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
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.

TABLE OF CONTENTS

EXECUTIVE SUMMARY ........................................................................................................... 1
Introduction ............................................................................................................................... 1
Methods ................................................................................................................................. 1
  Data Sources and Searches ................................................................................................. 1
  Study Selection ................................................................................................................... 1
  Data Abstraction and Quality Assessment ............................................................................ 2
  Data Synthesis and Analysis ............................................................................................... 2
Results ..................................................................................................................................... 2
  Results of Literature Search ............................................................................................... 2
  Acute Low Back Pain without Sciatica ................................................................................ 3
  Acute Low Back Pain with Sciatica .................................................................................... 4
  Adverse Events .................................................................................................................. 4
Summary of Results for Key Questions and Strength of Evidence ....................................... 5
Discussion ........................................................................................................................... 6
  Limitations ......................................................................................................................... 6
  Study Quality .................................................................................................................... 6
  Heterogeneity ..................................................................................................................... 6
  Applicability of Findings to the VA Population ................................................................. 6
  Research Gaps/Future Research ....................................................................................... 7
EVIDENCE REPORT ................................................................................................................... 8
INTRODUCTION ...................................................................................................................... 8
METHODS ............................................................................................................................... 9
  Topic Development ........................................................................................................... 9
  Search Strategy .................................................................................................................. 9
  Study Selection ................................................................................................................ 10
  Data Abstraction .............................................................................................................. 10
  Quality Assessment ......................................................................................................... 11
  Data Synthesis ................................................................................................................ 11
  Rating the Body of Evidence ......................................................................................... 12
  Peer Review ....................................................................................................................... 12
RESULTS .................................................................................................................................. 13
  Literature Flow ............................................................................................................... 13
  Description of the Evidence .............................................................................................. 15
Effectiveness and Harms of Spinal Manipulative Therapy Evidence-based Synthesis Program for the Treatment of Acute Neck and Lower Back Pain

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? ........... 17

  Acute Low Back Pain without Sciatica ........................................................................ 17
  Acute Low Back Pain with Sciatica ........................................................................... 32
  Adverse Events ....................................................................................................... 32
  Summary of Findings .............................................................................................. 38
  Quality of Evidence for Key Question 1 ...................................................................... 38

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? .... 38

  Quality of Evidence for Key Question 1A ................................................................. 39

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? ........... 39

  Studies of a Clinical Prediction Rule for SMT for Neck Pain ..................................... 44
  Adverse Events ....................................................................................................... 44
  Summary of Findings .............................................................................................. 44
  Quality of Evidence for Key Question 2 ...................................................................... 44

KEY QUESTION 2A: What is the relationship between the use of spinal manipulation/chiropractic services for neck pain and the use of opiate medication? ........... 44

  Summary of Findings .............................................................................................. 44
  Quality of Evidence for Key Question 2A ................................................................. 44

SUMMARY AND DISCUSSION .................................................................................. 45

Summary of Evidence by Key Question ................................................................... 45

  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? ...... 45

  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? 45

  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? .................. 45

  KEY QUESTION 2A: What is the relationship between the use of spinal manipulation/chiropractic services for neck pain and the use of opiate medication? .... 46

Limitations .................................................................................................................. 46

Publication Bias ........................................................................................................ 46

Study Quality .......................................................................................................... 46

Heterogeneity ......................................................................................................... 46
Effectiveness and Harms of Spinal Manipulative Therapy
for the Treatment of Acute Neck and Lower Back Pain

Applicability of Findings to the VA Population ................................................................. 46
Research Gaps/Future Research .......................................................................................... 46

REFERENCES .......................................................................................................................... 47

TABLES
Table 1. Quality Scores of RCTs of SMT for Acute Low Back Pain ........................................ 16
Table 2. Adverse events reported in randomized clinical trials of effectiveness
of spinal manipulative therapy for acute low back pain ....................................................... 33
Table 3. Results from cohort studies and randomized clinical trials focused on
adverse events of spinal manipulative therapy ....................................................................... 36
Table 4. Evidence Table for Neck Pain Articles ................................................................... 41

FIGURES
Figure 1. Literature Flow Chart .......................................................................................... 14
Figure 2. Effect of SMT on Immediate term Pain ................................................................. 18
Figure 3. Effect of SMT on Immediate term Function ............................................................ 20
Figure 4. Effect of SMT on Short term Pain ......................................................................... 22
Figure 5. Effect of SMT on Short term Function .................................................................... 24
Figure 6. Effect of Thrust Compared to Non-thrust SMT, by Outcome ................................. 27
Figure 7. Quality Scores ...................................................................................................... 28

APPENDIX A. SEARCH STRATEGIES ............................................................................. 54
1. Systematic Review Search Strategies ............................................................................... 54
2. Update Search Strategies ............................................................................................... 54

APPENDIX B. COCHRANE BACK GROUP RISK OF BIAS TOOL .................................... 61
Cochrane Back Review Group (CBRG) Criteria .................................................................. 61

APPENDIX C. PEER REVIEW COMMENTS/AUTHOR RESPONSES ............................... 63

APPENDIX D. EVIDENCE TABLE OF 26 RANDOMIZED CLINICAL TRIALS OF
SPINAL MANIPULATIVE THERAPY FOR ACUTE LOW BACK PAIN ..................... 65
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. The most recent Cochrane Review on the subject, however, concluded that SMT is not more effective than any other intervention or sham. 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:

**Participants:** 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 12.09 and R3.2.2. The Begg’s rank correlation10 and Egger regression asymmetry test11 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 studies12,13 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
Effectiveness and Harms of Spinal Manipulative Therapy Evidence-based Synthesis Program for the Treatment of Acute Neck and Lower Back Pain

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 our 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.
Effectiveness and Harms of Spinal Manipulative Therapy Evidence-based Synthesis Program for the Treatment of Acute Neck and Lower Back Pain

Figure 1. Literature Flow Chart

- Systematic review search results: 182
  - Excluded at title screen: 89
    - Titles retrieved from systematic reviews: 136
      - Titles for review: 1,790
        - Excluded at title screen: 1,564
          - Full texts for review: 198
            - Excluded at abstract screen: 28
              - Full Text Excludes: 150
                - 77 Not acute low back pain
                - 38 Not RCT
                - 10 Duplicates
                - 9 Background information only
                - 7 Not SMT
                - 3 No relevant outcome
                - 3 Other reasons
                - 2 Not ambulatory adult patients
                - 1 Not available

- Update search results: 1,639
  - Expert suggestion: 15
    - Titles for review: 1,790
      - Excluded at title screen: 1,564
        - Abstracts for review: 226
          - Excluded at abstract screen: 28

- Titles for review: 1,790
  - Abstracts for review: 93
    - Excluded at abstract screen: 13
      - Full texts for review: 80
        - Full Text Excludes: 31
          - 25 Duplicates
          - 1 Not SMT
          - 1 Unable to retrieve
          - 2 Cost effectiveness
          - 2 Background information only

- Systematic reviews eligible for title screening: 49
  - Titles retrieved from systematic reviews: 136
    - Abstracts for review: 226
      - Excluded at abstract screen: 28

- Included articles: 48*
  - Effectiveness: 40 articles
    - 3 sciatica subpopulation
    - 2 clinical prediction rule evaluations
    - 2 no eligible outcome data
    - 1 clinically unique patient population
    - 6 findings presented in another reference included in analyses
  - Adverse Events: 8 articles
  - Included in analyses: 26 articles
Effectiveness and Harms of Spinal Manipulative Therapy Evidence-based Synthesis Program for the Treatment of Acute Neck and Lower Back Pain

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

<table>
<thead>
<tr>
<th>Article</th>
<th>Randomization</th>
<th>Concealment</th>
<th>Blinding, provider</th>
<th>Blinding, patient</th>
<th>Blinding, outcome</th>
<th>Dropouts</th>
<th>Timing</th>
<th>ITT</th>
<th>Baseline Differences</th>
<th>Co-interventions</th>
<th>Compliance</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bergquist-Ullman 1977&lt;sup&gt;25&lt;/sup&gt;</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td></td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Blomberg&lt;sup&gt;17-21&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>6</td>
</tr>
<tr>
<td>Cherkin 1998&lt;sup&gt;3&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td></td>
<td>?</td>
<td>6</td>
</tr>
<tr>
<td>Childs 2004&lt;sup&gt;26&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>Cruser 2012&lt;sup&gt;28&lt;/sup&gt;</td>
<td>?</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>Delitto 1993&lt;sup&gt;29&lt;/sup&gt;</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td></td>
<td>?</td>
<td>4</td>
</tr>
<tr>
<td>Erhard 1994&lt;sup&gt;30&lt;/sup&gt;</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td></td>
<td>?</td>
<td>3</td>
</tr>
<tr>
<td>Fritz 2015&lt;sup&gt;32&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>Glover 1974&lt;sup&gt;33&lt;/sup&gt;</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td></td>
<td>?</td>
<td>2</td>
</tr>
<tr>
<td>Godfrey 1984&lt;sup&gt;12&lt;/sup&gt;</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td></td>
<td>+</td>
<td>6</td>
</tr>
<tr>
<td>Goertz 2013&lt;sup&gt;34&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>Grunnesjö&lt;sup&gt;22-24&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>Hallegraeff 2009&lt;sup&gt;13&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td>+</td>
<td></td>
<td>+</td>
<td>6</td>
</tr>
<tr>
<td>Hancock 2007&lt;sup&gt;4&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>9</td>
</tr>
<tr>
<td>Heymann 2013&lt;sup&gt;36&lt;/sup&gt;</td>
<td>?</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td></td>
<td>?</td>
<td>6</td>
</tr>
<tr>
<td>Hoiriis 2004&lt;sup&gt;37&lt;/sup&gt;</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td></td>
<td>?</td>
<td>3</td>
</tr>
<tr>
<td>Juni 2009&lt;sup&gt;38&lt;/sup&gt;</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td></td>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>MacDonald 1990&lt;sup&gt;39&lt;/sup&gt;</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td></td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Morton 1999&lt;sup&gt;40&lt;/sup&gt;</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td></td>
<td>?</td>
<td>3</td>
</tr>
<tr>
<td>Skargren 1998&lt;sup&gt;43&lt;/sup&gt;</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>?</td>
<td></td>
<td>?</td>
<td>2</td>
</tr>
</tbody>
</table>

* + = 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 I² = 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

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcome</th>
<th>Sample Size</th>
<th>Mean [95% Confidence Interval]</th>
<th>Comparator</th>
<th>Sample Size</th>
<th>Mean [95% Confidence Interval]</th>
<th>Mean Difference</th>
<th>95% Confidence Interval</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hancock, 2007</td>
<td>mns</td>
<td>119</td>
<td>*</td>
<td>120</td>
<td>*</td>
<td></td>
<td>-4.00</td>
<td>[-9.50; 1.50]</td>
<td>5</td>
</tr>
</tbody>
</table>

Comparison Group = Sham

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcome</th>
<th>Sample Size</th>
<th>Mean [95% Confidence Interval]</th>
<th>Comparator</th>
<th>Sample Size</th>
<th>Mean [95% Confidence Interval]</th>
<th>Mean Difference</th>
<th>95% Confidence Interval</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jull, 2010</td>
<td>mns</td>
<td>52</td>
<td>*</td>
<td>52</td>
<td>*</td>
<td></td>
<td>5.00</td>
<td>[-2.00; 12.00]</td>
<td>7</td>
</tr>
<tr>
<td>Houw, 2014</td>
<td>VAS</td>
<td>34</td>
<td>24 [17.31]</td>
<td>36</td>
<td>27 [20.34]</td>
<td></td>
<td>-0.90</td>
<td>[-13.55; 7.50]</td>
<td>5</td>
</tr>
</tbody>
</table>

Random effects model

- * = outcome data not reported by group, only between group data reported
- Mean difference represents effect size based on the random-effects meta-analysis
- High score = worse pain
- Quality score uses the Cochrane Bias Criteria (0-11)
- VAS = Visual Analog Scale (0-100) or (0-80); converted to 0-100
- NRS = Numeric Rating Scale (0-10); converted to 0-100
- n = 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

Table: Effectiveness of SMT for Acute Neck and Lower Back Pain

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Outcome</th>
<th>Sample Size</th>
<th>Mean</th>
<th>Standardized Mean Difference</th>
<th>95% Confidence Interval</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hancock, 2007</td>
<td>RMDQ</td>
<td>118</td>
<td>*</td>
<td>-0.14</td>
<td>[-0.26, -0.01]</td>
<td>9</td>
</tr>
<tr>
<td>Haake, 2004</td>
<td>ODI</td>
<td>46</td>
<td>126</td>
<td>-0.17</td>
<td>[-0.57, 0.24]</td>
<td>3</td>
</tr>
<tr>
<td>Halder, 1987</td>
<td>RMDQ</td>
<td>36</td>
<td>200</td>
<td>-0.14</td>
<td>[-0.49, 0.20]</td>
<td>3</td>
</tr>
<tr>
<td>Random effects model</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison Group = Sham

Comparison Group = all other therapies

Avni, 2006 | RMDQ    | 52          | 6 (4.8) | 62 | 5 (3.7) | 0.09 | [-0.22, 0.40] | 7 |
| Haake, 2004 | OD1    | 46          | 17 (13.21) | 46 | 17 (14.29) | 0.00 | [-0.40, 0.41] | 3 |
| Goetz, 2013 | RMDQ    | 46          | 13 (11.15) | 46 | 13 (11.2) | -0.03 | [-1.12, 0.06] | 7 |
| Heylmann, 2013 | RMDQ    | 36          | 0 (6.0) | 37 | 0 (8.12) | -0.67 | [-1.10, -0.24] | 7 |
| Hallgren, 2009 | OD1    | 21          | 21 (9.20) | 23 | 14 (10.19) | 0.00 | [-0.48, 0.46] | 6 |
| Cromer, 1993 | OD1    | 17          | 10 (9.32) | 10 | 16 (9.23) | -0.16 | [-1.18, 0.86] | 3 |
| Morton, 1990 | RMDQ    | 15          | 0 (5.7) | 14 | 0 (5.11) | -0.42 | [-1.16, 0.22] | 3 |

Random effects model

* = estimate not reported by group, only between group data reported.

Quality score uses the Common Base Group test (0-11):

RMDQ=Radiology and Medicine Disability Questionnaire (0-34)
OD1=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. 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 and 3 studies concluded it did not.
Effectiveness and Harms of Spinal Manipulative Therapy
for the Treatment of Acute Neck and Lower Back Pain

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

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>N Manip</th>
<th>Mean Manip</th>
<th>N Comp</th>
<th>Mean Comp</th>
<th>Quality Score</th>
<th>SMD 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hancock, 2007</td>
<td>59</td>
<td>60</td>
<td>8</td>
<td></td>
<td></td>
<td>-0.18 [-0.34; 0.02]</td>
</tr>
<tr>
<td>Holits, 2004</td>
<td>46</td>
<td>12</td>
<td>48</td>
<td>16</td>
<td>3</td>
<td>-0.35 [-0.76; 0.06]</td>
</tr>
</tbody>
</table>

Comparison Group = Sham

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>N Manip</th>
<th>Mean Manip</th>
<th>N Comp</th>
<th>Mean Comp</th>
<th>Quality Score</th>
<th>SMD 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skargren, 1989</td>
<td>172</td>
<td>139</td>
<td>2</td>
<td></td>
<td></td>
<td>-0.04 [-0.15; 0.07]</td>
</tr>
<tr>
<td>Chekwa, 1988</td>
<td>118</td>
<td>4</td>
<td>60</td>
<td>5</td>
<td>3</td>
<td>-0.37 [-0.66; -0.06]</td>
</tr>
<tr>
<td>Holits, 2004</td>
<td>46</td>
<td>12</td>
<td>47</td>
<td>16</td>
<td>3</td>
<td>-0.29 [-0.70; 0.12]</td>
</tr>
<tr>
<td>Goertz, 2013</td>
<td>45</td>
<td>8</td>
<td>46</td>
<td>12</td>
<td>7</td>
<td>-0.67 [-1.08; -0.24]</td>
</tr>
<tr>
<td>Cruser, 2012</td>
<td>39</td>
<td>4</td>
<td>30</td>
<td>7</td>
<td>7</td>
<td>-0.47 [-0.98; 0.04]</td>
</tr>
<tr>
<td>Morton, 1990</td>
<td>15</td>
<td>2</td>
<td>14</td>
<td>6</td>
<td>3</td>
<td>-1.09 [-1.78; -0.32]</td>
</tr>
</tbody>
</table>

Random effects model

Total N = 929
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
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. These 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, 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. 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. 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. 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.

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. 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, 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. 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\textsuperscript{25,43} and one each reported shorter and longer sick leave for the SMT group\textsuperscript{24,44}), one study reported no differences in SF-12 outcomes,\textsuperscript{36} and 2 studies reported utilization data.\textsuperscript{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. 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. 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 week for 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.
Table 2. Adverse events reported in randomized clinical trials of effectiveness of spinal manipulative therapy for acute low back pain

<table>
<thead>
<tr>
<th>Author/Year</th>
<th>Sample Size</th>
<th>Method for assessing adverse events</th>
<th>Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blomberg, 1993&lt;sup&gt;21&lt;/sup&gt;</td>
<td>N=149</td>
<td>Closed end questionnaires about side effects given to patients at 1, 2, and 4 months</td>
<td>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.</td>
</tr>
<tr>
<td>Fritz, 2015&lt;sup&gt;32&lt;/sup&gt;</td>
<td>N=220</td>
<td>Open and closed end questionnaire about side effects given to patients at 4 weeks</td>
<td>“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).”</td>
</tr>
<tr>
<td>Goertz, 2013&lt;sup&gt;34&lt;/sup&gt;</td>
<td>N=91</td>
<td>Adverse event data collection method not specified</td>
<td>“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]</td>
</tr>
<tr>
<td>Hancock, 2007&lt;sup&gt;4&lt;/sup&gt;</td>
<td>N=240</td>
<td>Spontaneous reporting and open-ended questions</td>
<td>“No participants reported serious adverse reactions associated with spinal manipulative therapy.”</td>
</tr>
<tr>
<td>Heymann, 2013&lt;sup&gt;36&lt;/sup&gt;</td>
<td>N=100</td>
<td>Adverse event data collection method not specified</td>
<td>“Safety analysis did not show any unexpected untoward events in either group.”</td>
</tr>
<tr>
<td>Juni, 2009&lt;sup&gt;38&lt;/sup&gt;</td>
<td>N=104</td>
<td>Adverse event data collection method not specified</td>
<td>“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.”</td>
</tr>
<tr>
<td>Author/Year</td>
<td>Sample Size</td>
<td>Method for assessing adverse events</td>
<td>Adverse Events</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Morton, 1999&lt;sup&gt;10&lt;/sup&gt;</td>
<td>N=29</td>
<td>Adverse event data collection method not specified</td>
<td>“No adverse effects were documented for either group.”</td>
</tr>
<tr>
<td>Waterworth, 1985&lt;sup&gt;44&lt;/sup&gt;</td>
<td>N=108</td>
<td>Adverse event data collection method not specified</td>
<td>“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 than group 1.”</td>
</tr>
</tbody>
</table>

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. 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

<table>
<thead>
<tr>
<th>Article/Study Design</th>
<th>Sample Size</th>
<th>Method for assessing adverse events</th>
<th>Interventions</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barrett, AJ., Breen, AC., 2000(^{57}) Prospective cohort</td>
<td>68 patients 11 chiropractors</td>
<td>Collected from questionnaires to be given to 12 consecutive new patients</td>
<td>All received spinal manipulation</td>
<td>53% reported an adverse event, mostly increased or radiating pain.</td>
</tr>
<tr>
<td>Cagnie, B., 2004(^{58}) Prospective cohort</td>
<td>465 patients 51 manipulating clinicians</td>
<td>Collected by questionnaires to be given to 15 consecutive new patients</td>
<td>All received spinal manipulation</td>
<td>283 patients (61%) reported at least 1 reaction. Headache, stiffness, aggravation of complaints, and radiating discomfort accounted for 2/3 of reactions.</td>
</tr>
<tr>
<td>Leboeuf-Yde, C., 1997(^{59}) Prospective cohort</td>
<td>625 patients 66 chiropractors</td>
<td>Collected from questionnaires to be given to 10 consecutive patients</td>
<td>All received spinal manipulation</td>
<td>Treatment reactions were common, but benign and short lasting</td>
</tr>
<tr>
<td>Rubinstein, S. 2008(^{60}) Prospective cohort</td>
<td>529 Patients with neck pain 79 chiropractors</td>
<td>Collected from questionnaires completed by patients at regularly scheduled visits</td>
<td>All received spinal manipulation</td>
<td>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.</td>
</tr>
<tr>
<td>Senstad, O., 1997(^{61}) Prospective cohort</td>
<td>1050 patients 102 chiropractors</td>
<td>Collected by the chiropractor asking 12 consecutive patients a set of standardized questions</td>
<td>All received spinal manipulation</td>
<td>At least one reaction was reported by 580 patients (55%), 53% reported reactions were local discomfort.</td>
</tr>
<tr>
<td>Maiers, M., 2014(^{62}) Randomized clinical trial</td>
<td>194 elderly patients with neck pain</td>
<td>Collected by standardized solicitation by clinicians, unsolicited reporting of patients and qualitative interviews with patients</td>
<td>Patients were randomized to receive spinal manipulative therapy, home exercise, or supervised rehabilitation exercise.</td>
<td>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).</td>
</tr>
<tr>
<td>Article/Study Design</td>
<td>Sample Size</td>
<td>Method for assessing adverse events</td>
<td>Interventions</td>
<td>Findings</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
<td>-------------------------------------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>Paanalahti, K., 2014&lt;sup&gt;56&lt;/sup&gt; Randomized clinical trial</td>
<td>767 patients</td>
<td>Collected by questionnaires given to patients in the waiting room at each return visit</td>
<td>Patients were randomized to spinal manipulative therapy, manual therapy without spinal manipulation, and manual therapy without stretching</td>
<td>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.</td>
</tr>
<tr>
<td>Walker, B.F., 2013&lt;sup&gt;63&lt;/sup&gt; Randomized clinical trial</td>
<td>198 patients 12 chiropractors</td>
<td>Collected by questionnaires completed within 48 hours of treatment</td>
<td>Patients were randomized to usual chiropractic care (96% received spinal manipulative therapy) or a sham.</td>
<td>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.</td>
</tr>
</tbody>
</table>
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. 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.” 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. A third study reported “drug consumption” as an outcome, but this was not further specified. One study reported the proportion of patients
taking opiate at baseline, but this was not measured as an outcome.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Setting: Patients</th>
<th>Type of outcome</th>
<th>Tx arms</th>
<th>Baseline</th>
<th>Follow-up</th>
<th>Quality Score*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Howe 1983&lt;sup&gt;1&lt;/sup&gt;</td>
<td>General practice at one 2-physician practice in England</td>
<td>Pain (dichotomous at baseline, follow-up measured as number of patients showing improvement)</td>
<td>Azapropazone plus thrust manipulation and/or steroid/local anesthetic injection (only 2 patients received injections)</td>
<td>N = 26 19/26 (73%)</td>
<td>Immediate improvement: 13/19 (68%) 1 week improvement: 14/19 (74%) 3 week improvement: 13/17 (76%)</td>
<td>V1 = ?  V2 = +  V3 = +  V4 = -  V5 = -  V6 = -  V7 = ?  V8 = +  V9 = ?  V10 = +  V11 = +  Total = 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Azapropazone</td>
<td>N = 26 17/26 (65%)</td>
<td>Immediate improvement: 1/17 (6%) 1 week improvement: 9/15 (60%) 3 week improvement: 7/12 (58%)</td>
<td></td>
</tr>
<tr>
<td>Nordemar 1981&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Department of physical medicine and rehabilitation in Sweden</td>
<td>Visual Analogue Scale (VAS)</td>
<td>Non-thrust manual therapy and neck collar</td>
<td>N = 10 97 (46 SD)</td>
<td>1 week: 18 (25 SD)</td>
<td>V1 = ?  V2 = ?  V3 = ?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Neck collar</td>
<td>N = 10 90 (26 SD)</td>
<td>1 week: 35 (45 SD)</td>
<td>V4 = -  V5 = -  V6 = -  V7 = ?  V8 = +  V9 = +  V10 = +  V11 = +  Total = 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Transcutaneous nerve stimulation</td>
<td>N = 10 83 (26 SD)</td>
<td>1 week: 17 (19 SD)</td>
<td></td>
</tr>
<tr>
<td>Author, Year</td>
<td>Setting; Patients</td>
<td>Type of outcome</td>
<td>Tx arms</td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Quality Score*</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>---------------</td>
</tr>
<tr>
<td>Pikula 1999&lt;sup&gt;83&lt;/sup&gt;</td>
<td>Private chiropractic office in Canada</td>
<td>VAS</td>
<td>Thrust spinal manipulation, ipsilateral</td>
<td>N = 12 42.5 (19.8 SD)</td>
<td>Immediate: 23.6 (18.6 SD)</td>
<td>V1 = + V2 = + V3 = ? V4 = - V5 = - V6 = - V7 = + V8 = + V9 = ? V10 = + V11 = + Total = 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Thrust spinal manipulation, contralateral</td>
<td>N = 12 44.1 (27.5 SD)</td>
<td>Immediate: 41.4 (28.4 SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Detuned ultrasound</td>
<td>N = 12 50.4 (22.5 SD)</td>
<td>Immediate: 46.5 (21.8 SD)</td>
<td></td>
</tr>
<tr>
<td>Gonzalez-Iglesias 2009&lt;sup&gt;85&lt;/sup&gt;</td>
<td>Physical therapy clinic in Spain</td>
<td>Numerical pain rate scale</td>
<td>Transcutaneous nerve stimulation and thermotherapy</td>
<td>N = 22 5.37 (0.6 SD)</td>
<td>1 week: 4.3 (0.8 SD)</td>
<td>V1 = + V2 = + V3 = + V4 = - V5 = - V6 = - V7 = ? V8 = + V9 = + V10 = + V11 = + Total = 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Thoracic spine thrust manipulation</td>
<td>N = 23 5.6 (0.9SD)</td>
<td>1 week: 2.3 (1 SD)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Northwick Park Neck Pain Questionnaire</td>
<td>Transcutaneous nerve stimulation and thermotherapy</td>
<td>N = 2227.1 (2.7 SD)</td>
<td>1 week: 22.9 (2.9 SD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Thoracic spine thrust manipulation</td>
<td>N = 23 27.8 (3.1 SD)</td>
<td>1 week: 15.2 (4.1 SD)</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Setting; Patients</td>
<td>Type of outcome</td>
<td>Tx arms</td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Quality Score*</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------</td>
<td>-------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Gonzalez- Iglesias 2009&lt;sup&gt;84&lt;/sup&gt;</td>
<td>Physical therapy clinic in Spain</td>
<td>VAS</td>
<td>Transcutaneous nerve stimulation and thermotherapy with thrust thoracic spine manipulation</td>
<td>N = 23 54.7 (8.2 SD)</td>
<td>Immediate: 20.2 (7.8 SD) 2 week: 26.4 (11.8 SD)</td>
<td>V1 = + V2 = + V3 = + V4 = - V5 = - V6 = - V7 = ? V8 = ?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Transcutaneous nerve stimulation and thermotherapy</td>
<td>N = 22 52.7 (5.5 SD)</td>
<td>Immediate: 44.7 (5.5 SD) 2 week: 41.2 (6.1 SD)</td>
<td></td>
</tr>
<tr>
<td>Northwick Park</td>
<td>Neck Pain Questionnaire</td>
<td>Transcutaneous</td>
<td>Transcutaneous nerve stimulation and thermotherapy with thrust thoracic spine manipulation</td>
<td>N = 23 27.9 (3.0 SD)</td>
<td>Immediate: 15.2 (3.9 SD) 2 week: 14.7 (2.8 SD)</td>
<td>V9 = + V10 = + V11 = + Total = 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nerve stimulation and thermotherapy</td>
<td>Transcutaneous nerve stimulation and thermotherapy</td>
<td>N = 22 27.0 (3.1 SD)</td>
<td>Immediate: 23.1 (3.2 SD) 21.8 (3.3 SD)</td>
<td></td>
</tr>
</tbody>
</table>

*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.

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


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


86. Puente-dua 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
