

Does spinal manipulative therapy help people with chronic low back pain?

Manuela L Ferreira, Paulo H Ferreira, Jane Latimer, Robert Herbert and Christopher G Maher

The University of Sydney

A systematic review of randomised clinical trials was conducted to assess the effect of spinal manipulative therapy on clinically relevant outcomes in patients with chronic low back pain. Databases searched included EMBASE, CINAHL, MEDLINE and PEDro. Methodological assessment of the trials was performed using the PEDro scale. Where there was sufficient homogeneity, a meta-analysis was conducted. Nine trials of mostly moderate quality were included in the review. Two trials were pooled comparing spinal manipulative therapy and placebo treatment, and two other trials were pooled comparing spinal manipulative therapy and non-steroidal anti-inflammatory drugs (NSAIDs). Spinal manipulative therapy reduced pain by 7mm on a 100mm visual analogue scale (95% CI 1 to 14) at one month follow-up when compared with placebo treatment, and by 14mm (95% CI -11 to 40) when compared with NSAIDs. Spinal manipulative therapy reduced disability by 6 points (95% CI 1 to 12) on a 100-point disability questionnaire when compared with NSAIDs. It is concluded that spinal manipulation does not produce clinically worthwhile decreases in pain compared with sham treatment, and does not produce clinically worthwhile reductions in disability compared with NSAIDs for patients with chronic low back pain. It is not clear whether spinal manipulation is more effective than NSAIDs in reducing pain of patients with chronic low back pain. [Ferreira ML, Ferreira PH, Latimer J, Herbert R and Maher CG (2002): Does spinal manipulative therapy help people with chronic low back pain? *Australian Journal of Physiotherapy* 48: 48: 277-284]

Key words: Low Back Pain; Manipulation, Orthopedic; Meta-Analysis; Spine

Introduction

It is believed that roughly 75-90% of patients with acute low back pain recover within six weeks (Waddell 1998). However, the other 10-25% are at risk of developing chronic low back pain, defined as pain and disability persisting for more than three months. These patients consume more than 80% of all health care for back trouble, and treatment for this population has a low success rate (Waddell 1998).

Patients with chronic low back pain present with impaired psychomotor functioning such as decreased speed of information processing and poor postural control (Luoto et al 1999 and 1996). In a comparative study, it has also been shown that these patients experience more frequent and severe pain and have poorer scores for physical and social functioning than non-chronic low back pain patients (Miedema et al 1998). These findings suggest that patients with chronic low back pain have adaptive changes related to long-term dysfunction. Therefore, when systematically reviewing treatments for low back pain, chronic low back pain should be considered separately from acute and sub-acute low back pain.

Spinal manipulative therapy has been widely used in the treatment of low back pain. This therapy includes high velocity, low amplitude ("thrust") joint manipulation, low velocity, small or large amplitude joint mobilisation,

manual traction and craniosacral therapy. However, the clinical effect of spinal manipulative therapy for chronic low back pain has not been firmly established. Previous systematic reviews of spinal manipulative therapy have produced inconsistent conclusions (Assendelft et al 1996, Koes et al 1996, Van Tulder et al 1997a). Van Tulder et al (1997a) reported results which support the effectiveness of spinal manipulative therapy for chronic low back pain, whereas Assendelft et al (1996) and Koes et al (1996) found no clear evidence of the effectiveness of this intervention for low back pain.

Most previous reviews have used a qualitative approach to synthesising trial findings (Furlan et al 2001). The most common qualitative methods in spinal manipulative therapy systematic reviews are the "hierarchical order" system, which draws conclusions based only on the best methodological quality trials, and the "levels of evidence" system where again the methodological quality of each trial is used to determine the strength of the evidence (strong, moderate or limited evidence; Furlan et al 2001). However, this approach considers only the number of trials favouring the experimental or control group to synthesise the findings, leaving the question of the size of the effect unanswered. Also, it has been shown recently that the agreement between similar qualitative approaches to classifying strength of evidence is poor (Ferreira et al 2002). Moreover, methods of synthesis based on counts of significant studies are extremely insensitive to true

treatment effects (Hedges and Olkin 1980).

We conducted a systematic review of randomised clinical trials of spinal manipulative therapy for low back pain persisting for more than three months to evaluate the level of support for this intervention. We attempted a quantitative synthesis of the findings of the studies of spinal manipulative therapy in the treatment of patients with chronic low back pain.

Methods

Criteria for considering trials for the review To be included, a study had to fulfil several criteria. Only full journal papers describing randomised or quasi-randomised controlled trials were accepted. There was no language restriction. Participants had to be adults with non-specific low back pain of more than three months duration as reported by the median duration of symptoms. The trials must have investigated one or more of the following types of spinal manipulative therapy: high velocity, low amplitude manipulation; low velocity, small or large amplitude joint mobilisation; manual traction; or craniosacral therapy. Trials must have reported at least one of the following outcome measures: disability; pain; quality of life; adverse events; return to work; global perceived effect; or patient satisfaction with therapy.

Identification and selection of studies Searches were conducted on MEDLINE, EMBASE, CINAHL, and PEDro up to March 2001. A combination of subject headings and text words related to the domains of randomised controlled trials and back pain (as described by the Cochrane Back Review Group), and variants of the words manual therapy, mobilisation, manipulation, Maitland, chiropractic, osteopathic, craniosacral, traction and passive movement were used as search terms (Van Tulder et al 1997b). A table with all the subject headings and text words is available from the authors. A manual search of reference lists of all previous systematic reviews assessing spinal manipulative therapy for low back pain was conducted. Two reviewers independently screened search results, selected the trials to be included and reviewed each article for inclusion. A third independent reviewer resolved disagreement. If any features of the selection criteria were not clearly described in the article, the article was withheld and the author written to for clarification. In situations where authors were not traceable or did not reply, a consensus of the investigators was invoked.

Assessment of validity The PEDro scale (Moseley et al 2002) was used to assess methodological quality of the trials. This scale, based on the Delphi list (Verhagen et al 1998), scores trials on the presence or absence of 10 methodological criteria (Moseley et al 2002). In the primary analysis, a PEDro score of less than three was used to exclude trials from analysis.

Table 1. Outcome measures in eligible trials.

Outcome measure	Number of trials
Pain intensity on a VAS	8
Number of patients pain-free	1
Main complaint	1
Disability	5
Return to work	1
Physical functioning	1
Adverse events	0
Global perceived effects	1
Patient satisfaction with therapy	1

Two PEDro raters independently rated methodological quality of each trial. PEDro raters are trained and assessed to maximise the accuracy of the methodological score. The two scores are compared by a third rater. Where the two scores disagree, the third rater adjudicates. The reliability for consensus ratings for the total PEDro score is moderate (ICC (1,1) = 0.68 (95% CI 0.57 to 0.76); unpublished data).

Analysis of data Independent reviewers extracted and analysed continuous and dichotomous outcome data. Where there were inconsistencies, a consensus was invoked.

Measures of effect in individual trials To determine the treatment effect for each outcome we calculated the mean and 95% confidence interval for the between-group differences (Herbert 2000). Between-group differences in endpoints or within-group change scores were used according to the data provided by each trial (Green et al 2001). For two trials that did not provide standard deviations, an estimate of that value was calculated. One trial provided the range, and the standard deviation was estimated as one quarter of the range (Gibson et al 1985). Another trial provided interquartile values and, in this case, the standard deviation was estimated as three-quarters of the interquartile range (Giles and Muller 1999). Pain reduction of 20mm or more on a 100mm visual analogue scale (Farrar et al 2001) and disability reduction of 30 points or more on a 100 point disability scale were considered clinically worthwhile values.

Relative risks were calculated for dichotomous data (Oxman 1994). A relative risk of 0.70 or less was considered clinically significant, where values of less than 1.0 favour manipulation.

Pooled analysis Trials were grouped according to the type of intervention, outcome measures and follow-up time. Where there were not multiple trials with sufficient homogeneity, the effect size of the individual trials was reported. Pooling was carried out using a random effects model to estimate pooled effect sizes (Fleiss 1993). The

Table 2. Effect of spinal manipulative therapy on outcomes measured on a continuous scale.

Outcome measure (reference – time of follow-up)	Sample size (total)	Methodological score (PEDro score/10)	Mean change difference [95% CI]
SMT vs placebo: pain (0-100mm VAS)			
Gibson et al 1985 - 2 weeks	69	4	-3 [-14 to 8]
Gibson et al 1985 - 4 weeks	68	4	-6 [-16 to 4]
Gibson et al 1985 - 12 weeks	51	4	7 [-6 to 20]
Triano et al 1995 - 2 weeks	86	6	-6 [-13 to 1]
Triano et al 1995 - 4 weeks	86	6	-8 [-17 to 0]
Waagen et al 1986 - 2 weeks	19	5	-2 [N/A]
SMT vs placebo: pain/physical impairment/disability (range 5–32 points)			
Postacchini et al 1988 - 3 weeks	95	3	-2 [N/A]
Postacchini et al 1988 - 2 months	95	3	-2 [N/A]
Postacchini et al 1988 - 6 months	95	3	-3 [N/A]
SMT vs placebo: disability (Oswestry 100 point scale)			
Triano et al 1995 - 2 weeks	79	6	-6 [-10 to -2]*
Triano et al 1995 - 4 weeks	79	6	-3 [-9 to 2]
SMT vs SWD: pain (0-100mm VAS)			
Gibson et al 1985 - 2 weeks	68	4	-10 [-21 to 1]
Gibson et al 1985 - 4 weeks	67	4	-7 [-18 to 4]
Gibson et al 1985 - 12 weeks	53	4	-12 [-25 to 1]
SMT vs acupuncture: pain (0-100mm VAS)			
Giles et al 1999 - 4 weeks	50	4	-33 [-51 to -15]*
SMT vs acupuncture: disability (Oswestry 100 point scale)			
Giles et al 1999 - 4 weeks	48	4	-10 [-17 to -2]*
SMT vs back school: pain (0-100mm VAS)			
Triano et al 1995 - 2 weeks	90	6	-6 [-6 to 1]
Triano et al 1995 - 4 weeks	90	6	-2 [-9 to 6]
SMT vs back school: pain/physical impairment/disability (range 5–32 points)			
Postacchini et al 1988 - 3 weeks	102	3	-2 [N/A]
Postacchini et al 1988 - 2 months	102	3	0 [N/A]
Postacchini et al 1988 - 6 months	102	3	-4 [N/A]
SMT vs back school: disability (Oswestry 100 point scale)			
Triano et al 1995 - 2 weeks	77	6	-3 [-6 to 1]
Triano et al 1995 - 4 weeks	77	6	-1 [-6 to 4]
Manual therapy vs NSAIDs: pain (0-100 mm)			
Bronfort et al 1996 - 5 weeks	105	7	-2 [-10 to 6]
Bronfort et al 1996 - 11 weeks	96	7	-8 [-17 to 1]
Giles et al 1999 - 4 weeks	52	4	-28 [-46 to -10]*
Manual therapy vs NSAIDs: pain/physical impairment/disability (range 5–32 points)			
Postacchini et al 1988 - 3 weeks	99	3	1 [N/A]
Postacchini et al 1988 - 2 months	99	3	1 [N/A]
Postacchini et al 1988 - 6 months	99	3	0 [N/A]
Manual therapy vs NSAIDs: disability (Roland Morris-Oswestry 100 point scale)			
Bronfort et al 1996 - 5 weeks	105	7	-2 [-9 to 6]
Bronfort et al 1996 - 11 weeks	96	7	-6 [-13 to 1]
Giles et al 1999 - 4 weeks	52	4	-9 [-14 to -4]*

Continued over

Continued from previous page

Outcome measure (reference – time of follow-up)	Sample size (total)	Methodological score (PEDro score/10)	Mean change difference [95% CI]
Manual therapy vs NSAIDs: general health status (COOP charts: 0-100 scale)			
Bronfort et al 1996 - 5 weeks	105	7	-2 [-8 to 3]
Bronfort et al 1996 - 11 weeks	96	7	0 [-5 to 5]
Manual therapy vs physiotherapy: main complaint (10 point-scale)			
Koes et al 1992 - 12 weeks	58	8	1 [N/A]
Koes et al 1992 - 12 months	58	8	2 [N/A]
Manual therapy vs physiotherapy: pain/physical impairment/disability (range 5–32 points)			
Postacchini et al 1988 - 3 weeks	99	3	2 [N/A]
Postacchini et al 1988 - 2 months	99	3	2 [N/A]
Postacchini et al 1988 - 6 months	99	3	2 [N/A]
Manipulation vs no treatment: pain (daily score on a 4-point scale, during three weeks)			
Evans et al 1978 - 3 weeks	22	6	-1 [N/A]

Mean differences between groups are given for each trial. Effect size and confidence intervals were calculated where there were enough data. * Statistically significant. N/A – data not available. Negative estimates favour manipulation. Sample sizes are those at time of outcome measure. SMT, spinal manipulative therapy. SWD, short wave diathermy. NSAIDs, non-steroidal anti-inflammatory drugs.

random effects model assumes that the studies are a random sample from a larger population of studies, there being a mean population effect size about which study-specific effect sizes vary. Even if each study's results were based on very large sample sizes, with standard error being zero, there would be study-to-study variation because each study would have its own effect size (Fleiss 1993).

Calculation of the pooled estimate requires estimates of the variance of between-group differences in means, but some trials did not supply these data. To determine the effect of excluding trials that did not provide enough data to estimate standard deviations, another analysis was performed using the sample sizes as weights (Hunter and Smith 1990).

Results

Twelve of the 117 identified papers met the inclusion criteria (Bronfort et al 1996, Evans et al 1978, Gibson et al 1985, Giles and Muller 1999, Koes et al 1992a, 1992b and 1992c, Koes et al 1993, Postacchini et al 1988, Rupert et al 1985, Triano et al 1995, Waagen et al 1986). The CINAHL database yielded three of the 12 eligible papers; EMBASE yielded eight, MEDLINE yielded nine and PEDro yielded all 12 eligible papers. Four papers published data from the same trial (Koes et al 1992a, 1992b and 1992c, Koes et al 1993), giving a total of nine trials. The included trials assessed manipulation (Gibson et al 1985, Postacchini et al 1988, Rupert et al 1985, Triano et al 1995, Waagen et al 1986), and a combination of manipulation and mobilisation techniques (Bronfort et al 1996, Evans et al 1978, Giles and Muller 1999, Koes et al 1992a, 1992b and 1992c, Koes et

al 1993). The mean (SD) duration of pain of subjects admitted to these trials was 28.1 months (24.0). Details of all included randomised clinical trials are presented in the Appendix. The study by Rupert et al (1985) is not in the Appendix. Data from that study were not analysed because it had a low methodological score.

Assessment of outcome Table 1 shows the number of trials reporting relevant outcome measures according to our criteria. None of the included trials reported adverse events or quality of life measures. One trial combined measures of pain, physical impairment and disability in a single scale (Postacchini et al 1988). In another trial, most outcome measures were reported for subjects with back and neck pain together, and only some data concerning low back patients could be extracted and analysed (Koes et al 1992a, 1992b and 1992c, Koes et al 1993).

Methodological quality of individual trials One trial (Rupert et al 1985) scored 2 on the PEDro scale and thus did not reach the methodological quality score threshold of 3 or greater. Therefore, only eight trials were analysed in the review. Only five trials scored 5 or more on the methodological assessment scale. The most common flaws were lack of explicit subject and therapist blinding (nine trials each), failure to explicitly use an intention-to-treat analysis (eight trials) and failure to explicitly conceal allocation (seven trials). Blinding of subjects and therapists is difficult or impossible in trials of spinal manipulative therapy.

Treatment efficacy Effect sizes of individual trials are shown in Table 2. Confidence intervals could not be calculated for four trials, and thus only mean differences

Table 3. Effects of manipulation on dichotomous outcomes.

Comparison and outcome Trial - duration of follow-up	Relative Risk [95% Confidence Interval]
Manipulation vs NSAIDs: - Remaining off work Bronfort et al 1996 - 11 weeks	0.79 [0.36 to 1.77]
Spinal manipulation vs no treatment: patients' reports of no benefit from treatment Evans et al 1978 - 3 weeks	0.49 [0.25 to 0.94]*
Spinal manipulation vs shortwave diathermy: number of patients remaining in pain Gibson et al 1985 - 12 weeks	0.82 [0.57 to 1.19]
Spinal manipulation vs shortwave diathermy: number of patients remaining off work Gibson et al 1985 - 12 weeks	0.71 [0.11 to 4.74]
Spinal manipulation vs placebo: number of patients remaining in pain Gibson et al 1985 - 12 weeks	1.03 [0.68 to 1.55]
Spinal manipulation vs placebo: remaining off work Gibson et al 1985 - 12 weeks	0.28 [0.06 to 1.30]

Relative risks less than 1.0 favour manipulation. * Statistically significant ($p < 0.05$).

between groups are reported for these trials. Dichotomous data are shown in Table 3. Most results of individual trials could not be pooled due to heterogeneity of outcome measures or comparison groups, or insufficient reported data. In general, effect sizes were small and not significant for individual trials.

Pooling of effects was performed with two trials that compared manipulation with placebo treatment, and which reported pain on a visual analogue scale (0-100mm) at two and four weeks follow-up (Gibson et al 1985, Triano et al 1995). The pooled effect was to reduce pain by 5mm (95% CI: -1 to 11) on a 100mm visual analogue pain scale at two weeks ($n = 155$), and to reduce pain by 7mm (95% CI: 1 to 14) at four weeks follow-up ($n = 154$).

Adding a third trial (Waagen et al 1986) and using the sample sizes as weights resulted in a 1mm decrease in the effect on pain at two weeks follow-up (ie pooled estimate of 4mm on a 100mm visual analogue pain scale, $n = 174$). A confidence interval could not be obtained using this method.

We also pooled the effects of two trials comparing spinal manipulative therapy with NSAIDs at four weeks follow-up (Bronfort et al 1996, Giles and Muller 1999). The pooled estimate of the effect was that spinal manipulative therapy reduced pain by 14mm (-11 to 40) on a 100mm visual analogue pain scale ($n = 156$) and reduced disability by 6 points (-1 to 12) on a 100-point disability scale ($n = 157$).

Discussion

This is the first systematic review since 1992 to quantitatively synthesise the findings of trials of spinal manipulative therapy for chronic low back pain. The results suggest that spinal manipulative therapy does not produce a clinically significant reduction in pain when compared with sham treatment, nor a significant improvement in disability when compared with NSAIDs in patients with chronic low back pain. Recent reviews have shown that there are other evidence-based physiotherapy management options for chronic low back patients, including exercise, back school, multidisciplinary rehabilitation, and cognitive behavioural treatment (Van Tulder et al 2001a, 2001b and 2001c). Of course, in contemporary physiotherapy management, some of these options may be provided in conjunction with spinal manipulative therapy.

The results also suggest that spinal manipulative therapy is not more effective than NSAIDs in reducing chronic low back pain. However this is less conclusive, since the most optimistic confidence limit for the treatment effect is clinically worthwhile (40mm reduction on a 100mm pain scale).

Considering that patients with chronic low back pain typically have higher levels of pain and disability than patients with acute low back pain (Miedema et al 1998), the estimated effects of most individual trials seem too small to be clinically significant. However there are two trials that report large effect sizes (Evans et al 1978, Giles and Muller

1999). The trial by Giles and Muller (1999) is a low quality study that found a large reduction in pain with spinal manipulative therapy compared with acupuncture. This finding may simply reflect the fact that low quality trials are associated with an increased estimated effect (Moher et al 1998). In the trial by Evans et al (1978), more patients in the spinal manipulative therapy group believed they benefited from treatment, but it is unclear what the benefit was, as the pain results suggest that the spinal manipulative therapy group experienced almost no change in pain compared with the no treatment group (Table 2).

Furlan et al (2001) conducted a critical review of earlier systematic reviews in chronic low back pain, identifying nine reviews in spinal manipulative therapy. Eight performed qualitative analysis of the trials, the most common approach being the popular "levels of evidence" approach, in which the strength of evidence is classified according to the number and quality of trials. One of the problems with the levels of evidence approach is that varying definitions of "levels of evidence" generate different conclusions on treatment efficacy (Ferreira et al 2002).

In our review, we present the size of effect for each trial. Although heterogeneity of the trials has prevented pooling of some of the results, we believe presenting the size of the effect provides more information than qualitatively analysing the trials. In the future, improving methodological quality of the trials is essential if one wishes to draw more conclusions on the efficacy of spinal manipulative therapy for chronic low back pain.

Conclusion

Spinal manipulative therapy is not substantially more effective than sham treatment in reducing pain, nor is it more effective than NSAIDs in improving disability of patients with chronic low back pain. It is not clear whether spinal manipulative therapy is more effective than NSAIDs in reducing pain in chronic low back pain patients.

Correspondence Manuela L Ferreira, School of Physiotherapy, The University of Sydney, PO Box 170, Lidcombe New South Wales 1825. E-mail: mfer5142@mail.usyd.edu.au.

References:

Assendelft WJJ, Koes BW, Van der Heijden GJ and Bouter LM (1996): The effectiveness of chiropractic for treatment of low back pain: an update and attempt at statistical pooling. *Journal of Manipulative and Physiological Therapeutics* 19: 499-506.

Bronfort G, Goldsmith CH, Nelson CF, Boline PD and Anderson AV (1996): Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomised, observer-blinded clinical trial. *Journal of Manipulative and Physiological Therapeutics* 19: 570-8282.

Evans DP, Burke MS, Lloyd KN, Roberts EE and Roberts GM (1978): Lumbar spinal manipulation on trial. Part I - clinical assessment. *Rheumatology and Rehabilitation* 17: 46-53.

Farrar JT, Young JP, LaMoreaux L, Werth JL and Poole RM (2001): Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 94: 149-158.

Ferreira PH, Ferreira ML, Maher C, Latimer J, Herbert R and Refshauge K (2002): Effect of applying different "levels of evidence" criteria on conclusions of Cochrane reviews of interventions for low back pain. *Journal of Clinical Epidemiology* 55: 1111-1114.

Fleiss JL (1993): The statistical basis of meta-analysis. *Statistical Methods in Medical Research* 2: 121-145.

Furlan AD, Clarke J, Esmail R, Sinclair S and Irvin E (2001): A critical review of reviews on the treatment of chronic low back pain. *Spine* 26: E155-E163.

Gibson T, Grahame R, Harkness J, Woo P, Blagrove P and Hills R (1985): Controlled comparison of short-wave diathermy treatment with osteopathic treatment in non-specific low back pain. *Lancet* 1: 1258-1261.

Giles LGF and Muller R (1999): Chronic spinal pain syndromes: a clinical pilot trial comparing acupuncture, a non-steroidal anti-inflammatory drug, and spinal manipulation. *Journal of Manipulative and Physiological Therapeutics* 22: 376-381.

Green S, Deeks J and Savio F (2001): Meta-analysis of continuous data: does the end justify the mean? Proceedings of the 9th International Cochrane Colloquium, Oct 9-13; Lyon, France.

Hedges LV and Olkin I (1980): Vote-counting methods in research synthesis. *Psychological Bulletin* 88: 359-369.

Herbert RD (2000): How to estimate treatment effects from reports of clinical trials: I: continuous outcomes. *Australian Journal of Physiotherapy* 46: 229-235.

Hunter JE and Schmidt FL (1990): Methods of integrating findings across studies. In Hunter JE and Schmidt FL (Eds): *Methods of meta-analysis - correcting error and bias in research findings*. London: Sage Publications, pp. 468-489.

Koes BW, Assendelft WJJ, van der Heijden and Bouter LM (1996): Spinal manipulation for low back pain - an updated systematic review of randomised clinical trials. *Spine* 21: 2860-2873.

Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GJ, Hofhuizen DM, Houben JP and Knipschild PG (1992a): A randomised clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: results of one-year follow-up. *BMJ* 304: 601-605.

Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GJ, Hofhuizen DM, Houben JP and Knipschild PG (1992b): A blinded randomised clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: physical outcome measures. *Journal of Manipulative and Physiological Therapeutics* 15: 16-23.

Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GJ, Hofhuizen DM, Houben JP and Knipschild PG (1992c): The effectiveness of manual therapy,

- physiotherapy, and treatment by the general practitioner for non-specific back and neck complaints. *Spine* 17: 28-36.
- Koes BW, Bouter LM, van Mameren H, Essers AH, Verstegen GJ, Hofhuizen DM, Houben JP and Knipschild PG (1993): A randomised clinical trial of manual therapy and physiotherapy for persistent back and neck complaints: subgroup analysis and relationship between outcome measures. *Journal of Manipulative and Physiological Therapeutics* 16: 211-219.
- Luoto S, Taimela S, Hurri H, Aalto H, Pyykkö I and Alaranta H (1996): Psychomotor speed and postural control in chronic low back pain patients. *Spine* 21: 2621-2627.
- Luoto S, Taimela S, Hurri H and Alaranta H (1999): Mechanisms explaining the association between low back trouble and deficits in information processing. *Spine* 24: 255-261.
- Miedema HS, Chorus AM, Wevers CWJ and van der Linden S (1998): Chronicity of back problems during working life. *Spine* 23: 2021-2028.
- Moher D, Pham B, Jones A, Cook DJ and Jadad A (1998): Does quality of reports of randomized trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 352: 609-613.
- Moseley AM, Herbert RD, Sherrington C, Maher CG (2002): Evidence for physiotherapy practice: A survey of the Physiotherapy Evidence Database (PEDro). *Australian Journal of Physiotherapy* 48: 43-49.
- Oxman A (1994): Preparing and maintaining systematic reviews. In *Cochrane Collaboration Handbook*. Oxford: The Cochrane Library.
- Postacchini F, Facchini M and Palieri P (1988): Efficacy of various forms of conservative treatment in low back pain. A comparative study. *Neuro-Orthopedics* 61: 28-35.
- Rupert RL, Wagon R, Thompson P and Ezzeldin MT (1985): Chiropractic adjustments: results of a controlled clinical trial in Egypt. *ICA International Review of Chiropractic Winter*: 58-60.
- Triano JJ, McGregor M, Hondras MA and Brennan PC (1995): Manipulative therapy versus education programs in chronic low back pain. *Spine* 20: 948-955.
- Van Tulder MW, Koes BW and Bouter LM (1997a): Conservative treatment of acute and chronic non-specific low back pain. *Spine* 22: 2128-2136.
- Van Tulder M, Assendelft WJJ, Koes BW and Bouter LM (1997b): Method guidelines for systematic reviews in the Cochrane Collaboration back review group for spinal disorders. *Spine* 22: 2323-2330.
- Van Tulder MW, Esmail R, Bombardier C and Koes BW (2001a): Back schools for non-specific low back pain. The Cochrane Library, Issue 4. Oxford: Update Software.
- Van Tulder MW, Malmivaara A, Esmail R and Koes BW (2001b): Exercise therapy for low back pain. The Cochrane Library, Issue 4. Oxford: Update Software.
- Van Tulder MW, Ostelo RWJG, Vlaeyen JWS, Linton SJ, Morley SJ and Assendelft WJJ (2001c): Behavioural treatment for chronic low back pain. The Cochrane Library, Issue 4. Oxford: Update Software.
- Verhagen AP, de Vet HCW, de Bie RA, Kessels AG, Boers M, Bouter LM and Knipschild PG (1998): The Delphi list: a criteria list for quality assessment of randomised clinical trials for conducting systematic reviews developed by Delphi consensus. *Journal of Clinical Epidemiology* 51: 1235-1241.
- Waagen GN, Haldeman S, Cook G, Lopez D and DeBoer KF (1986): Short-term trial of chiropractic adjustments for the relief of chronic low back pain. *Manual Medicine* 2: 63-67.
- Waddell G (1998): *The Back Pain Revolution*. London: Churchill Livingstone, pp. 103-117, 351-367.

Appendix 1. Description of studies included in the review.

Reference	Spinal manipulative therapy	Reference treatment	Treatment period	Relevant outcome measures	Results
Bronfort et al 1996	1. Chiropractic and strengthening exercise	2. NSAID therapy and strengthening exercise	10 sessions during 5 weeks	pain, disability (Roland Morris)	No significant differences.
Evans et al 1978*	1. Rotational thrust bilaterally and NSAID therapy	2. NSAID therapy	3 weeks	pain, treatment effectiveness	No significant differences in pain. Significant number of patients assessed SMT as more effective.
Gibson et al 1985	1. Osteopathy	2. SWD 3. Detuned SWD	4 weeks	pain, medication use, ability to work	No significant difference among groups.
Giles et al 1999	1. Chiropractic	2. Acupuncture 3. NSAIDs	6 sessions in 3-4 weeks	pain, disability (Oswestry)	Significant reductions in pain and disability in SMT group.
Koes et al 1992	1. Manipulation and mobilisation	2. Physiotherapy: exercise, massage, modalities 3. GP: medication, advice 4. Placebo: detuned SWD	Maximum of 3 months	severity of main complaint, global perceived effect, pain, functional status	Improvement of main complaint for 1 and 2 greater than for 3 and 4. Highest mean scores for global perceived effect for 1 and 2 (no difference between 1 and 2). No statistical difference for pain and functional status among groups.
Postacchini et al 1988	1. Chiropractic	2. NSAID therapy 3. Physiotherapy: light massage, modalities. 4. Placebo: anti-oedema gel 5. Back school	1. 12 sessions 2. 15 to 20 days 3. daily for 2 to 3 weeks 4. twice a day for 2 weeks 5. 2 months	pain, perceived ability to perform daily activities	No significant changes on combined outcome measures.
Triano et al 1995	1. Manipulation	2. Sham: high velocity, low force mimic 3. Back school	2 weeks	pain, disability (Oswestry)	Significant improvement in pain scores favouring manipulation group. No significant differences in disability among groups.
Waagen et al 1986	1. Chiropractic	2. Sham adjustment	4 to 6 sessions	pain	Pain significantly improved in Group 1.

* cross-over design. NSAID, non-steroidal anti-inflammatory drug. SWD, short-wave diathermy. SMT, spinal manipulation therapy. GP, general medical practitioner.