual therapists believe that patient selection is critical to the application of SMT. Other than the set of studies dealing with the clinical prediction rule, though, evidence that this is true has been lacking. We therefore examined each study to see if the authors reported having selected patients based on certain a priori criteria they believed made patients more likely to benefit from SMT.

Study quality. Prior research has shown that in treatment of patient with back pain studies with lower quality, as determined by a summary score of 5 or 6 on the Cochrane Back Group quality checklist, had larger effect sizes than studies of higher quality. We therefore classified studies as higher or lower quality, based on a threshold of 6 points, and compared the results between the 2 categories.

Outcome and Timing

There were 15 studies that measured both pain and function outcomes, and 8 studies that measured outcomes at both early and later time points. This allowed us to do within-study comparisons of these 2 hypotheses about sources of heterogeneity. Within-study comparisons are, in general, less prone to bias than across-study comparisons, since all the "study-level" differences such as particulars of the treatment, patients, etcetera, are controlled for. In the 15 studies reporting both pain and function outcomes, the average effect size for the pain outcomes was -0.49, while the average effect size for the function outcome was -0.44. The difference between these was -0.05 (95% CI, -0.22, 0.12) meaning that the effect of SMT in pain outcomes tended to be slightly larger, but the difference was not statistically significant. In the 6 studies that presented both short-term and long-term outcomes the difference was 0.03 (95% CI, -0.23, 0.18), meaning that long-term outcomes were slightly larger, but the difference was not statistically significant.

Thrust versus Non-thrust SMT

We kept the comparisons of thrust versus non-thrust SMT separate for the different outcomes. Figure 6 presents the results of these comparisons, using an effect size analysis for both pain and function to have all studies on the same scale. Only for the outcome/timing of pain at 2 weeks or less were there sufficient studies in both categories to support a pooled result. These were not statistically significantly different, although the pooled effect size for thrust SMT studies was nearly twice as great as it was for non-thrust SMT studies (-0.44 versus -0.23). This pattern was not repeated for the outcome of function at 2 weeks or less, where the pooled effect size for thrust SMT studies (-0.18) was between the effect size for either of the 2 non-thrust SMT studies (-0.19 and -0.42; studies could not be pooled since 3 studies is the minimum needed for random effects pooling). The pattern was repeated for the outcome of pain at 3-6 weeks, where the pooled effect size for the thrust SMT studies (-0.30) was much larger than the effect size for the one study of non-thrust SMT (ES = -0.02). For the outcome of function at 3-6 weeks, the pooled effect size from 5 studies of thrust SMT (-0.30) was again about twice as big as the only estimate of effect from a non-thrust SMT study (-0.16). In 3 of the 4 outcome/time analyses the results support (but do not prove) that thrust SMT may be more effective than non-thrust SMT for patients with acute low back pain.

The Patients

As we did for the interventions, we kept the comparison of the effect that patient selection may have on outcome separate for the different outcomes and times. Unfortunately, only one study explicitly reported having selected patients based on an increased probability of response to SMT (outside of the clinical prediction rule studies, see below), and no conclusions could be drawn.

Study Quality

Figure 7 presents the results of the studies stratified by quality. There were no statistically significant differences between groups, but in general studies of higher quality reported larger effect sizes than studies of lower quality. From this, we conclude that the overall result of a beneficial effect of SMT in patients with low back pain is not due to lower-quality studies reporting more beneficial effects.

Figure 6. Effect of Thrust Compared to Non-thrust SMT, by Outcome

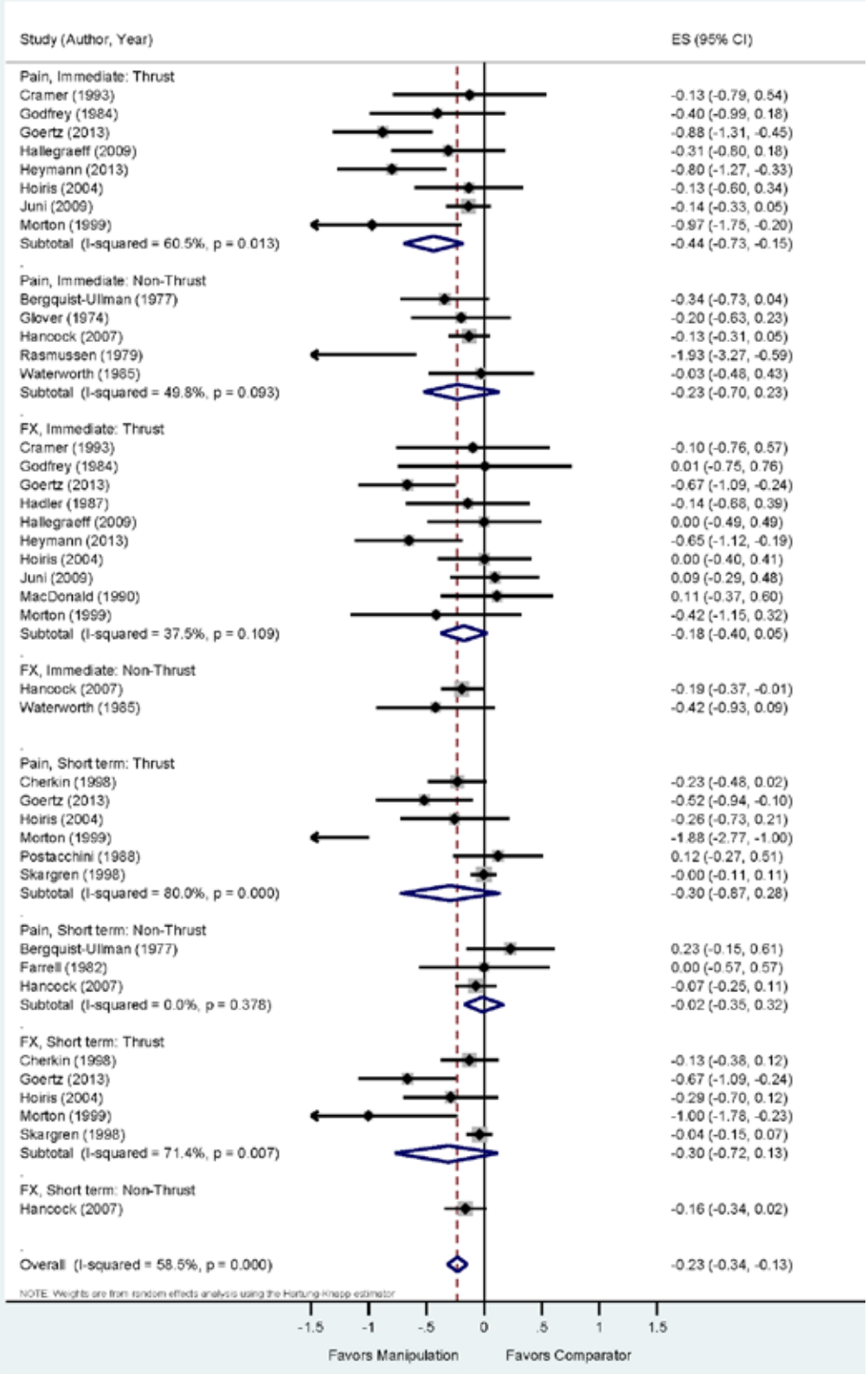
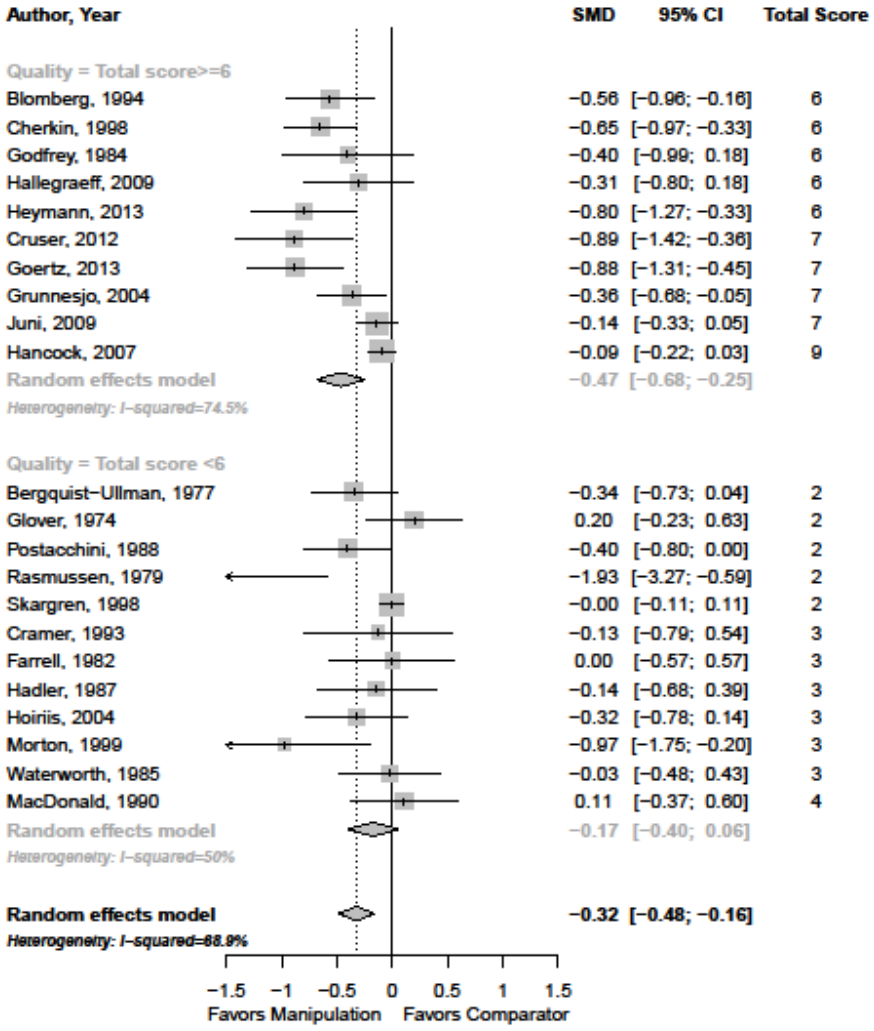


Figure 7. Quality Scores



Studies of a Clinical Prediction Rule for SMT for Low Back Pain

There were 4 studies meeting all eligibility criteria that we did not include in the pooled analysis because they all shared some common characteristics, which were: 1) All used a method to select patients as more likely to benefit from a specific kind of manual therapy; 2) All used the same SMT technique; 3) All studies were authored by professionally related physical therapists; 4) All studies used the same outcome measures, the Oswestry Disability Questionnaire; and 5) These studies reported the 3 largest effect sizes for their primary outcome, short-term function (effect sizes 3x-9x greater than the average for other SMT studies). Thus these "extraordinarily effective" studies are most appropriate to discuss as their own group.

The first 2 studies were authored by the same group of researchers at the University of Pittsburgh and Missouri. Both studies were relatively small (N = 24 in each). Patients were selected from a larger pool of eligible patients if they were judged to be more likely to respond to extension-mobilization therapy. This judgement was made on the basis of specific variables related to movement and physical signs that focused on pelvic alignment. In the second study, Waddell's tests for non-organic physical signs were used as an additional exclusion criterion.

In the first study, patients were then randomized to receive McKenzie-type extension exercises plus a manipulation "purported to affect the sacroiliac joint." The comparison group received Williams-type flexion exercises. In the second study, patients were randomized to receive the manipulation plus hand-heel rocking and the comparison group received McKenzie-type extension exercises. Both studies used the Oswestry Questionnaire as the outcome, at various times up to one week following the treatment. Both studies reported large benefits in favor of the patients receiving the manipulation therapy (effect size of -1.49 and -1.63). Both studies concluded that selecting patients according to the classification schemes and then treating with manipulation was effective.[29,30] Both studies were categorized as low quality on the Cochrane Back Group checklist.

The third study was a randomized trial of a clinical prediction rule to identify patients most likely to benefit from spinal manipulation. The patients and many of the authors were active members of the US Air Force. Building on their earlier work, which used a prospective cohort to identify variables,[48] the authors proposed 5 criteria, any 4 of which identified a patient as much more likely to benefit from spinal manipulation: duration of episode < 16 days, no symptoms radiating below the knee, less than 19 points on the Fear Avoidance Belief Questionnaire work subscale, and 2 physical findings, one for hypermobile lumbar spine segments and the other related to hip internal range of motion. Among a large number of potentially eligible patients with back pain, 70 were randomized to manipulation and 61 were randomized to low-stress aerobic and lumbar spine strengthening exercises. Patients were further divided within this group according to whether or not they met the clinical prediction rule for benefit from SMT.

The outcome measured was the Oswestry Disability Questionnaire. At 1 week, 4 weeks, and 6 months, the patients who were treated with SMT and who were positive on the clinical prediction rule had far better outcomes than patients who were negative on the rule (and treated with manipulation) or either exercise group. Comparing patients who were positive on the clinical prediction rule, patients treated with SMT had an effect size at 1 week that was -4.76.[26] This study was classified as high-quality on the Cochrane Back Group Checklist.

A fourth RCT reported results from participants selected using a similar clinical prediction rule and treated with the same type of thrust manipulation. Although this study found statistically significant benefits in both pain and function in patients treated with SMT, the size of the benefit was smaller than in the prior 3 studies.[32] This discrepancy was attributable to better outcomes in the non-SMT treated patients in this study compared to the prior 3 studies.

These results need to be placed in the context of the larger literature on clinical prediction rules in patients with low back pain. An attempt at an independent evaluation of this clinical prediction was done by a retrospective assessment of the data from an RCT included in our analysis.[4] The authors compared outcomes in their placebo-controlled trial of spinal manipulative therapy versus diclofenac, stratified by whether or not they met the threshold of 4 or more positive findings on the clinical prediction rule. There was no difference between SMT and placebo either for patients positive on the clinical prediction rule or for patients negative on the rule. These authors concluded the clinical prediction rule did not generalize to patients in primary care with acute low back pain.[49]

However, the SMT used in this "failed" evaluation was non-thrust manipulation. An RCT performed by the authors of the clinical prediction rule randomized patients to receive one of 3 therapies: 2 different kinds of thrust-type SMT or non-thrust SMT.[50] In this trial of 112 patients, patients randomized to either of the 2 thrust-type SMT treatments had better outcomes than patients who received non-thrust SMT, implying that an alternative hypothesis for the "failed" independent evaluation of the clinical prediction rule was that the type of SMT used in that study was ineffective in patients positive on the clinical prediction rule.

Two other RCTs have assessed this clinical prediction rule, one finding no difference in outcomes when thrust-type SMT was compared to "pragmatic non-thrust" SMT,[51] and the other finding no difference between outcomes when thrust-type SMT was compared to "mechanical diagnosis and therapy," which was postulated as a more effective intervention than the exercises used as the comparison in the original clinical prediction rule.[52] However, in the former study, the enrolled patient population had a much longer duration of pain (mean duration = 26 weeks) than in the clinical prediction rule validation study (median duration of symptoms = 4 weeks), raising questions about the comparability of the patient populations. The latter study is not a replication because the authors explicitly posit that the "mechanical diagnosis and therapy" is more effective than prior non-SMT-based physical therapy. These authors speculate that the clinical prediction rule simply identifies patients who are more likely to have a very favorable prognosis regardless of therapy. However, this hypothesis does not explain why patients in the original validation study who were positive on the clinical prediction rule had much better outcomes when treated with SMT than with non-SMT physical therapy.

In summary, there is RCT evidence that a clinical prediction rule helps identify patients more likely to respond well to thrust-type SMT, and "failed" independent evaluations have had substantial differences from the original study in terms of patients, the intervention, or the comparison group, limiting their conclusion that the clinical prediction rule is not valid. Nevertheless, it is a serious limitation that the clinical prediction rule results have not yet been replicated by an independent research team. A recent test of the clinical prediction rule did not report effects as large as the first 3 studies. Possible hypotheses include that the comparison group (usual care along with education and reassurance based on The Back Book) was more effective than the exercises given to the comparison groups in the prior studies; or patient

selection, as the most recent study recruited patients directly from primary care and not from patients already referred to physical therapy (and therefore possibly having less successful spontaneous improvement). The recent study also selected patients using a modification of the prediction rule that is more pragmatic for clinical implementation but is known to sacrifice specificity in identifying likely SMT responders.

Other Outcomes

Too few studies included outcomes other than pain and function to allow us to draw conclusions. Four studies reported return-to-work or duration of sick leave (2 of which reported no differences between groups[25,43] and one each reported shorter and longer sick leave for the SMT group[24,44]), one study reported no differences in SF-12 outcomes,[36] and 2 studies reported utilization data.[3,43]

Acute Low Back Pain with Sciatica

We found 3 randomized controlled clinical trials using SMT in patients with back pain and sciatica. Two of these, authored by Mathews et al, presented results from the same pool of patients.[53,54] Patients were divided into treatment arms based on their clinical presentation of duration of symptoms. In the manipulation versus heat arm, manipulation was given for up to 2 weeks, daily if indicated, at the discretion of the physiotherapist. Overpressure, rotation, and straight thrust techniques were used based on clinical symptoms. Infrared heat was applied to the low back of control patients for 15 minutes, 3 times weekly. Patients with low back pain with limited straight-leg raising (SLR) showed significant improvement of pain with manipulation when compared to heat ($p<0.01$). Patients with low back pain without limitation in SLR did not benefit from manipulation ($p>0.1$). Patients with low back pain and positive SLR also had a statistically significant decrease in pain when compared to heat ($p<0.01$).

In the other study by Santilli, patients with acute back pain and sciatica with disc protrusion were randomized to active spinal manipulation or simulated (sham) spinal manipulation.[55] Acute low back pain was defined as pain for less than 10 days in a patient who had been pain-free in the previous 3 months. To be included in the study, patients had to complain of moderate to severe intensity pain, moderate to severe radiating pain to one leg, and MRI evidence of disc protrusion with or without disc degeneration in the spinal segments involved in the pain. Patients were then randomized blindly to active or simulated manipulations. Patients received SMT by experienced chiropractors 5 days per weekfor up to 20 treatments and were followed at regular intervals for 180 days post treatment. Active manipulations consisted of examining the range of motion of the back, followed by soft tissue manipulations and brisk rotational thrusting away from the greatest restriction. Patients undergoing active SMT had a higher percentage of pain-free cases, (local pain, ($p<0.05$), radiating pain, ($p<0.0001$)), fewer days of pain ($p<0.005$), and fewer days of severe pain ($p<0.05$) compared to patients undergoing simulated treatments.

Adverse Events

Low Back Pain

In the 26 RCTs of SMT for acute low back pain included in our pooled analyses, 18 publications made no mention of any assessment of adverse events, 3 publications made general comments about adverse events (“no adverse effects were documented…”), and 5 publications reported on specific adverse events (Table 2), none of which were judged to be related to the treatment except for “the treatment hurts” being statistically more common in the group of patients receiving SMT (as part of a package of therapies) compared to those receiving conventional medical care.[20]

Table 2. Adverse events reported in randomized clinical trials of effectiveness of spinal manipulative therapy for acute low back pain

Author/Year	Sample Size	Method for assessing adverse events	Adverse Events
Blomberg, 1993[21]	N=149	Closed end questionnaires about side effects given to patients at 1, 2, and 4 months	Has a table of side effects by group. "The treatment hurts" was statistically significantly more likely in the group treated with spinal manipulative therapy than continued medical care.
Fritz, 2015[32]	N=220	Open and closed end questionnaire about side effects given to patients at 4 weeks	"12.0% (of patients) reported a total of 20 adverse effects from treatment including increased pain (1 mild, 2 severe, and no severity given), stiffness (2 mild, 3 moderate, 1 severe, and 1 no severity given), spasm (1 severe and 1 no severity given), shooting pain (1 moderate and 1 no severity given), and fatigue (1 mild)."
Goertz, 2013[34]	N=91	Adverse event data collection method not specified	"There were no serious adverse events." [2 mild adverse events were reported in spinal manipulative therapy group, both were pain that resolved in 24-48 hours]
Hancock, 2007[4]	N=240	Spontaneous reporting and open-ended questions	"No participants reported serious adverse reactions associated with spinal manipulative therapy."
Heymann, 2013[36]	N=100	Adverse event data collection method not specified	"Safety analysis did not show any unexpected untoward events in either group."
Juni, 2009[38]	N=104	Adverse event data collection method not specified	"Two serious adverse events occurred in the experimental group (4%) and two in the control group (4%). In the experimental group there was one patient with an acute loss of motor and sensory function due to a herniated disk after randomization, but before any spinal manipulative therapy treatment was initiated. In the control group, there was one patient with symptomatic cholelithiasis and one patient with a femoroacetabular impingement syndrome."

Author/Year	Sample Size	Method for assessing adverse events	Adverse Events
Morton, 1999[40]	N=29	Adverse event data collection method not specified	"No adverse effects were documented for either group."
Waterworth, 1985[44]	N=108	Adverse event data collection method not specified	"Adverse experiences with therapy were not specifically itemized, but their seriousness and drug relationship were recorded. Group 3 [spinal manipulative therapy] patients experienced less adverse reactions to treatments on the second assessment then group 1."

Text in quotations indicates text taken directly from the original article

SMT in General

We identified 8 studies that prospectively assessed adverse events in patients receiving SMT, generally by asking consecutive patients receiving SMT from a sample of manual therapy clinicians to complete a survey. The results of these studies, which ranged from 68 patients to 1,058 patients, are broadly consistent. Mild, transient adverse events were reported by 50%-60% of patients, with the most common reported events being local discomfort or an increase in pain (Table 3). Interestingly, in one randomized trial focused on SMT adverse events, while approximately 50% of patients receiving SMT reported adverse events, this was not statistically different than the reporting of adverse events in patients randomized to receive manual therapy without SMT or manual therapy without stretching exercises.[56] This suggests that these mild transient adverse events may be related to manual therapy in general and not spinal manipulation specifically. No serious adverse events were reported in any of these studies.

Table 3. Results from cohort studies and randomized clinical trials focused on adverse events of spinal manipulative therapy

Article/Study Design	Sample Size	Method for assessing adverse events	Interventions	Findings
Barrett, AJ., Breen, AC., 2000[57] Prospective cohort	68 patients 11 chiropractors	Collected from questionnaires to be given to 12 consecutive new patients	All received spinal manipulation	53% reported an adverse event, mostly increased or radiating pain.
Cagnie, B., 2004[58] Prospective cohort	465 patients 51 manipulating clinicians	Collected by questionnaires to be given to 15 consecutive new patients	All received spinal manipulation	283 patients (61%) reported at least 1 reaction. Headache, stiffness, aggravation of complaints, and radiating discomfort accounted for 2/3 of reactions.
Leboeuf-Yde, C., 1997[59] Prospective cohort	625 patients 66 chiropractors	Collected from questionnaires to be given to 10 consecutive patients	All received spinal manipulation	Treatment reactions were common, but benign and short lasting
Rubinstein, S. 2008[60] Prospective cohort	529 Patients with neck pain 79 chiropractors	Collected from questionnaires completed by patients at regularly scheduled visits	All received spinal manipulation	All patients were treated for neck pain. 56% of patients reported at least one adverse event. More than 70% of reported adverse events were musculoskeletal or pain.
Senstad, O., 1997[61] Prospective cohort	1050 patients 102 chiropractors	Collected by the chiropractor asking 12 consecutive patients a set of standardized questions	All received spinal manipulation	At least one reaction was reported by 580 patients (55%), 53% reported reactions were local discomfort.
Maiers, M., 2014[62] Randomized clinical trial	194 elderly patients with neck pain	Collected by standardized solicitation by clinicians, unsolicited reporting of patients and qualitative interviews with patients	Patients were randomized to receive spinal manipulative therapy, home exercise, or supervised rehabilitation exercise.	Overall, 130 patients (67%) reported at least one adverse event. Spinal Manipulative Therapy patients reported about twice as many adverse events as patients randomized to home exercise (74 vs. 40).

Article/Study Design	Sample Size	Method for assessing adverse events	Interventions	Findings
Paanalahti, K., 2014[56] Randomized clinical trial	767 patients	Collected by questionnaires given to patients in the waiting room at each return visit	Patients were randomized to spinal manipulative therapy, manual therapy without spinal manipulation, and manual therapy without stretching	About 50% of patients reported an adverse event. The most common adverse event was soreness in muscles, followed by increased pain, stiffness, and tiredness. There were no differences between patients receiving spinal manipulative therapy, manual therapy without spinal manipulative therapy, or manual therapy without stretching.
Walker, B.F., 2013[63] Randomized clinical trial	198 patients 12 chiropractors	Collected by questionnaires completed within 48 hours of treatment	Patients were randomized to usual chiropractic care (96% received spinal manipulative therapy) or a sham.	42% of usual care and 33% of sham care patients reported an adverse event. The most common adverse events were increased pain, muscle stiffness, headache and radiating discomfort.

Serious Adverse Events

There have been numerous case reports, collections of case reports, and systematic and non-systematic reviews of serious adverse events of SMT, of SMT for low back pain, and of SMT for neck pain.[64-76] All of these are based on case reports, or claims data, or both. These have been the subject of prior reviews and were not re-reviewed for this key question. The limitations of not being able to assess causality and not being able to calculate frequency have not been overcome.

Summary of Findings

Twenty-six studies of SMT treatments for acute low back pain found overall statistically significant evidence of a clinical benefit that was, on average, modest. However, there was substantial heterogeneity in results, with some studies reporting much larger effects and some studies reporting no effect at all. We explored 6 potential sources of heterogeneity, and while there were some non-statistically significant differences that may be signals of possible effects of type of manipulation, selection of patients, and study quality, most of the differences in outcome between studies remains unexplained.

Mild transient musculoskeletal adverse events are common following SMT, although these may be equally common following non-SMT manual therapy. Serious adverse events have been the subject of case reports, but assessing causality has proved challenging.

There were too few studies of SMT in patients with acute back pain and sciatica to draw conclusions.

Quality of Evidence for Key Question 1

We judged the quality of evidence as moderate that treatment with SMT improved the outcomes of pain and function in patients with acute low back pain, due to heterogeneity of results.

We judged the quality of evidence as high that transient minor musculoskeletal adverse events are common following SMT, although they may be equally common following non-SMT manual therapy.

We judged the quality of evidence as insufficient regarding SMT and outcomes for patients with low back pain and sciatica.

KEY QUESTION 1A: What is the relationship between the use of spinal manipulation/chiropractic services for lower back pain and the use of opiate medication?

Among the 26 studies included in our pooled analysis only one specifically reported on the use of opiate medications. In that study, about 9% of patients were prescribed opiate medications during the follow-up period, and the authors state “regimens were similar in the experimental and control groups.”[38] A second study reported use of “schedule II” medications that included cyclobenzaprine and acetaminophen with codeine. The authors reported no difference between groups in the use of schedule II drugs.[28] A third study reported “drug consumption” as an outcome, but this was not further specified.[20] One study reported the proportion of patients

taking opiate at baseline, but this was not measured as an outcome.[34] The remaining studies did not report drug consumption unless the drug was the comparison group (*eg*, a specific NSAID).

A number of studies have reported on the association of chiropractic care and opioid use using claims data. While these studies have reported lower use of opioids in patients also or first receiving chiropractic care, because of their observational design the studies are not able to control for selection bias and therefore were not considered as evidence for this report.[77-80]

Quality of Evidence for Key Question 1A

With only a single study reporting this outcome and that one not reporting the actual use by treatment group, we classified the quality of evidence as insufficient for this outcome.

KEY QUESTION 2: What are the benefits and harms of spinal manipulation/chiropractic services for acute neck pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

We found 5 randomized controlled clinical trials using spinal manipulative therapy (SMT) in patients with cervical (neck) pain.[81-85]

Howe et al studied 52 patients presenting with neck pain to a two-physician practice in Gwent, UK. Patients were randomized to thrust manipulation and azapropazone or azapropazone treatment alone. All patients except 6 had pain for less than 4 weeks. The 6 patients who had pain longer than 4 weeks were, by chance, allocated to the manipulation arm. Two patients in the manipulation arm were given an injection of a mixture of lignocaine and hydrocortisone prior to their manipulation to allow them to tolerate the manipulation better. Each group had goniometric assessments of rotation and lateral flexion on the day of randomization. Patients were asked to return for 2 follow-up visits at one and 3 weeks after the initial randomization and treatment. One of 17 patients (6%) described immediate improvement of neck pain in the control group; whereas 13/19 (68%) described immediate improvement in the manipulation group ($p<0.001$). A higher proportion of patients in the manipulation group continued to show improvement at the one- and 3-week visits, but the improvement over the control group was not statistically significant.

Nordemar and Thorner treated thirty patients with acute cervical pain of less than 3 days duration in a Physical Medicine and Rehabilitation Clinic in Sweden. Patients were randomized into one of 3 groups and treated with either neck collar alone, transcutaneous nerve stimulation, or manual non-thrust manipulation. Each of the latter 2 groups also received a neck collar. All groups were allowed to take analgesic pain medication. A physiotherapist performed the non-thrust mobilization and all patients were seen in follow-up after 1, 2, or 6 weeks, and after 3 months. All patients completed the study but the majority were so much improved after the first week that they did not need the second week of treatment. At one week of follow-up, transcutaneous nerve stimulation had comparable improvement in pain scores compared to the manipulation group. Both were better than the neck collar alone but the differences were not statistically significant. At 6 weeks, all patients were fully recovered.

Pikula randomly allocated 36 patients with acute unilateral neck pain (less than 2 weeks in duration) to one of 3 treatment groups: 1) SMT applied to the same side as the pain (ipsilateral), 2) SMT applied to the opposite side as the pain (contralateral), 3) a placebo group receiving only detuned ultrasound therapy. The patients were seen in a private chiropractic office in Canada and all treatments were provided by a chiropractor. Patients received a single high velocity, low amplitude thrust manipulation to the cervical spine to either the ipsilateral or contralateral neck with reference to the side of the cervical pain. The remaining 12 received 8 minutes of detuned ultrasound. A visual analogue pain score was determined both pretreatment and immediately post-treatment. In the ipsilateral SMT, pain scores improved from 42.5 to 23.6. In the contralateral CMT group, pain scores improved from 44.1 to 41.4 and in the placebo group, pain scores improved from 50.4 to 46.5. This pilot study demonstrated greater improvement in immediate pain scores using ipsilateral SMT than contralateral SMT ($p<0.05$).

Gonzales-Iglesias and colleagues completed 2 similar studies evaluating the effect of thoracic spine distraction thrust manipulation on patients with acute neck pain. In the first study, 45 patients aged 22-44 with acute neck pain (less than 3 weeks) were randomly allocated to one of 2 groups. The control group was treated at a physical therapy clinic in Spain with 6 sessions of TENS, superficial thermotherapy and soft tissue massage over a 3-week period. The experimental group received the same treatment as the control group and additionally received thoracic thrust manipulation once a week for 3 consecutive weeks. Pain was rated using the numerical pain rate scale (NPRS) at baseline and at 1 week after discharge. The level of disability (Northwick Park Neck Pain Questionnaire) and neck flexion were also assessed. Patients receiving thoracic spine manipulation experienced greater improvement in neck pain than the control group (32.8% (95% CI 29.9-35.8) vs 9.4% (95% CI 7.2-11.4); $p<.001$). Disability scores also showed significant improvement in the experimental group compared to the control group (12.6% (95% CI 11.4-13.8) vs 4.1% (95% CI 3.4-4.8); $p<0.001$).

In the second study by Gonzales-Iglesias, 45 patients with acute neck pain (less than 4 weeks) were randomly allocated to one of 2 groups. Both groups were treated at a physical therapy clinic in Spain with 5 sessions using standard electro/thermal therapy over a 3-week period. The program consisted of an infrared lamp for 15 minutes followed by 20 minutes of transcutaneous electrical nerve stimulation. The experimental group also received thoracic thrust manipulation for 3 consecutive Mondays. Pain was rated using a visual analogue scale (VAS) at baseline, immediately after the final treatment session (5th week), and at the 2- and 4-week follow-up visits. Level of disability (Northwick Park Neck Pain Questionnaire) and neck flexion were also assessed. Patients receiving thoracic manipulation experienced greater improvements in pain at the fifth week (final) treatment session and the 2- and 4-week follow-up visits ($p<0.001$), with pain improvement scores in the manipulation group of 16.8 mm and 26.5 mm greater than those in the control group at the 2- and 4-week visits, respectively. The experimental group also experienced significantly greater improvement improvements in disability with a between-group difference of 8.8 points (95% CI 7.5, 10.1); $p<0.001$ at the fifth visit and 8.0 points (95% CI 5.8, 10.2, $p<0.001$) at the 2-week follow-up.

Table 4. Evidence Table for Neck Pain Articles

Author, Year	Setting; Patients	Type of outcome	Tx arms	Baseline	Follow-up	Quality Score*
Howe 1983[81]	General practice at one 2-physician practice in England	Pain (dichotomous at baseline, follow-up measured as number of patients showing improvement)	Azapropazone plus thrust manipulation and/or steroid/local anesthetic injection (only 2 patients received injections)	N = 26 19/26 (73%)	Immediate improvement: 13/19 (68%) 1 week improvement: 14/19 (74%) 3 week improvement: 13/17 (76%)	V1 = ? V2 = + V3 = + V4 = - V5 = - V6 = - V7 = ? V8 = + V9 = ? V10 = + V11 = + Total = 5
			Azapropazone	N = 26 17/26(65%)	Immediate improvement: 1/17 (6%) 1 week improvement: 9/15 (60%) 3 week improvement: 7/12 (58%)	
Nordemar 1981[82]	Department of physical medicine and rehabilitation in Sweden	Visual Analogue Scale (VAS)	Non-thrust manual therapy and neck collar	N = 10 97 (46 SD)	1 week: 18 (25 SD)	V1 = ? V2 = ? V3 = ? V4 = - V5 = - V6 = - V7 = ? V8 = + V9 = + V10 = + V11 = + Total = 4
			Neck collar	N = 10 90 (26 SD)	1 week: 35 (45 SD)	
			Transcutaneous nerve stimulation	N = 10 83 (26 SD)	1 week: 17 (19 SD)	

Author, Year	Setting; Patients	Type of outcome	Tx arms	Baseline	Follow-up	Quality Score*
Pikula 1999[83]	Private chiropractic office in Canada	VAS	Thrust spinal manipulation, ipsilateral	N = 12 42.5 (19.8 SD)	Immediate: 23.6 (18.6 SD)	V1 = + V2 = + V3 = ? V4 = - V5 = - V6 = - V7 = + V8 = + V9 = ? V10 = + V11 = + Total = 6
			Thrust spinal manipulation, contralateral	N = 12 44.1 (27.5 SD)	Immediate: 41.4 (28.4 SD)	
			Detuned ultrasound	N = 12 50.4 (22.5 SD)	Immediate: 46.5 (21.8 SD)	
Gonzalez-Iglesias 2009[85]	Physical therapy clinic in Spain	Numerical pain rate scale	Transcutaneous nerve stimulation and thermotherapy	N = 22 5.37 (0.6 SD)	1 week: 4.3 (0.8 SD)	V1 = + V2 = + V3 = + V4 = - V5 = - V6 = - V7 = ? V8 = + V9 = + V10 = + V11 = + Total = 7
			Thoracic spine thrust manipulation	N = 23 5.6 (0.9SD)	1 week: 2.3 (1 SD)	
		Northwick Park Neck Pain Questionnaire	Transcutaneous nerve stimulation and thermotherapy	N = 2227.1 (2.7 SD)	1 week: 22.9 (2.9 SD)	
			Thoracic spine thrust manipulation	N = 23 27.8 (3.1 SD)	1 week: 15.2 (4.1 SD)	

Author, Year	Setting; Patients	Type of outcome	Tx arms	Baseline	Follow-up	Quality Score*
Gonzalez-Iglesias 2009[84]	Physical therapy clinic in Spain	VAS	Transcutaneous nerve stimulation and thermotherapy with thrust thoracic spine manipulation	N = 23 54.7 (8.2 SD)	Immediate: 20.2 (7.8 SD) 2 week: 26.4 (11.8 SD)	V1 = + V2 = + V3 = + V4 = - V5 = - V6 = - V7 = ? V8 = ? V9 = + V10 = + V11 = + Total = 6
			Transcutaneous nerve stimulation and thermotherapy	N = 22 52.7 (5.5 SD)	Immediate: 44.7 (5.5 SD) 2 week: 41.2 (6.1 SD)	
		Northwick Park Neck Pain Questionnaire	Transcutaneous nerve stimulation and thermotherapy with thrust thoracic spine manipulation	N = 23 27.9 (3.0 SD)	Immediate: 15.2 (3.9 SD) 2 week: 14.7 (2.8 SD)	
			Transcutaneous nerve stimulation and thermotherapy	N = 22 27.0 (3.1 SD)	Immediate: 23.1 (3.2 SD) 21.8 (3.3 SD)	

*Quality Criteria listed in Appendix B; + = yes, - = no, ? = unsure/don't know

Studies of a Clinical Prediction Rule for SMT for Neck Pain

There have been attempts to develop clinical prediction rules to identify patients with neck pain who are more likely to benefit from SMT. These studies are not as advanced compared to the studies of a clinical prediction rule for lower back pain.[86,87]

Adverse Events

No included neck pain studies reported any adverse events. For data about adverse events of SMT in general, please see the adverse events subheading under acute low back pain.

Summary of Findings

Only 5 studies were identified of SMT compared to a non-SMT treatment group for patients with acute neck pain. Although each study reported favorable results on at least one outcome, in total only 198 have been studied in total.

Quality of Evidence for Key Question 2

We rated the evidence as low that SMT improves outcomes in patients with acute neck pain due to study quality concerns and imprecision of results (too few studies).

KEY QUESTION 2A: What is the relationship between the use of spinal manipulation/chiropractic services for acute neck pain and the use of opiate medication?

Summary of Findings

None of the included studies reported on the use of analgesic medications or opiate medication as an outcome.

Quality of Evidence for Key Question 2A

With no evidence from included studies, we rated this evidence as insufficient.

SUMMARY AND DISCUSSION

SUMMARY OF EVIDENCE BY KEY QUESTION

KEY QUESTION 1: What are the benefits and harms of spinal manipulation/chiropractic services for acute lower back pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Twenty-six studies of SMT treatments for acute low back pain found overall statistically significant evidence of a clinical benefit that was, on average, modest. However, there was substantial heterogeneity in results, with some studies reporting much larger effects and some studies reporting no effect at all. We explored 6 potential sources of heterogeneity, and while there were some non-statistically significant differences that may be signals of possible effects of type of manipulation, selection of patients, and study quality, most of the differences in outcome between studies remain unexplained.

Mild transient musculoskeletal adverse events are common following SMT, although these may be equally common following non-SMT manual therapy. Serious adverse events have been the subject of case reports, but assessing causality has proved challenging.

There were too few studies of SMT in patients with acute back pain and sciatica to draw conclusions.

Mild transient musculoskeletal adverse events are common following SMT, although these may be equally common following non-SMT manual therapy. Serious adverse events have been the subject of case reports, but assessing causality has proved challenging.

There were too few studies of SMT in patients with acute back pain and sciatica to draw conclusions.

KEY QUESTION 1A: What is the relationship between the use of spinal manipulation/chiropractic services for lower back pain and the use of opiate medication?

Among the 26 studies included in our analysis only one specifically reported on the use of opiate medications.

KEY QUESTION 2: What are the benefits and harms of spinal manipulation/chiropractic services for acute neck pain (less than 6 weeks duration) compared to usual care or other forms of acute pain management?

Only 5 studies were identified of SMT compared to a non-SMT treatment group for patients with acute neck pain. Although each study reported favorable results on at least one outcome, in total only 198 patients have been studied.

KEY QUESTION 2A: What is the relationship between the use of spinal manipulation/chiropractic services for neck pain and the use of opiate medication?

None of the included studies reported on the use of analgesic medications or opiate medication as an outcome.

LIMITATIONS

Publication Bias

In general we did not find evidence of publication bias, although no evidence of bias is not the same as evidence of no publication bias.

Study Quality

Study quality was highly variable and our pooled analysis is split about equally between studies considered "high" and studies considered "low" quality. Our analysis found no evidence to support a hypothesis that our results are due to low-quality studies with inflated effect sizes.

Heterogeneity

Heterogeneity in the results is the primary limitation of this analysis. The statistical evidence of heterogeneity was significant and visual inspection of the forest plots illustrated this: some studies of SMT found, for the same outcome, found positive results, while others found essentially no benefit (ES = 0, ES = 0.06, *etc*). Our investigation of multiple potential sources of heterogeneity yielded no results that were statistically significant, although visually there were suggestions that the comparison group, the patients, and the type of SMT may be important. Nevertheless, the majority of heterogeneity remains unexplained and this larger degree of heterogeneity may limit the enthusiasm of some clinicians and policymakers for advocating more widespread use of SMT.

Applicability of Findings to the VA Population

We identified no studies specific to VA population. Nevertheless, acute back pain in primary care is probably quite similar within VA to outside VA, and these results have to be considered at least moderately applicable to VA populations.

RESEARCH GAPS/FUTURE RESEARCH

There continues to be a great deal of unexplained heterogeneity in results of SMT for acute low back pain, so a research gap is better understanding what contributes to patient selection and intervention to improve the consistency of the result. This could include an attempt at replication of the clinical prediction rule RCT or new RCTs with more detailed data collection on the patient clinical characteristics and details of the SMT intervention. For neck pain, there are simply too few studies to draw firm conclusions. Additional RCTs are warranted. Attention should be paid to collecting clinical variables and details of the intervention to use in the exploration of possible heterogeneity of treatment effects.

REFERENCES

1. Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal Manipulative Therapy for Low Back Pain: A Meta-Analysis of Effectiveness Relative to Other Therapies. *Annals of Internal Medicine.* 2003(11):871-881. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12003008464/frame.html.
2. Rubinstein SM, Terwee CB, Assendelft WJ, Boer MR, Tulder MW. Spinal manipulative therapy for acute low back pain: an update of the Cochrane review (Structured abstract). *Spine.* 2013(3):E158-e177. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12013013786/frame.html.
3. Cherkin DC, Deyo RA, Battie M, Street J, Barlow W. 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.* Oct 8 1998;339(15):1021-1029.
4. Hancock MJ, Maher CG, Latimer J, et al. Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. *Lancet.* Nov 10 2007;370(9599):1638-1643.
5. van Tulder MW, Suttorp M, Morton S, Bouter LM, Shekelle P. Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. *Spine (Phila Pa 1976).* Jul 15 2009;34(16):1685-1692.
6. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC medical research methodology.* 2014;14(1):25.
7. Sánchez-Meca J, Marín-Martínez F. Confidence intervals for the overall effect size in random-effects meta-analysis. *Psychological Methods.* 2008;13(1):31.
8. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal.* 2003;327(7414):557.
9. StataCorp. *Stata Statistical Software: Release 12.* College Station, TX: StataCorp LP;2011.
10. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 1994;50:1088-1101.
11. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629-634.
12. Godfrey CM, Morgan PP, Schatzker J. A randomized trial of manipulation for low-back pain in a medical setting. *Spine (Phila Pa 1976).* Apr 1984;9(3):301-304.
13. Hallegraeff JM, de Greef M, Winters JC, Lucas C. Manipulative therapy and clinical prediction criteria in treatment of acute nonspecific low back pain. *Percept Mot Skills.* Feb 2009;108(1):196-208.
14. Rubinstein Sidney M, Terwee Caroline B, Assendelft Willem JJ, de Boer Michiel R, van Tulder Maurits W. Spinal manipulative therapy for acute low-back pain. *Cochrane Database of Systematic Reviews.* 2012(9). http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008880.pub2/abstract.
15. Olsen MF, Bjerre, E., Hansen, M. D., Hilden, J., Landler, N. E., Tendal, B., & Hróbjartsson, A. . Pain relief that matters to patients: systematic review of empirical

studies assessing the minimum clinically important difference in acute pain. *BMC Medicine* 2017;15(35).

16. GRADE working group. 2014; http://www.gradeworkinggroup.org/. Accessed 09/25/15.

17. Blomberg S, Hallin G, Grann K, Berg E, Sennerby U. Manual therapy with steroid injections--a new approach to treatment of low back pain. A controlled multicenter trial with an evaluation by orthopedic surgeons. *Spine (Phila Pa 1976).* Mar 1 1994;19(5):569-577.

18. Blomberg S, Svardsudd K, Mildenberger F. A controlled, multicentre trial of manual therapy in low-back pain. Initial status, sick-leave and pain score during follow-up. *Scand J Prim Health Care.* Sep 1992;10(3):170-178.

19. Blomberg S, Svardsudd K, Tibblin G. Manual therapy with steroid injections in low-back pain. Improvement of quality of life in a controlled trial with four months' follow-up. *Scand J Prim Health Care.* Jun 1993;11(2):83-90.

20. Blomberg S, Svardsudd K, Tibblin G. A randomized study of manual therapy with steroid injections in low-back pain. Telephone interview follow-up of pain, disability, recovery and drug consumption. *Eur Spine J.* 1994;3(5):246-254.

21. Blomberg S, Tibblin G. A controlled, multicentre trial of manual therapy with steroid injections in low-back pain: functional variables, side effects and complications during four months follow-up. *Clinical Rehabilitation.* 1993;7(1):49-62.

22. Grunnesjo MI, Bogefeldt JP, Svardsudd KF, Blomberg SI. A randomized controlled clinical trial of stay-active care versus manual therapy in addition to stay-active care: functional variables and pain. *J Manipulative Physiol Ther.* Sep 2004;27(7):431-441.

23. Grunnesjö MI, Bogefeldt JP, Blomberg SI, Strender LE, Svärdsudd KF. A randomized controlled trial of the effects of muscle stretching, manual therapy and steroid injections in addition to 'stay active' care on health-related quality of life in acute or subacute low back pain. *Clinical rehabilitation.* 2011;25(11):999-1010. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/578/CN-00865578/frame.html.

24. Bogefeldt J, Grunnesjo MI, Svardsudd K, Blomberg S. Sick leave reductions from a comprehensive manual therapy programme for low back pain: the Gotland Low Back Pain Study. *Clin Rehabil.* Jun 2008;22(6):529-541.

25. Bergquist-Ullman M, Larsson U. Acute low back pain in industry. A controlled prospective study with special reference to therapy and confounding factors. *Acta Orthop Scand.* 1977(170):1-117.

26. Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med.* Dec 21 2004;141(12):920-928.

27. Cramer GD, Humphreys CR, Hondras MA, McGregor M, Triano JJ. The Hmax/Mmax ratio as an outcome measure for acute low back pain. *J Manipulative Physiol Ther.* Jan 1993;16(1):7-13.

28. Cruser A, Maurer D, Hensel K, Brown SK, White K, Stoll ST. A randomized, controlled trial of osteopathic manipulative treatment for acute low back pain in active duty military personnel. *Journal of manual & manipulative therapy.* 2012;20(1):5-15. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/708/CN-00897708/frame.html.

29. Delitto A, Cibulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther.* Apr 1993;73(4):216-222; discussion 223-218.

30. Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Phys Ther.* Dec 1994;74(12):1093-1100.

31. Farrell JP, Twomey LT. Acute low back pain. Comparison of two conservative treatment approaches. *Med J Aust.* Feb 20 1982;1(4):160-164.

32. Fritz JM, Magel JS, McFadden M, et al. Early physical therapy vs usual care in patients with recent-onset low back pain: A randomized clinical trial. *JAMA - Journal of the American Medical Association.* 13 2015;314(14):1459-1467.

33. Glover JR, Morris JG, Khosla T. Back pain: a randomized clinical trial of rotational manipulation of the trunk. *Br J Ind Med.* Jan 1974;31(1):59-64.

34. Goertz CM, Long CR, Hondras MA, et al. Adding chiropractic manipulative therapy to standard medical care for patients with acute low back pain: results of a pragmatic randomized comparative effectiveness study. *Spine.* 2013;38(8):627-634. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/298/CN-00967298/frame.html.

35. Hadler NM, Curtis P, Gillings DB, Stinnett S. A benefit of spinal manipulation as adjunctive therapy for acute low-back pain: a stratified controlled trial. *Spine (Phila Pa 1976).* Sep 1987;12(7):702-706.

36. Heymann WJ, Schloemer P, Timm J, Muehlbauer B. Spinal high-velocity low amplitude manipulation in acute nonspecific low back pain: a double-blinded randomized controlled trial in comparison with diclofenac and placebo. *Spine.* 2013;38(7):540-548. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/716/CN-00878716/frame.html.

37. Hoiriis KT, Pfleger B, McDuffie FC, et al. A randomized clinical trial comparing chiropractic adjustments to muscle relaxants for subacute low back pain. *J Manipulative Physiol Ther.* Jul-Aug 2004;27(6):388-398.

38. Juni P, Battaglia M, Nuesch E, et al. A randomised controlled trial of spinal manipulative therapy in acute low back pain. *Ann Rheum Dis.* Sep 2009;68(9):1420-1427.

39. MacDonald RS, Bell CM. An open controlled assessment of osteopathic manipulation in nonspecific low-back pain. *Spine (Phila Pa 1976).* May 1990;15(5):364-370.

40. Morton JE. Manipulation in the treatment of acute low back pain. *Journal of Manual & Manipulative Therapy.* 1999;7(4):182-189.

41. Postacchini F, Facchini M, Palieri P. Efficacy of various forms of conservative treatment in low back pain: a comparative study. *Neuro-orthopedics.* 1988;6(1):28-35.

42. Rasmussen G. Manipulation in treatment of low back pain: a randomized clinical trial. *Man Med.* 1979;1:8-10.

43. Skargren EI, Carlsson PG, Oberg BE. One-year follow-up comparison of the cost and effectiveness of chiropractic and physiotherapy as primary management for back pain. Subgroup analysis, recurrence, and additional health care utilization. *Spine (Phila Pa 1976).* Sep 1 1998;23(17):1875-1883; discussion 1884.

44. Waterworth RF, Hunter IA. An open study of diflunisal, conservative and manipulative therapy in the management of acute mechanical low back pain. *N Z Med J.* May 22 1985;98(779):372-375.

45. Roelofs PD, Deyo RA, Koes BW, Scholten RJ, van Tulder MW. Nonsteroidal anti-inflammatory drugs for low back pain: an updated Cochrane review. *Spine (Phila Pa 1976).* Jul 15 2008;33(16):1766-1774.

46. Cherkin DC, Deyo RA, Street JH, Hunt M, Barlow W. Pitfalls of patient education. Limited success of a program for back pain in primary care. *Spine (Phila Pa 1976).* Feb 1 1996;21(3):345-355.

47. Machado GC, Maher CG, Ferreira PH, et al. Efficacy and safety of paracetamol for spinal pain and osteoarthritis: systematic review and meta-analysis of randomised placebo controlled trials. *BMJ.* 2015;350:h1225.

48. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine (Phila Pa 1976).* Dec 15 2002;27(24):2835-2843.

49. Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. *Eur Spine J.* Jul 2008;17(7):936-943.

50. Cleland JA, Fritz JM, Kulig K, et al. Comparison of the effectiveness of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule: a randomized clinical trial. *Spine (Phila Pa 1976).* Dec 1 2009;34(25):2720-2729.

51. Learman K, Showalter C, O'Halloran B, Donaldson M, Cook C. No Differences in Outcomes in People with Low Back Pain Who Met the Clinical Prediction Rule for Lumbar Spine Manipulation When a Pragmatic Non-thrust Manipulation Was Used as the Comparator. *Physiotherapy Canada.* 2014;66(4):359-366.

52. Schenk R, Dionne C, Simon C, Johnson R. Effectiveness of mechanical diagnosis and therapy in patients with back pain who meet a clinical prediction rule for spinal manipulation. *Journal of manual & manipulative therapy.* 2012;20(1):43-49. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/754/CN-01018754/frame.html.

53. Mathews JA, Mills SB, Jenkins VM, et al. Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol.* Dec 1987;26(6):416-423.

54. Mathews W, Morkel M, Mathews J. Manipulation and traction for lumbago and sciatica: physiotherapeutic techniques used in two controlled trials. *Physiotherapy Theory and Practice.* 1988;4(4):201-206.

55. Santilli V, Beghi E, Finucci S. Chiropractic manipulation in the treatment of acute back pain and sciatica with disc protrusion: a randomized double-blind clinical trial of active and simulated spinal manipulations. *Spine J.* Mar-Apr 2006;6(2):131-137.

56. Paanalahti K, Holm LW, Nordin M, Asker M, Lyander J, Skillgate E. Adverse events after manual therapy among patients seeking care for neck and/or back pain: a randomized controlled trial. *BMC Musculoskeletal Disorders.* 2014;15:77.

57. Barrett AJ, Breen AC. Adverse effects of spinal manipulation. *J R Soc Med.* May 2000;93(5):258-259.

58. Cagnie B, Vinck E, Beernaert A, Cambier D. How common are side effects of spinal manipulation and can these side effects be predicted? *Man Ther.* Aug 2004;9(3):151-156.

59. Leboeuf-Yde C, Hennius B, Rudberg E, Leufvenmark P, Thunman M. Side effects of chiropractic treatment: a prospective study. *J Manipulative Physiol Ther.* Oct 1997;20(8):511-515.

60. Rubinstein SM, Leboeuf-Yde C, Knol DL, de Koekkoek TE, Pfeifle CE, van Tulder MW. Predictors of adverse events following chiropractic care for patients with neck pain. *Journal of Manipulative and Physiological Therapeutics.* 2008;31(2):94-103.

61. Senstad O, Leboeuf-Yde C, Borchgrevink C. Frequency and characteristics of side effects of spinal manipulative therapy. *Spine (Phila Pa 1976).* Feb 15 1997;22(4):435-440; discussion 440-431.

62. Maiers M, Evans R, Hartvigsen J, Schulz C, Bronfort G. Adverse events among seniors receiving spinal manipulation and exercise in a randomized clinical trial. *Manual Therapy.* 1 2015;20(2):335-341.

63. Walker BF, Hebert JJ, Stomski NJ, et al. Outcomes of usual chiropractic. The OUCH randomized controlled trial of adverse events. *Spine.* Sep 15 2013;38(20):1723-1729.

64. Ernst E. Ophthalmological adverse effects of (chiropractic) upper spinal manipulation: evidence from recent case reports (Structured abstract). *Acta Ophthalmologica Scandinavica.* 2005;83(5):581-585. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12005004755/frame.html.

65. Ernst E. Deaths after chiropractic: a review of published cases (Provisional abstract). *International Journal of Clinical Practice.* 2010;64(8):1162-1165. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12011007078/frame.html.

66. Oliphant D. Safety of spinal manipulation in the treatment of lumbar disk herniations: a systematic review and risk assessment (Provisional abstract). *Journal of Manipulative and Physiological Therapeutics.* 2004;27(3):197-210. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12004001048/frame.html.

67. Hebert JJ, Stomski NJ, French SD, Rubinstein SM. Serious adverse events and spinal manipulative therapy of the low back region: a systematic review of cases (Provisional abstract). *Journal of Manipulative and Physiological Therapeutics.* 2013(4):epub. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12013036142/frame.html.

68. Stevinson C, Ernst E. Risks associated with spinal manipulation (Structured abstract). *American Journal of Medicine.* 2002(7):566-571. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12002001305/frame.html.

69. Carlesso LC, Gross AR, Santaguida PL, Burnie S, Voth S, Sadi J. Adverse events associated with the use of cervical manipulation and mobilization for the treatment of neck pain in adults: a systematic review (Provisional abstract). *Manual Therapy.* 2010(5):434-444. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12010006882/frame.html.

70. Chung CL, Cote P, Stern P, L'Esperance G. The association between cervical spine manipulation and carotid artery dissection: a systematic review of the literature (Provisional abstract). *Journal of Manipulative and Physiological Therapeutics.* 2014(4). http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12014005028/frame.html.

71. Ernst E. Manipulation of the cervical spine: a systematic review of case reports of serious adverse events, 1995 - 2001 (Structured abstract). *Medical Journal of Australia.*

2002(8):376-380. http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12002001060/frame.html.

72. Walker BF, Hebert JJ, Stomski NJ, Losco B, French SD. Short-term usual chiropractic care for spinal pain: a randomized controlled trial. *Spine.* 2013;38(24):2071-2078. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/967/CN-00914967/frame.html.

73. Kosloff TM, Elton D, Tao J, Bannister WM. Chiropractic care and the risk of vertebrobasilar stroke: results of a case-control study in U.S. commercial and Medicare Advantage populations. *Chiropr Man Therap.* 2015;23:19.

74. Whedon JM, Mackenzie TA, Phillips RB, Lurie JD. Risk of Traumatic Injury Associated With Chiropractic Spinal Manipulation in Medicare Part B Beneficiaries Aged 66 to 99 Years. *Spine.* 2015;40(4):264-270.

75. Cassidy JD, Boyle E, Côté P, et al. Risk of vertebrobasilar stroke and chiropractic care: results of a population-based case-control and case-crossover study. *Journal of Manipulative and Physiological Therapeutics.* 2009;32(2):S201-S208.

76. Nielsen SM, Tarp S, Christensen R, Bliddal H, Klokker L, Henriksen M. The risk associated with spinal manipulation: an overview of reviews. *Syst Rev.* Mar 24 2017;6(1):64.

77. Rhee Y, Taitel MS, Walker DR, Lau DT. Narcotic drug use among patients with lower back pain in employer health plans: a retrospective analysis of risk factors and health care services. *Clinical therapeutics.* 2007;29(11):2603-2612.

78. Vogt MT, Kwoh CK, Cope DK, Osial TA, Culyba M, Starz TW. Analgesic usage for low back pain: impact on health care costs and service use. *Spine.* 2005;30(9):1075-1081.

79. Franklin GM, Rahman EA, Turner JA, Daniell WE, Fulton-Kehoe D. Opioid use for chronic low back pain: A prospective, population-based study among injured workers in Washington state, 2002-2005. *The Clinical journal of pain.* 2009;25(9):743-751.

80. Allen H, Wright M, Craig T, et al. Tracking Low Back Problems in a Major Self-Insured Workforce: Toward Improvement in the Patient's Journey. *Journal of Occupational and Environmental Medicine.* 2014;56(6):604-620.

81. Howe DH, Newcombe RG, Wade MT. Manipulation of the cervical spine--a pilot study. *J R Coll Gen Pract.* Sep 1983;33(254):574-579.

82. Nordemar R, Thorner C. Treatment of acute cervical pain--a comparative group study. *Pain.* Feb 1981;10(1):93-101.

83. Pikula JR. The effect of spinal manipulative therapy (SMT) on pain reduction and range of motion in patients with acute unilateral neck pain: a pilot study. *The Journal of the Canadian Chiropractic Association.* 1999;43(2):111.

84. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Gutierrez-Vega Mdel R. Thoracic spine manipulation for the management of patients with neck pain: a randomized clinical trial. *J Orthop Sports Phys Ther.* Jan 2009;39(1):20-27.

85. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Alburquerque-Sendin F, Palomeque-del-Cerro L, Mendez-Sanchez R. Inclusion of thoracic spine thrust manipulation into an electro-therapy/thermal program for the management of patients with acute mechanical neck pain: a randomized clinical trial. *Man Ther.* Jun 2009;14(3):306-313.

86. Puentedura EJ, Cleland JA, Landers MR, Mintken P, Louw A, Fernández-de-las-Peñas C. Development of a clinical prediction rule to identify patients with neck pain likely to

benefit from thrust joint manipulation to the cervical spine. *journal of orthopaedic & sports physical therapy.* 2012;42(7):577-592.

87. Dunning JR, Cleland JA, Waldrop MA, et al. Upper cervical and upper thoracic thrust manipulation versus nonthrust mobilization in patients with mechanical neck pain: a multicenter randomized clinical trial. *journal of orthopaedic & sports physical therapy.* 2012;42(1):5-18.

88. Furlan AD, Pennick V, Bombardier C, van Tulder M. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine.* 2009;34(18):1929-1941.

APPENDIX A. SEARCH STRATEGIES

1. SYSTEMATIC REVIEW SEARCH STRATEGIES

SEARCH STRATEGY FOR “CHIROPRACTIC” SYSTEMATIC REVIEWS
DATABASE SEARCHED:
Cochrane Database of Systematic Reviews and Other Reviews

NO DATE OR LANGUAGE LIMITATIONS

SEARCH STRATEGY:
'chiroprac* in Title, Abstract, Keywords
Cochrane Reviews (17)
Other Reviews (44)

SEARCH STRATEGY:
"Manipulation, Spinal"

Cochrane Database Search Strategy #2:
spine or spinal or neck or back or cervi*
and
(smt or manipulat* or chiropract*):ti,ab,kw

Dates:
2011-present,

Limit to the Cochrane Systematic Reviews, Other Reviews (DARE), Technology Assessments, and Economic Evaluations databases.

Forward search on:
Hurwitz EL, Aker PD, Adams AH, Meeker WC, Shekelle PG. Manipulation and mobilization of the cervical spine. A systematic review of the literature. Spine (Phila Pa 1976). Aug 1 1996;21(15):1746-1759; discussion 1759-1760.

2. UPDATE SEARCH STRATEGIES

SPINAL MANIPULATION THERAPY – 2015 UPDATE
SEARCH METHODOLOGY

DATABASE SEARCHED & TIME PERIOD COVERED:
COCHRANE CENTRAL – 1/1/2011-2/06/2017

SEARCH STRATEGY:

#1 MeSH descriptor: [Back] explode all trees

#2 MeSH descriptor: [Buttocks] this term only

#3 MeSH descriptor: [Leg] this term only

#4 MeSH descriptor: [Back Pain] explode all trees
#5 MeSH descriptor: [Back Pain] 1 tree(s) exploded
#6 MeSH descriptor: [Back Injuries] explode all trees
#7 MeSH descriptor: [Low Back Pain] this term only
#8 MeSH descriptor: [Sciatica] this term only
#9 low next back next pain
#10 lbp
#11 #1 or #2 or #3 or #5 or #6 or #7 or #8 or #9 or #10
#12 MeSH descriptor: [Musculoskeletal Manipulations] explode all trees
#13 MeSH descriptor: [Chiropractic] explode all trees
#14 manip*
#15 MeSH descriptor: [Osteopathic Medicine] explode all trees
#16 osteopath*
#17 chiropract*
#18 #12 or #13 or #14 or #15 or #16 or #17
#19 #11 and #18
= =

DATABASE SEARCHED & TIME PERIOD COVERED:
MEDLINE ON OVID – 1/1/2011-2/06/2017

SEARCH STRATEGY:
1 Clinical Trial.pt.
2 randomized.ab,ti.
3 placebo.ab,ti.
4 dt.fs.
5 randomly.ab,ti.
6 trial.ab,ti.
7 groups.ab,ti.
8
1 or 2 or 3 or 4 or 5 or 6 or 7
9 Animals/
10 Humans/
11
9 not (9 and 10) Including Related Terms
12
8 not 11

13 dorsalgia.ti,ab.

14 exp Back Pain/

15 backache.ti,ab.

16 (lumbar adj pain).ti,ab.

17 coccyx.ti,ab.

18 coccydynia.ti,ab.

19 sciatica.ti,ab.

20 sciatica/

21 spondylosis.ti,ab.

22 lumbago.ti,ab.

23 exp low back pain/

24

13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23

25 exp Manipulation, Chiropractic/

26 exp Manipulation, Orthopedic/

27 exp Manipulation, Osteopathic/

28 exp Manipulation, Spinal/

29 exp Musculoskeletal Manipulations/

30 exp Chiropractic/

31 manipulation.mp.

32 manipulate.mp.

33 exp Orthopedics/

34 exp Osteopathic Medicine/

35

25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34

36

12 and 24 and 35

37

36 and 2011:2015.(sa_year).

= =

DATABASE SEARCHED & TIME PERIOD COVERED:
EMBASE – 1/1/2011-2/06/2017

SEARCH STRATEGY:
#2 'clinical article'/exp OR 'clinical study'/exp OR 'clinical trial'/de OR 'controlled study'/de OR 'randomized controlled trial'/de OR 'major clinical study'/de OR 'double blind procedure'/de OR 'multicenter study'/de OR 'single blind procedure'/de OR 'phase 3 clinical trial'/de OR 'phase 4 clinical trial'/de OR 'crossover procedure'/de OR 'placebo'/de

#6 allocat*

#7 assign*

#8 blind*

#12 clinical NEAR/25 (study OR trial*)

#13 compar*

#14 control*

#17 'cross over'

#18 'cross-over'

#19 'crossover'

#20 factorial

#21 'follow up'

#22 follow* NEAR/3 up

#23 'follow up'

#24 placebo*

#25 prospectiv*

#26 random*

#27 (singl* OR doubl* OR trebl* OR tripl*) NEAR/25 (blind* OR mask*)

#28 trial

#29 versus OR vs

#30

#6 OR #7 OR #8 OR #12 OR #13 OR #14 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23

OR #24 OR #25 OR #26 OR #27 OR #28 OR #29

#31

#2 OR #30

#34 dorsalgia

#35 'back pain'

#36 lumbar NEAR/2 pain

#37 coccyx

#38 coccydynia

#39 sciatica

#40 spondylosis

#41 lumbago

#42 'backache'/exp OR 'ischialgia'/exp OR 'low back pain'/exp

#43

#34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42

#44 'chiropractic'/exp OR 'orthopedic manipulation'/exp OR 'manipulative medicine'/exp OR 'osteopathic medicine'/exp OR 'orthopedics'/exp

#45 manipulation

#46 manipulate

#47 osteopathy

#48

#44 OR #45 OR #46 OR #47

#49

#31 AND #43 AND #48

#50

#31 AND #43 AND #48 AND [humans]/lim

#51

#31 AND #43 AND #48 AND [humans]/lim AND [2011-2015]/py

= =

DATABASE SEARCHED & TIME PERIOD COVERED:
CINAHL – 1/1/2011-2/06/2017

SEARCH STRATEGY:

Search modes - Find all search terms (For all search statements)

S1 randomized controlled trials

S2 randomized controlled trials

S3 PT clinical trial

S4 (MH "Clinical Trials+")

S5 clin* n25 trial*

S6 (singl* or doubl* or trebl* or tripl*) n25 (blind* or mask*)

S7 (MH "Placebos")

S8 (MH "Study Design+")

S9 (MH "Comparative Studies")

S10 (MH "Evaluation Research+")

S11 (MH "Prospective Studies+")

S12 "follow up studies" OR "follow-up studies" OR "followup studies" OR "follow-up study" OR "follow up study" OR "followup study"

S13 control* or prospectiv* or volunteer*

S14 placebo* OR random* OR (latin n2 square*)

S15

S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14

S18 TI dorsalgia OR AB dorsalgia

S19 (MH "Back Pain+")

S20 TI backache OR AB backache

S21 TI lumbar n2 pain OR AB lumbar n2 pain

S22 TI coccyx pain OR AB lumbar n2 pain

S23 TI coccyx OR AB coccyx

S24 TI coccydynia OR AB coccydynia

S25 TI sciatica OR AB sciatica

S26 (MH "Sciatica")

S27 TI spondylosis OR AB spondylosis

S28 TI lumbago cronico OR AB spondylosis

S29 TI lumbago OR AB lumbago

S30 (MH "Low Back Pain")

S31

S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29

OR S30

S32 (MH "Chiropractic+")

S33 (MH "Manipulation, Chiropractic")

S34 (MH "Manipulation, Orthopedic")

S35 (MH "Manipulation, Osteopathic")

S36 (MH "Manual Therapy+")

S37 (MH "Orthopedics")

S38 (MH "Osteopathy+")

S39 manipulation

S40 manipulate

S41

S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40

S42

S15 AND S31 AND S41

S43

S15 AND S31 AND S41

= =

DATABASE SEARCHED & TIME PERIOD COVERED:

PubMed – 1/1/2015-2/06/2017

SEARCH STRATEGY:
Manipulation, Chiropractic[mh] OR Manipulation, Orthopedic[mh] OR Manipulation, Osteopathic[mh] OR Manipulation, Spinal[mh] OR Musculoskeletal Manipulations[mh] OR Chiropractic[mh] OR Orthopedics OR Osteopathic Medicine
AND
"Low Back Pain"[Mesh] OR low back pain*[tiab] OR "Back"[Mesh] OR dorsalgia[tiab] OR Back Pain[mh] OR backache[tiab] OR “lumbar pain”[tiab] OR coccyx[tiab] OR coccydynia[tiab] OR sciatica[tiab] OR sciatica[mh] OR spondylosis[tiab] OR lumbago[tiab]
AND
Randomized Controlled Trial" [Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR random*[tiab] OR rct* OR systematic[tiab] OR systematic[sb] OR Clinical Trial[pt] OR randomized[tiab] OR placebo[tiab] OR randomly[tiab] OR trial[tiab] OR groups[tiab]

APPENDIX B. COCHRANE BACK GROUP RISK OF BIAS TOOL

COCHRANE BACK REVIEW GROUP (CBRG) CRITERIA

2003 Version[88]

Domain	Operationalization of the Criteria List	Reviewers' judgment
V1. Randomization	A random (unpredictable) assignment sequence. Examples of adequate methods are computer generated random number table and use of sealed opaque envelopes. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.	Was the method of randomization adequate? Yes / No / Don't Know
V2. Concealment	Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient.	Was the treatment allocation concealed? Yes / No / Don't Know
V3. Baseline differences	In order to receive a yes, groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with neurologic symptoms, and value of main outcome measure(s). [adapt as required by topic]	Were the groups similar at baseline regarding the most important prognostic indicators? Yes / No / Don't Know
V4. Blinding - patient	The reviewer determines if enough information about blinding is given in order to score a yes.	Was the patient blinded to the intervention? Yes / No / Don't Know
V5. Blinding – care provider	The reviewer determines if enough information about blinding is given in order to score a yes.	Was the care provider blinded to the intervention? Yes / No / Don't Know
V6. Blinding - outcome	The reviewer determines if enough information about blinding is given in order to score a yes.	Was the outcome assessor blinded to the intervention? Yes / No / Don't Know
V7. Co-interventions	Co-interventions should either be avoided in the trial design or similar between the index and control groups.	Were co-interventions avoided or similar? Yes / No / Don't Know
V8.	The reviewer determines if the compliance to the interventions is	Was the compliance

Compliance	acceptable, based on the reported intensity, duration, number and frequency of sessions for both the index intervention and control intervention(s).	acceptable in all groups? Yes / No / Don't Know
V9. Dropouts	The number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given. If the percentage of withdrawals and drop-outs does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias a yes is scored. (N.B. these percentages are arbitrary, not supported by literature)	Was the drop-out rate described and acceptable? Yes / No / Don't Know
V10. Timing	Timing of outcome assessment should be identical for all intervention groups and for all important outcome assessments.	Was the timing of the outcome assessment in all groups similar? Yes / No / Don't Know
V11. ITT	All randomized patients are reported/analyzed in the group they were allocated to by randomization for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and co-interventions.	Did the analysis include an intention-to-treat analysis? Yes / No / Don't Know

APPENDIX C. PEER REVIEW COMMENTS/AUTHOR RESPONSES

Comment	Response
Recommend considering change to the time labels given to the outcome periods. Currently these labels are "short term" and "long term." Traditionally "long term" outcomes for low back pain would be considered to be 3-12 months rather than 3-6 weeks. To the casual reader, this may be misleading.	We have updated these time labels so that less than 2 weeks is now “immediate” instead of “short” and 3 to 6 weeks is “short” instead of “long” term. These labels were chosen to align with terms used by a previous publication: Chou et al. Epidural corticosteroid injections for radiculopathy and spinal stenosis. Ann Intern Med. 2015; 163 (5): 373-381.
I recommend considering addition of a measure of clinical importance to the outcomes. For example, use of the Minimum Clinically Important Difference (MCID) may help the reader to interpret the clinical significance of relatively small changes in pain and function.	We have now incorporated a discussion of MCID into the Data Synthesis section. Choosing an MCID for the Roland Scale proved challenging, and in the report we explain this in detail.
Also consider including a reference or explanation of what is statistically and clinically significant change in pain and function for the reader	
Consider including a reference for readers to refer to and learn about forest plots. Some readers may not be familiar with this idea. The extensive use of forest plots would support educating readers on this subject.	We have now incorporated a reference at our first mention of forest plots in the Data Synthesis section: Greenhalgh T. How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses). BMJ. 1997;315(7109):672-675.
Page 3, first paragraph - the sentence is confusing. “six studies each scored 3 points and 2 points.” Maybe change to “six studies scored 3 points and the other 6 studies scored 2 points”.	This change has been made.
Page 8, line 22 “ESP” is not defined. Consider including what this acronym is.	This change has been made.
Page 9, line 9 – Participants are defined as “Adults”. Consider defining what an adult is such as 18+ and children as less than 18 years of age.	This change has been made.
Page 9, line 16 – Spinal manipulation is not clearly defined. Consider adding the HVLA definition here or referring the reader to the second paragraph under data extraction on page 9.	We have added language that refers readers to the mentioned data abstraction section.
Page 9, line 30 – settings: Why were hospital settings excluded? The VA is a hospital-based setting. I suggest including the rationale for excluding these studies.	The decision to focus only on ambulatory patients was made at the outset by the topic nominators; this was specified in the Topic Nomination Brief.
Page 9, line 55 – was TEP defined earlier? If not define this acronym.	This change has been made

Comment	Response
Page 26, line 15 the word enrollment is misspelled.	This change has been made.
Published or unpublished studies that may have been overlooked: (1) Puentedura CPR for cervical manipulation versus non-thrust (validation pending), (2) Dunning CPR for cervical AND thoracic thrust versus cervical and thoracic non-thrust. (3) Cook 2012 Manual Therapy article also directly compares thrust to non-thrust manipulation for LBP.	(1 & 2) These articles did not meet inclusion criteria because the pain was not acute in their patient populations, and are now in the "exclude-background" group for their relevance to clinical prediction rules. They have been incorporated into the text. (3) This article was identified by our searches but excluded from our review because it did not focus on acute pain.
Other ESPs have included cohort studies in the analysis, particularly in the area of assessing harms. I realize that you have included some prospective cohort studies in the harms analysis. However I would suggest including some of the higher quality retrospective cohort studies that have looked at harms and opiate use. Understood that this would be lower level evidence, however even a sidebar discussion of this can help provide some better information, particularly in light of the current situation regarding opiate use for musculoskeletal pain.	We have now incorporated the 7 articles identified by the reviewer into either the Serious Adverse Events section (Cassidy, Kosloff, Whedon) or the Key Question 2 section (Rhee, Vogt, Franklin, Allen) with discussion where appropriate.
I understood (perhaps incorrectly) at the outset that this ESP would also include review / analysis of retrospective cohort studies of harms and opiate use? If this was accomplished, I may have overlooked this information given there are only a few - but important - such reports. Our office considers this information relevant and important, particularly given the increased scrutiny on use of opiates for pain management, and rehabilitative alternatives (such as chiropractic, CAM, and other various treatment modalities).	

APPENDIX D. EVIDENCE TABLE OF 26 RANDOMIZED CLINICAL TRIALS OF SPINAL MANIPULATIVE THERAPY FOR ACUTE LOW BACK PAIN

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Bergquist-Ullman, et al, 1977[25]	Industry	87% male	34 years	14% of patients had a straight leg raise test positive at less than 60 degrees	Pain index	43	back school (instruction and exercise)	N=44	10 day median: 20 3 week median: 19 6 week median: 22
						42	non-thrust manipulation	N=50	10 day median: 22 3 week median: 18 6 week median: 21
						42	diathermy according to Cyriax, Kaltenborn, Lewit, and Janda	N=56	10 day median: 28 3 week median: 25 6 week median: 17
Blomberg, et al, 1994[17-21]	Primary care	52% male	37 years	10% with “true radicular pain”	Disability Rating Score (function)	no baseline data	usual medical care	N=48	3 days mean: 4.6 1 week mean: 3.9 2 week mean: 3.2 3 week mean: 3
							mix of thrust and non-thrust manipulation, some patients also got steroid injections of the parasacrococcygeal structures as described by Cyriax	N=53	3 days mean: 3.5 1 week mean: 2.6 2 week mean: 1.8 3 week mean: 1.4
					Pain score		usual medical care	N=48	3 days mean: 4.8 1 week mean: 4.2 2 week mean: 3.4 3 week mean: 3.4
							mix of thrust and non-thrust manipulation, some patients also got steroid injections of the parasacrococcygeal structures as described by Cyriax	N=53	3 days mean: 3.8 1 week mean: 3.1 2 week mean: 2 3 week mean: 1.7

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Cherkin, et al, 1998[3]	Primary care patients from health maintenance organization	52% male	41 years	Sciatica excluded	Roland Morris Disability questionnaire (function)	12.1 (CI: 11.2-13.1)	thrust manipulation	N=122	4 week mean: 3.7 (2.9 SD)
						11.7 (CI: 10.4-13.0)	physical therapy according to McKenzie	N=136	4 week mean: 4.1 (3.3 SD)
						11.7 (CI: 10.4-13.0)	educational booklet	N=66	4 week mean: 4.9 (3.8 SD)
					Bothersomeness of symptoms (pain)	5.5 (CI: 5.1-5.8)	thrust manipulation	N=122	4 week mean: 1.9 (1.5 SD)
						6 (CI: 5.6-6.5)	physical therapy according to McKenzie	N=136	4 week mean: 2.3 (1.9 SD)
						5.3 (CI: 4.9-5.7)	educational booklet	N=66	4 week mean: 3.1 (2.4 SD)
Childs, et al, 2004[26]	8 physical therapy clinics in the United States	58% male	34 years	24% had "symptoms distal to knee"	Oswestry disability questionnaire (function)	41.4 (10.1 SD)	thrust manipulation	N=70	1 week mean: 14.6 4 week mean: 8.4
						40.9 (10.8 SD)	low stress aerobic exercise and lumbar spine strengthening program according to Agency for Health Care Policy and Research guidelines	N=61	1 week mean: 35 4 week mean: 23
Cramer, et al, 1993[27]	Clinical chiropractic college	57% male	Not reported	Patients with "compressive neuropathy" we excluded	Visual Analogue Scale (pain)	71.8 (14.8 SD)	non-thrust manipulation and electrical stimulation and cold pack	N=17	10 day mean: 38.6 (25.2 SD)
						72 (19.2 SD)	detuned ultrasound and cold pack	N=18	10 day mean: 42 (28.8 SD)
					Oswestry disability questionnaire (function)	17.6 (11.9 SD)	non-thrust manipulation and electrical stimulation and cold pack	N=17	10 day mean: 7.3 (6.8 SD)
						14.9 (5.0 SD)	detuned ultrasound and cold pack	N=18	10 day mean: 8.0 (7.6 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Cruser, et al, 2012[28]	United States military facility	55% male	27 years	Not reported	Visual Analogue Scale (pain)	5.2 (2.1 SD)	mix of thrust and non-thrust manipulation, soft tissue stretching, myofascial release, counterstrain muscle energy, sacroiliac articulation	N=30	4 week mean: 2.0 (1.5 SD)
						5.5 (2.2 SD)	usual medical care	N=30	4 week mean: 3.7 (2.4 SD)
					Roland Morris Disability questionnaire (function)	12.4 (5.3 SD)	mix of thrust and non-thrust manipulation, soft tissue stretching, myofascial release, counterstrain muscle energy, sacroiliac articulation	N=30	4 week mean: 4.4 (5.9 SD)
						12.5 (6.0 SD)	usual medical care	N=30	4 week mean: 7.31 (6.3 SD)
Delitto, et al, 1993[29]	Physiotherapy department	58% male	33 years	21% had “leg symptoms”	Oswestry disability questionnaire (function)	33 (5 SD)	thrust manipulation and extension exercises according to McKenzie and hand-heel rock exercise	N=14	3 day mean: 20 (5 SD) 5 day mean: 10 (5 SD)
						41 (5 SD)	flexion exercises according to Williams	N=10	3 day mean: 36 (5 SD) 5 day mean: 32 (4 SD)
Erhard, et al, 1994[30]	Physiotherapy department	62% male	44 years	8% had “leg symptoms”	Oswestry disability questionnaire (function)	45 (12 SD)	thrust manipulation and extension exercises according to McKenzie	N=12	3 day mean: 20 (8 SD) 5 day mean: 8 (8 SD)
						40 (12 SD)	extension exercises according to McKenzie	N=12	3 day mean: 35 (8 SD)5 day mean: 25 (14 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Farrell, et al, 1982[31]	Setting unclear	62% male	42 years	Not reported	Subjective pain rating	4.95	non-thrust manipulation according to Stoddart and Maitland	N=24	3 week mean: 0.3
						5.3	physical therapy and diathermy, isometric abdominal exercises and ergonomic instructions	N=24	3 week mean: 0.3
Fritz, et al, 2015[32]	Primary care	48% male	37 years	Patients with presence of pain or numbness distal to the knee were excluded	Numeric pain rating of low back pain severity	no baseline data	thrust manipulation and exercises	N=108	4 week mean: 1.7 (1.9 SD)
						no baseline data	standard medical care and self-help booklet	N=112	4 week mean: 2.1 (1.9 SD)
					Oswestry disability questionnaire (function)	no baseline data	thrust manipulation and exercises	N=108	4 week mean: 11.1 (12.5 SD)
						no baseline data	standard medical care and self-help booklet	N=112	4 week mean: 14.5 (13.2 SD)
Glover, et al, 1974[33]	Work medical center	89% male	39 years	Not reported	Percent pain relief	no baseline data	diathermy	N=41	3 day mean: 56 1 week mean: 80
						no baseline data	non-thrust manipulation	N=43	3 day mean: 50 1 week mean: 75
Godfrey, et al, 1984[12]	Patients referred from primary care	Not reported	42 years	Not reported	General symptomatology (number of patients with marked improvement) (pain)	no baseline data	thrust manipulation according to Maigne		2-3 week: 14/39 (35.9%)
						no baseline data	light effleurage and minimal electrostimulation		2-3 week: 7/33 (21.2%)
					Activities of Daily Living (number of	no baseline data	thrust manipulation according to Maigne		2-3 week: 7/24 (29.2%)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
					patients with moderate improvement) (function)	no baseline data	light effleurage and minimal electrostimulation		2-3 week: 5/17 (29.4%)
Goertz, et al, 2013[34]	United States army medical center	86% male	26 years	43% had "radicular signs"	Numerical pain rating scale	5.8 (2.1 SD)	standard medical care and brief massage, ice or heat, McKenzie exercises, stretching exercises	N=46	2 week mean: 6.1 4 week mean: 5.2
						5.8 (1.5 SD)	thrust manipulation	N=45	2 week mean: 3.9 4 week mean: 3.9
					Roland Morris Disability questionnaire (function)	12.7 (5.1 SD)	standard medical care and brief massage, ice or heat, McKenzie exercises, strengthening exercises	N=46	2 week mean: 12.9 4 week mean: 12
						11 (4.2 SD)	thrust manipulation	N=45	2 week mean: 8.9 4 week mean: 8
Grunnesjö, et al, 2004[22-24]	Nine primary health care and one outpatient orthopedic hospital department	56% male	41 years	8% had "verified herniations"	Pain last 24 hours	52.2 (CI: 46.7-57.8)	stay active	N=71	5 week mean: 29.7 (25.8 SD)
						54.7 (CI: 49.8-59.6)	mix of thrust and non-thrust manipulation and stay active and in some patients a steroid injection in the parasacrococcygeal region	N=89	5 week mean: 20.8 (23.3 SD)
					All disability rating variables	52 (CI: 47.4-56.6)	stay active	N=71	5 week mean: 31.9 (21.9 SD)
						57.8 (CI: 53.7-61.8)	mix of thrust and non-thrust manipulation and stay active and in some patients a steroid injection in the parasacrococcygeal region	N=89	5 week mean: 25.8 (22.1 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Hadler, et al, 1987[35]	Primary care	57% male	Not reported	Not reported	Roland Morris Disability questionnaire (function)	no baseline data	mobilization	N=28	9 day mean: 4.5 12 day mean: 3.7
						no baseline data	thrust manipulation	N=26	9 day mean: 3.7 12 day mean: 3.4
Hallegraeff, et al, 2009[13]	Three physical therapy and manual therapy centers	55% male	39 years	Patients with symptoms distal to the knee were excluded	Oswestry disability questionnaire (function)	0.24 (0.18 SD)	thrust manipulation	N=31	2.5 week mean: 0.14 (0.17 SD)
						0.26 (0.12 SD)	physical therapy	N=33	2.5 week mean: 0.14 (0.12 SD)
					Visual Analogue Scale (pain)	42.7 (18.4 SD)	thrust manipulation	N=31	2.5 week mean: 19 (16.9 SD)
						54 (17.5 SD)	physical therapy	N=33	2.5 week mean: 24.8 (20.1 SD)
Hancock, et al, 2007[4]	Patients referred from primary care	56% male	41 years	Patients with "nerve root compromise" were excluded	Numerical pain rating scale negative effect size favors manipulation	no baseline data	non-thrust manipulation	N=59	1 week effect size: 0.2 (CI: -0.3-0.7) 2 week effect size: -0.4 (CI: -1.0, 0.1) 4 week effect size: -0.2 (CI: -0.7, 0.3)
						no baseline data	detuned pulsed ultrasound (sham)	N=60	
					Roland Morris Disability questionnaire (function) negative effect size favors manipulation	no baseline data	non-thrust manipulation	N=59	1 week effect size: -0.7 (CI: -2.1, 0.6) 2 week effect size: -1.4 (CI: -2.7, -0.1) 4 week effect size: -1 (CI: -2.1, 0.1)
						no baseline data	detuned pulsed ultrasound (sham)	N=60	

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Heymann, et al, 2013[36]	5 orthopedic or general practices	60% male	37 years	Not reported	Roland Morris Disability questionnaire (function)	13.5 (5.6 SD)	thrust manipulation	N=38	1 week mean: 5.8
						14.4 (4.8 SD)	analgesic (diclofenac)	N=37	1 week mean: 9.7
						15 (3.8 SD)	sham	N=25	no data provided
					Visual Analogue Scale (pain)	no baseline data	thrust manipulation	N=38	1 week mean: 10
						no baseline data	analgesic	N=37	1 week mean: 30
						no baseline data	sham	N=25	1 week mean: no data provided
Hoiriis, et al, 2004[37]	Patients recruited via advertisement	57% male	42 years	Patients with "known or suspected disk herniation" were excluded	Visual Analogue Scale (pain)	4.52 (1.82 SD)	thrust manipulation	N=34	2 week mean: 2.4 (2.2 SD) 4 week mean: 1.7 (1.9 SD)
						3.9 (2.0 SD)	muscle relaxants (cyclobenzaprine or carisoprodol or methocarbamol)	N=36	2 week mean: 2.7 (2.2 SD) 4 week mean: 2.2 (2.2 SD)
						3.8 (1.6 SD)	sham	N=40	2 week mean: 3.2 (2.4 SD) 4 week mean: 2.2 (2.0 SD)
					Oswestry disability questionnaire (function)	24.8 (11.5 SD)	thrust manipulation	N=46	2 week mean: 17.0 (13.8 SD) 4 week mean: 11.9 (11.9 SD)
						22.8 (12.9 SD)	muscle relaxants (cyclobenzaprine or carisoprodol or methocarbamol)	N=47	2 week mean: 17.0 (12.2 SD) 4 week mean:16.0 (16.1 SD)
						24.8 (11.7 SD)	sham	N=48	2 week mean: 19.3 (13.7 SD) 4 week mean: 16.3 (12.6 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Juni, et al, 2009[38]	Patients referred from emergency department or a general practice	64% male	35 years	Patients with "signs of nerve root irritation or compression" were excluded	Roland Morris Disability questionnaire (function)	12.8 (5.1 SD)	Mix of thrust and non-thrust manipulation	N=52 12.8 (5.1 SD)	2 week mean: 5.8 (5.7 SD)
						14.3 (4.9 SD)	analgesic (paracetamol, diclofenac, or dihydrocodeine)	N=52	2 week mean: 5.2 (7.0 SD)
					Pain intensity, BS-11 score positive favors manipulation	6.3 (2.2 SD)	mix of thrust and non-thrust manipulation	N=52	Difference of 0.5 (2.6 SD)
						6.8 (2.2 SD)	Analgesic (paracetamol, diclofenac, or dihydrocodeine)	N=52	
MacDonald, et al, 1990[39]	General practice	41% male	Not reported	Patients with "neurologic deficits" were excluded	Improvement in the disability index	6.4 (3 SD)	thrust manipulation and advice on posture, exercises and avoidance of occupational stress	N=36	2 week mean: 4.1 (3.5 SD)
						6.1 (2.5 SD)	advice on posture, exercise, and avoidance of occupational stress	N=30	2 week mean: 4.4 (3.5 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Morton, 1999[40]	Patients referred from primary care	34% male	44 years	Patients with "abnormalities on neurologic exam" were excluded	Roland Morris Disability questionnaire (function)	10.6 (5.2 SD)	thrust manipulation	N=15	1 week mean: 6.9 (4.1 SD) 2 week mean: 6.0 (2.3 SD) 3 week mean: 3.7 (3.7 SD) 4 week mean: 1.9 (2.5 SD)
						10.1 (6.4 SD)	spinal stabilizing exercises	N=14	1 week mean: 9.1 (5.9 SD) 2 week mean: 7.9 (6.3 SD) 3 week mean: 7 (6.1 SD) 4 week mean: 6 (5.2 SD)
					Visual Analogue Scale (pain)	49.7 (23.6 SD)	thrust manipulation	N=15	1 week mean: 27.6 (15.2 SD) 2 week mean: 17.4 (13.9 SD) 3 week mean: 7.5 (6.4 SD) 4 week mean: 2.4 (3 SD)
						46.6 (25.1 SD)	spinal stabilizing exercises	N=14	1 week mean: 46.4 (23.3 SD) 2 week mean: 36.6 (24.6 SD) 3 week mean: 34.5 (23 SD) 4 week mean: 25.4 (17.3 SD)

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Postacchini, et al, 1988[41]	Hospital outpatient department	51% male	38 years	Not reported	Improvement in low back pain from pre-treatment	no baseline data	thrust manipulation	N=53	3 week mean: 8.5
						no baseline data	back school	N=17	3 week mean: 10.4
						no baseline data	analgesics (diclofenac)	N=49	3 week mean: 9.4
						no baseline data	physiotherapy of light massage, analgesic currents, and diathermy	N=47	3 week mean: 8.1
						no baseline data	bed rest	N=29	3 week mean: 6.6
						no baseline data	topical gel	N=46	3 week mean: 5.8
Rasmussen, 1979[42]	Hospital department of physical medicine and rheumatology	Not reported	35 years	Patients with "signs of root pressure" were excluded	Number of patients with total restorement of all symptoms	no baseline data	non-thrust manipulation	N=12	11/12 (91.7%)
						no baseline data	diathermy	N=12	3/12 (25%)
Skargren, et al, 1998[43]	Primary care centers	38% male	41 years	Not reported	Visual Analogue Scale (pain) negative favors manipulation	56 (22 SD)	thrust manipulation	N=172	4-5 week difference: -0.16 (CI: -6.47, 6.15)
						61 (21 SD)	physiotherapy	N=144	
					Oswestry disability questionnaire (function) negative favors manipulation	35 (17 SD)	thrust manipulation	N=172	4-5 week difference: -1.49 (CI: -5.51, 2.54)
						37 (16 SD)	physiotherapy	N=144	

Author, Year	Setting	% Male	Mean Age	Presence of Leg Pain or Sciatica	Outcome	Baseline value	Treatment arms	Sample Size	Follow-up
Waterworth, et al, 1985[44]	General practice	62% male	36 years	Not reported	Score of lower back pain	2.1	non-thrust manipulation	N=38	12 day mean: 0.42
						2.1	analgesic (diflunisal)	N=36	12 day mean: 0.44
						2	physiotherapy including local heat, ultrasound, and flexion and extension exercises	N=34	12 day mean: 0.38
					Patient has overall improvement score of excellent	no baseline data	non-thrust manipulation	N=38	23/38 (60.5%)
						no baseline data	analgesic (diflunisal)	N=36	15/36 (41.7%)
						no baseline data	physiotherapy including local heat, ultrasound, and flexion and extension exercises	N=34	13/34 (38.2%)