Clinical Question

Does manipulation or mobilization of the cervical spine result in improved treatment outcomes compared with physical therapy, physiotherapy, or exercise in patients with neck pain?

1. Data Extraction

Each retrieved citation was reviewed by two independently working reviewers (D.F. and A.S.). Most articles were excluded on the basis of information provided by the title or abstract. Citations that appeared to be appropriate or those that could not be excluded unequivocally from the title and abstract were identified, and the corresponding full-text reports were reviewed by the two reviewers. Any disagreement between them was resolved by reviewer consensus. From the included articles, the following data were extracted: patient demographics, study population characteristics, intervention and control group procedures, outcomes measured, and results.

2. Study Quality

Determination of the class of evidence (CoE) provides the basis for critical appraisal of included studies and potential risk of bias in individual studies. The methods used for assessing the quality of evidence of individual studies as well as the overall quality of evidence incorporate aspects of rating scheme developed by the Oxford Centre for Evidence-based Medicine,[1] precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group,[2] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ)[3] including more recent guidance from the AHRQ on critical appraisal of studies for risk of bias.[4] This system accounts for features of methodological quality and important sources of bias, combining epidemiologic principles with characteristics of study design. Each individual study was rated by two different investigators against preset criteria that resulted in an evidence rating (CoE I, II, III, or IV). Disagreements were resolved through discussion.

3. Summary of Inclusion and Exclusion Criteria

Study component	Inclusion	Exclusion
Participants	• Patients with neck pain	• Cervical radiculopathy diagnosis • Cervical spinal stenosis diagnosis • Diagnosis of cervical myelopathic condition • Cervical disk herniation • Cervical postsurgical pain • History of cervical vertebral fracture(s) • History of tumor to cervical spine • Headache etiology of neck pain
Intervention(s)	• Cervical spinal manipulation therapy • Cervical spinal mobilization (manual therapy)	• Spinal manipulation directed at the thoracic spine only (i.e., thoracic thrust manipulation) • Multimodal therapy
Comparators	• Physical therapy (PT) • Physiotherapy/exercise • Feldenkrais method • Home exercises/mobilization	• Acupuncture • Electrical stimulation, including transcutaneous electrical nerve stimulation (TENS) • Treatment with injections • Surgical correction • Massage • Behavioral therapy • No treatment
Outcomes	• Pain reduction • Decreased disability • Symptom-free time • Time/procedure length until improvement • Improved quality of life • Complications of treatment • Costs of treatment	
Study design	• Randomized controlled trials or high-quality cohort studies (CoE II or higher)	• Studies with < 10 subjects • Low-quality studies (LoE III or lower)
Publication type	• Peer-reviewed studies published in English with abstracts	• White papers • Conference proceedings • Editorials, letters to editor • Preliminary or pilot studies • Multiple studies on the same patient population

4a: Critical Appraisal for Articles on Therapy

Methodological principle	Bronfort (2001)	Bronfort (2012)	Evans (2002)	Moretti (2004)	Hoving (2002)	Korthals-de Bos (2003)	Hoving (2006)
Study design							
Randomized controlled trial	✓	✓	✓	✓	✓	✓	✓
Cohort study							
Case series							
Random sequence generation[a]	✓	✓	✓		✓	✓	✓
Statement of concealed allocation[a]	✓	✓	✓		✓	✓	✓
Intention to treat[a]	[b]	✓	[b]		✓	✓	✓
Independent or blind assessment							
Co-interventions applied equally	✓	✓	✓	✓	✓	✓	✓
Complete follow-up of ≥80%	✓			✓	✓	✓	✓
Adequate sample size	✓	✓	✓	✓	✓	✓	✓
Controlling for possible confounding[c]	✓	✓	✓		✓	✓	✓
Evidence level	II	II	II	II	II	II	II

Note: Blank box indicates that the criterion was either not met or that it could not be determined.
[a]Applies to randomized controlled trials only.
[b]Primary outcome (pain) was assessed with the intent to treat for short-term outcome (11 wk), but not for long-term outcome (52, 104 wk).
[c]Groups must be comparable on baseline characteristics or evidence of control for confounding presented.

4b: Class of Evidence (CoE) Determination for Studies Comparing Treatments (Therapeutic Studies)

		Studies of therapy	
Class	Bias risk	Study design	Criteria
I	**Low risk:** Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias	Good-quality RCT	• Random sequence generation • Allocation concealment • Intent-to-treat analysis • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of >80% • Adequate sample size
II	**Moderately low risk:** Study has potential for some bias; study does not meet all criteria for class I, but deficiencies not likely to invalidate results or introduce significant bias	Moderate or poor-quality RCT	• Violation of one of the criteria for good-quality RCT
		Good-quality cohort	• Blind or independent assessment in a prospective study, or use of reliable data[a] in a retrospective study • Co-interventions applied equally • F/U rate of >80% • Adequate sample size • Controlling for possible confounding[b]

(*Continued*)

(*Continued*)

		Studies of therapy	
Class	**Bias risk**	**Study design**	**Criteria**
III	**Moderately High risk:** Study has significant flaws in design and/or execution that increase potential for bias that may invalidate study results	Moderate or poor-quality cohort	• Violation of any of the criteria for good-quality cohort
		Case control	• Any case–control design
IV	**High risk:** Study has significant potential for bias; lack of comparison group precludes direct assessment of important outcomes	Case series	• Any case series design

Abbreviations: F/U: follow up; RCT: randomized controlled trial.
[a]Outcome assessment is independent of healthcare personnel judgment. Reliable data are data such as mortality or re-operation.
[b]Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

5. Determination of Overall Strength of Evidence

After individual article evaluation, the overall body of evidence with respect to each outcome is determined based on precepts outlined by the GRADE Working Group[2] and recommendations made by the AHRQ.[3] Qualitative analysis is performed considering the following AHRQ required and additional domains.[5]

- *Risk of bias* is evaluated during the individual study evaluation described above. After individual article review, the literature evidence was rated as "HIGH" initially if the majority of the articles are Class I or II. It is rated as "LOW" if the majority were Class III or lower. This is the "baseline" strength of evidence, online supplementary "4a: Critical Appraisal for Articles on Therapy". The consistency, directness, precision, and subgroup effects are considered for potential "downgrading" the strength of the body of evidence (one or two levels depending on the degree and number of domain violations).

Criteria evaluated for "downgrading"

- *Consistency* refers to the degree of similarity in the effect sizes of different studies within an evidence base. If effect sizes indicate the same direction of effect and if the range of effect sizes is narrow, an evidence base was judged to be consistent. Single study evidence bases were judged "consistency unknown (single study)" and downgraded.
- *Directness* is concerned with whether the evidence being assessed reflected a single, direct link between the interventions of interest and the ultimate health outcome; that is, a determination of whether the most clinically relevant outcome was measured or if a surrogate outcome was assessed. Directness also applies to indirect comparisons of treatment when head-to-head comparisons of interest could not be made within individual studies.
- *Precision* of evidence pertains to the degree of certainty surrounding an estimate of effect for a specific outcome. This is based on whether the estimate of effect reached statistical significance and/or the inspection of confidence intervals around effect estimates. When there are only two subgroups, the overlap of the confidence intervals of the summary estimates of the two groups is considered. No overlap of the confidence intervals indicates statistical significance, but the confidence intervals can overlap to a small degree and the difference still is statistically significant.
- *Subgroup effects*. For evaluating subgroup effects (i.e., heterogeneity of treatment effects), we downgrade if the authors do not state a priori their plan to perform subgroup analyses and if there was no test for interaction.

Criteria used for "upgrading"

- Finally, if the strength of evidence is less than "HIGH," we "upgrade" the evidence if there is a dose–response association or a strong magnitude of effect.

The following four possible levels and their definition are reported:

- High—High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- Moderate—Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- Low—Low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and likely to change the estimate.
- Insufficient—Evidence either is unavailable or does not permit a conclusion.

Outcome	Strength of evidence	Conclusions and comments	Baseline	Downgrade	Upgrade
Spinal manipulation therapy vs. exercise					
Pain	Acute: LOW Chronic: LOW	• Acute: No short- or long-term pain improvement differences in manipulation therapy compared with home exercise treatment groups were reported in one study • Chronic: No short-term pain improvement differences were found in manipulation therapy vs. intense exercise treatment groups, though a long-term pain improvement was associated with exercise in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Disability	Acute: LOW Chronic: LOW	• Acute: No disability improvement was reported in manipulation therapy compared with home exercise in one study • Chronic: No disability improvement was reported in manipulation therapy compared with home exercise in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Treatment improvement	Acute: LOW Chronic: LOW	• Acute: No short- or long-term treatment improvement between mobilization therapy and home exercise groups were found in one study • Chronic: No short- or long-term treatment improvement differences between mobilization therapy and home exercise groups were found in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Treatment satisfaction	Acute: LOW Chronic: LOW	• Acute: Short- and long-term treatment satisfaction was associated with manipulation therapy compared with home exercise in one study • Chronic: No differences in treatment satisfaction were found between mobilization therapy and home exercise groups in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Health status	Acute: LOW Chronic: LOW	• Acute: No physical or mental health status change between manipulation therapy and exercise groups was found in one study • Chronic: No health status improvement was reported in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Functional improvement	Acute: LOW Chronic: LOW	• Acute: No short-term functional improvement differences in flexion/extension, rotation, or lateral flexion range of motion were found in manipulation therapy vs. home exercise groups in one study • Chronic: Short-term improvement in extension strength, but not flexion or rotation strength, and an improvement in flexion/extension range of motion, but not rotation or lateral flexion range of motion, were found in subjects who underwent exercise compared with mobilization therapy in one study	Acute: HIGH Chronic: HIGH	YES (2) consistency unknown, imprecise YES (2) consistency unknown, imprecise	NO NO
Mobilization therapy vs. physical therapy					
Pain	Acute: LOW Acute: LOW	• Acute: Short-term pain improvement was associated with mobilization therapy, compared with physical therapy, in one study, and there were no differences between groups in another study	Acute: HIGH Acute: HIGH	YES (2) inconsistent, imprecise YES (2) consistency unknown, imprecise	NO

(*Continued*)

(*Continued*)

Outcome	Strength of evidence	Conclusions and comments	Baseline	Downgrade	Upgrade
		• Acute: Long-term pain improvement was associated with physical therapy, compared with mobilization therapy, in one study and was not reported in another study			
Disability	Acute: LOW	• Acute: No disability improvement was reported in mobilization therapy compared with physical therapy in one study	Acute: HIGH	YES (2) consistency unknown, imprecise	NO
Treatment improvement	Acute: LOW	• Acute: Short-term perceived treatment recovery was associated with mobilization therapy, compared with physical therapy, in one study	Acute: HIGH	YES (2) consistency unknown, imprecise	NO
Health status	Acute: LOW	• Acute: Short-term health status improvement was associated with mobilization therapy, compared with physical therapy, in one study. No long-term utility (quality of life) improvement between groups was found in another study	Acute: HIGH	YES (2) consistency unknown, imprecise	NO
Functional improvement	Acute: MODERATE Acute: LOW	• Acute: No short-term functional improvement differences in flexion/extension, rotation, or lateral flexion range of motion were found in manipulation therapy vs. home exercise groups in two studies • Acute: No long-term functional improvement differences in flexion/extension, rotation, or lateral flexion range of motion were found in manipulation therapy vs. home exercise groups in one study	Acute: HIGH Acute: HIGH	YES (1) imprecise YES (2) consistency unknown, imprecise	NO NO

Note: All AHRQ "required" and "additional" domains[a] are assessed. Only those that influence the baseline grade are listed in the table.
Baseline strength: Risk of bias (including control of confounding) is accounted for in the individual article evaluations. HIGH = majority of articles Level I/II. LOW = majority of articles Level III/IV.
Downgrade: Inconsistency[b] of results (1 or 2); indirectness of evidence (1 or 2); imprecision of effect estimates (1 or 2); subgroup analyses not stated a priori and no test for interaction (2).
Upgrade: Large magnitude of effect (1 or 2); dose–response gradient (1).
[a]Required domains: risk of bias, consistency, directness, precision. Plausible confounding that would decrease observed effect is accounted for in our baseline risk of bias assessment through individual article evaluation. Additional domains: dose–response, strength of association, publication bias.
[b]Single study = "consistency unknown."

6. Excluded Articles

Author	Year	Reason for exclusion
Blunt KL, Rajwani MH, et al. The effectiveness of chiropractic management of fibromyalgia patients: a pilot study. J Manipulative Physiol Ther 1997; 20(6):389–399	1997	Study population and comparison groups did not meet inclusion criteria
Evans R, Bronfort G, et al. A pilot study for a randomized clinical trial assessing chiropractic care, medical care, and self-care education for acute and subacute neck pain patients. J Manipulative Physiol Ther 2003;26(7):403–411	2003	No between-group comparisons were performed due to the small sample size
Giles LG, Muller R. Chronic spinal pain syndromes: a clinical pilot trial comparing acupuncture, a nonsteroidal anti-inflammatory drug, and spinal manipulation. J Manipulative Physiol Ther 1999;22(6):376–381	1999	Study population did not meet inclusion criteria
Giles LG, Muller R. Chronic spinal pain: a randomized clinical trial comparing medication, acupuncture, and spinal manipulation. Spine (Phila Pa 1976) 2003;28(14):1490–502; discussion 1502–1503	2003	Study population did not meet inclusion criteria
Hemmila HM. Bone setting for prolonged neck pain: a randomized clinical trial. J Manipulative Physiol Ther 2005;28(7):508–515	2005	Comparison groups did not meet inclusion criteria
Howe DH, Newcombe RG, Wade MT. Manipulation of the cervical spine–a pilot study. J R Coll Gen Pract 1983;33(254):574–579	1983	Study population did not meet inclusion criteria
Jordan A, Bendix T, et al. Intensive training, physiotherapy, or manipulation for patients with chronic neck pain. A prospective, single-blinded, randomized clinical trial. Spine (Phila Pa 1976) 1998;23(3):311–318; discussion 319	1998	Study population did not meet inclusion criteria
Kanlayanaphotporn R, Chiradejnant A. et al. The immediate effects of mobilization technique on pain and range of motion in patients presenting with unilateral neck pain: a randomized controlled trial. Arch Phys Med Rehabil 2009;90(2):187–192	2009	Comparison groups did not meet inclusion criteria
Koes BW, Bouter LM, et al. A randomized clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: subgroup analysis and relationship between outcome measures. J Manipulative Physiol Ther 1993;16(4): 211–219	1993	Study population and comparison groups did not meet inclusion criteria
Koes BW, Bouter LM, et al. A blinded randomized clinical trial of manual therapy and physiotherapy for chronic back and neck complaints: physical outcome measures. J Manipulative Physiol Ther 1992;15(1):16–23	1992	Study population and comparison groups did not meet inclusion criteria
Koes BW, Bouter LM, et al. The effectiveness of manual therapy, physiotherapy, and treatment by the general practitioner for nonspecific back and neck complaints. A randomized clinical trial. Spine 1992;17(1):28–35	1992	Study population and comparison groups did not meet inclusion criteria
Koes BW, Bouter LM, et al. Randomised clinical trial of manipulative therapy and physiotherapy for persistent back and neck complaints: results of one year follow up. BMJ 1992;304(6827):601–605	1992	Study population and comparison groups did not meet inclusion criteria
Martel J, Dugas C, et al. A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain. BMC Musculoskelet Disord 2011;12:41	2011	Comparison groups did not meet inclusion criteria
Mealy K, Brennan H, Fenelon GC. Early mobilization of acute whiplash injuries. Br Med J (Clin Res Ed) 1986;292(6521):656–657	1986	Comparison groups did not meet inclusion criteria
Muller R, Giles LG, Long-term follow-up of a randomized clinical trial assessing the efficacy of medication, acupuncture, and spinal manipulation for chronic mechanical spinal pain syndromes. J Manipulative Physiol Ther 2005;28(1): 3–11	2005	Study population and comparison groups did not meet inclusion criteria
Murphy B, Taylor HH, et al. The effect of spinal manipulation on the efficacy of a rehabilitation protocol for patients with chronic neck pain: a pilot study. J Manipulative Physiol Ther 2010;33(3):168–177	2010	Comparison groups did not meet inclusion criteria
Palmgren PJ, Sandstrom PJ, et al. Improvement after chiropractic care in cervicocephalic kinesthetic sensibility and subjective pain intensity in patients with nontraumatic chronic neck pain. J Manipulative Physiol Ther 2006; 29(2):100–106	2006	Comparison groups did not meet inclusion criteria

(*Continued*)

(*Continued*)

Author	Year	Reason for exclusion
Pool JJ, Ostelo RW, et al. Is a behavioral graded activity program more effective than manual therapy in patients with subacute neck pain? Results of a randomized clinical trial. Spine (Phila Pa 1976) 2010;35(10):1017–1024	2010	Comparison groups did not meet inclusion criteria
Rogers RG. The effects of spinal manipulation on cervical kinesthesia in patients with chronic neck pain: a pilot study. J Manipulative Physiol Ther 1997;20(2):80–85	1997	Nonrandomized study, poor-quality cohort (CoE III)
Skargren EI, Carlsson PG, et al. One-year follow-up comparison of the cost and effectiveness of chiropractic and physiotherapy as primary management for back pain. Subgroup analysis, recurrence, and additional health care utilization. Spine (Phila Pa 1976) 1998;23(17):1875–1883; discussion 1884	1998	Study population did not meet inclusion criteria
Skargren EI, Oberg BE, et al. Cost and effectiveness analysis of chiropractic and physiotherapy treatment for low back and neck pain. Six-month follow-up. Spine (Phila Pa 1976) 1997;22(18):2167–2177	1997	Study population did not meet inclusion criteria
Sloop PR, Smith DS, et al Manipulation for chronic neck pain. A double-blind controlled study. Spine (Phila Pa 1976) 1982;7(6):532–535	1982	Comparison groups did not meet inclusion criteria
Sterling M, Jull G, Wright A. Cervical mobilisation: concurrent effects on pain, sympathetic nervous system activity and motor activity. Man Ther 2001;6 (2):72–81	2001	Crossover trial; did not assess outcomes of interest
Williams NH, Wilkinson C, et al. Randomized osteopathic manipulation study (ROMANS): pragmatic trial for spinal pain in primary care. Fam Pract 2003; 20(6):662–669	2003	Study population did not meet inclusion criteria
Ylinen J, Kautiainen H, et al. Stretching exercises vs. manual therapy in treatment of chronic neck pain: a randomized, controlled cross-over trial. J Rehabil Med 2007;39(2):126–132	2007	Poor-quality (CoE III) crossover trial

References

1 Phillips BBall CSackett D, et al. Levels of evidence and grades of recommendation2001. Available at: http://www.cebm.net/levels_of_evidence.asp. Accessed December 2, 2006

2 Atkins DBest DBriss PA, et al; and GRADE Working Group-Grading quality of evidence and strength of recommendationsBMJ 2004; 328(7454):1490

3 West SKing VCarey TS, et al. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment No. 47 (prepared by the Research Triangle Institute-University of North Carolina Evidence-Based Practice Center, Contract No. 290-97-0011)Rockville, MDAgency for Healthcare Research and Quality2002

4 Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EFRockville, MD; April 2012. Available at: www.effectivehealthcare.ahrq.gov

5 Owens DKLohr KNAtkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—agency for healthcare research and quality and the effective health-care programJ Clin Epidemiol 2010;63(5): 513–523