



Systematic review

The effectiveness of sub-group specific manual therapy for low back pain: A systematic review

Sarah L. Slater*, Jon J. Ford, Matthew C. Richards, Nicholas F. Taylor, Luke D. Surkitt, Andrew J. Hahne

Low Back Research Team, Musculoskeletal Research Centre, Department of Physiotherapy, La Trobe University, Bundoora, Victoria 3083, Australia

ARTICLE INFO

Article history:

Received 22 September 2011

Received in revised form

9 January 2012

Accepted 13 January 2012

Keywords:

Classification

Manual therapy

Low back pain

Sub-group

ABSTRACT

Background: Manual therapy is frequently used to treat low back pain (LBP), but evidence of its effectiveness is limited. One explanation may be sample heterogeneity and inadequate sub-grouping of participants in randomized controlled trials (RCTs) where manual therapy has not been targeted toward those likely to respond.

Objectives: To determine the effectiveness of specific manual therapy provided to sub-groups of participants identified as likely to respond to manual therapy.

Data sources: A systematic search of electronic databases of MEDLINE, EMBASE, CINAHL, and the Cochrane Central Register of Controlled trials (CENTRAL).

Trial eligibility criteria: RCTs on manual therapy for participants identified as belonging to a sub-group of LBP likely to respond to manual therapy were included.

Trial appraisal and synthesis methods: Identified trials were assessed for eligibility. Data from included trials were extracted by two authors independently. Risk of bias in each trial was assessed using the PEDro scale and the overall quality of evidence rated according to the GRADE domains. Treatment effect sizes and 95% confidence intervals were calculated for pain and activity.

Results: Seven RCTs were included in the review. Clinical and statistical heterogeneity precluded meta-analysis. Significant treatment effects were found favouring sub-group specific manual therapy over a number of comparison treatments for pain and activity at short and intermediate follow-up. However, the overall GRADE quality of evidence was very low.

Conclusions: This review found preliminary evidence supporting the effectiveness of sub-group specific manual therapy. Further high quality research on LBP sub-groups is required.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Low back pain (LBP) is a widespread problem experienced by up to two-thirds of the American population during their lifetime (Deyo et al., 2006). Contrary to previous reports of a favourable prognosis (Coste et al., 1994), 62% of people with first time LBP report persistent problems at 12 months (Hestbaek et al., 2003). Additional adverse impacts include increased healthcare costs (Waddell, 1996; Fritz et al., 2008), work absenteeism and reduced productivity (Dagenais et al., 2008) as well as individuals being excluded from work, family and social interactions (Kent & Keating, 2005a).

There are a wide variety of treatment options for LBP (Koes et al., 2006). Manual therapy is a commonly used clinical modality (Li &

Bombardier, 2001; Liddle et al., 2009) recommended in clinical guidelines (van Tulder et al., 2004; Chou et al., 2007; National Collaborating Centre for Primary Care, 2009). For the purposes of this review we defined manual therapy as a high velocity thrust manipulation or lower velocity mobilization directed at the vertebral articulations (Maitland, 2005). Systematic reviews evaluating the effectiveness of manual therapy for LBP have found little evidence of a clinically meaningful effect compared to other treatment options (Assendelft et al., 2004; Dagenais et al., 2010; Rubinstein et al., 2011), although there appears to be a small benefit when compared to placebo or no treatment (Ferreira et al., 2003; Assendelft et al., 2004).

Heterogeneity in randomized controlled trials (RCTs) investigating manual therapy may be an explanation for the limited evidence of effectiveness to date (Kent et al., 2005b). Sample heterogeneity is known to reduce the likelihood of a significant treatment effect due to the reduced numbers of participants in the sample for whom the treatment is appropriate (Delitto, 2005; Ford

* Corresponding author.

E-mail address: s Slater@internode.on.net (S.L. Slater).

et al., 2007). One way of overcoming this problem may be to recruit a homogenous sub-group of participants likely to respond to specific treatment. Such an approach replicates clinical practice where treatment for different sub-groups of LBP is determined by likely responsiveness (Kent & Keating, 2004; Maitland, 2005). Conducting research using sub-grouping principles has been recognized as a high research priority for LBP (Ford et al., 2007; Fritz et al., 2007).

To date no systematic review has evaluated the effectiveness of sub-group specific manual therapy for LBP. The aim of the current review was therefore to evaluate the effectiveness of specific manual therapy to reduce pain and increase activity in sub-groups of LBP participants identified as likely to respond to manual therapy.

2. Methods

2.1. Data sources and searches

A search for RCTs was conducted using a combination of the most recent Cochrane Back Review Group search strategies (Cochrane Back Review Group, 2008) and search terms identified by previous systematic reviews of manual therapy for LBP (Appendix 1). Databases searched included, MEDLINE (Ovid, 1950–October 2010), EMBASE (Ovid, 1988–October 2010), CINAHL (Ebsco, 1982–October 2010) and The Cochrane Central Register of Controlled Trials (CENTRAL) (to October 2010). Reference lists of included trials and previous systematic reviews were searched.

2.1.1. Trial selection

Selection of trials to be included in the review was completed independently by two reviewers (ST, MR) with disagreements resolved by consultation with a third reviewer (JF). Potentially relevant trials were screened against pre-determined eligibility criteria, initially based on title and abstract. Full text articles were assessed for all included trials and those where a definite decision on inclusion could not be made based on title and abstract. Only RCTs published in full by peer-reviewed journals written in English were included.

2.1.2. Participants

Trials involving participants aged 18–65 years, with LBP \pm leg pain attributable to mechanical causes were included in the review. Trials were excluded if LBP was due to serious or non-mechanical pathologies.

2.1.3. Interventions

Trials were included if at least one group received therapist-applied manual therapy including manipulation or mobilization. There is no “gold standard” method for developing LBP sub-group criteria given the wide variety of approaches used clinically and in the manual therapy literature to date (Ford et al., 2007; Kent et al., 2009). In order to identify the full diversity of sub-groups reported in this literature, the following eligibility criteria were applied. Firstly, trials needed to state clearly that participants recruited belonged to a LBP sub-group likely to respond to manual therapy. Secondly, trials where the sub-group was developed using statistical methods, (Riddle, 1998; Ford et al., 2007) such as clinical prediction rules, were only included if the treatment being investigated was consistent with the descriptions in the original developmental studies. Thirdly, in trials where judgmental methods (Riddle, 1998; Ford et al., 2007) were used to develop the sub-group criteria, a clear relationship between treatment effect and the nature of the sub-group had to be described. For example, a trial where authors describe a hypothesized causal relationship between an “opening” lumbar rotation technique and improvement in zygapophyseal joint pain normally aggravated by “closing down” movements would be included in this review as treatment

specific to the sub-group. Trials that described a sub-group as a condition for eligibility only, without any description of treatment specificity, were not included in the review. Finally, data had to be presented for two or more treatment groups, including sub-group specific manual therapy and at least one other comparison treatment.

2.1.4. Outcomes

Outcomes of interest were pain intensity measured on a visual analogue scale (VAS) or numerical pain-rating scale (NPRS), and activity limitation measured by low back specific scales (Deyo et al., 1998; Bombardier, 2000).

2.2. Data extraction and risk of bias assessment

Two reviewers independently extracted data from included trials using a standardized form (Furlan et al., 2009) developed for previous research within our group (Hahne et al., 2010). Data extracted included information on trial settings, population characteristics, sample size, treatment interventions and outcomes (mean scores and standard deviations). Where insufficient detail was provided additional information was sought from the trial authors.

The reviewers independently assessed risk of bias of included trials using the PEDro scale (Verhagen et al., 1998). This scale has demonstrated evidence of reliability (Maher et al., 2003) and validity (de Morton, 2009) and has been used in previous systematic reviews involving LBP (May & Johnson, 2008; Hahne et al., 2010). Trials scoring six or more out of 10 were considered to be of low risk of bias (Slade & Keating, 2006; Hahne et al., 2010) (Table 1).

2.3. Data synthesis and analysis

Group means and standard deviations were recorded for all outcomes at each follow-up. Where the mean was not provided, the median was taken as the next best estimate (Slade & Keating, 2006; Higgins & Green, 2011). Missing standard deviation scores were calculated using the standard error of the mean values (Higgins & Green, 2011). Treatment effect sizes and 95% confidence intervals were calculated for continuous data using Hedges adjusted-g standardized mean difference (SMD) and relative risk ratios for dichotomous data. SMD was chosen to allow easier comparison of effects measured by different outcome measures between trials.

Table 1
PEDro scale.

Item	Criteria
1	Eligibility criteria were specified
2	Subjects were randomly allocated to groups
3	Allocation was concealed
4	Groups were similar at baseline for the most important prognostic indicators
5	All participants were blinded
6	All participants who administered therapy were blinded
7	All assessors who measured at least one key outcome were blinded
8	Measures of at least one key outcome were obtained from more than 85% of the participants initially allocated to groups
9	All participants for whom outcome measures were available received the treatment or control condition as allocated, or, where this was not the case, data for a least one key outcome was analyzed by intention to treat
10	The results of between group statistical analysis are reported for at least one key outcome
11	The study provides both point measures and measures of variability for at least one key outcome.

Note: Points were awarded for criteria 2 to 11 providing an overall score out of 10.

Treatment effects favouring manual therapy were assigned a positive SMD value. A value of 0.2 was considered to be a small effect size, 0.5 moderate and 0.8 large (Cohen, 1988).

A meta-analysis of pooled data was planned if two or more trials were considered to be clinically homogenous, defined as sufficiently similar in participants characteristics, interventions and outcomes (Higgins & Green, 2011). In clinically homogenous trials, evaluation of statistical heterogeneity was calculated using Revman 5.0 (2008). Meta-analysis was not performed in trials with statistical heterogeneity as determined by P values of <0.1 on the χ^2 test, or I^2 values $>50\%$ (Hopley et al., 2010).

A qualitative analysis using the GRADE approach (Grading of Recommendations, Assessment, Development and Evaluation) (Atkins et al., 2004; Furlan et al., 2009; Higgins et al., 2011) was planned. This method provides a rating of the quality of evidence for each primary outcome based on five domains comprising: limitations of trial design, inconsistency, indirectness, imprecision of results, and publication bias (Atkins et al., 2004; Furlan et al., 2009) (Appendix 3).

3. Results

3.1. Trial selection

The process of selecting included trials for this review is described in Fig. 1. A list of excluded papers can be obtained from the author.

3.1.1. Participants and setting

A total of seven RCTs (Delitto et al., 1993; Erhard et al., 1994; Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009; Hallegraeff et al., 2009; Sutlive et al., 2009) were included randomizing 463 participants (Table 3). All seven of the trials were conducted in a physical therapy clinic setting. Six trials were

Table 2

Risk of bias assessment using the PEDro scale.

Trial	PEDro score											Total
	1	2	3	4	5	6	7	8	9	10	11	
Browder et al. (2007)	✓	✓	✓	—	—	—	—	✓	✓	✓	✓	6
Childs et al. (2004)	✓	✓	✓	✓	—	—	✓	✓	✓	✓	✓	8
Cleland et al. (2009)	✓	✓	✓	✓	—	—	✓	✓	✓	✓	✓	8
Delitto et al. (1993)	✓	✓	—	—	—	—	—	—	✓	✓	✓	4
Erhard et al. (1994)	✓	✓	—	✓	—	—	—	✓	—	✓	✓	5
Hallegraeff et al. (2009)	✓	✓	✓	—	—	—	—	✓	✓	✓	✓	6
Sutlive et al. (2009)	✓	✓	✓	✓	—	—	—	✓	✓	✓	✓	7

conducted in the USA (Delitto et al., 1993; Erhard et al., 1994; Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009; Sutlive et al., 2009) and one in the Netherlands (Hallegraeff et al., 2009).

Across the seven trials three sub-groups were defined by: (1) centralization of symptoms with repeated lumbar extension (Browder et al., 2007), (2) centralization of symptoms as well as symptom reproduction in three out of four provocative tests for sacroiliac joint pain (extension-mobilization sub-group) (Delitto et al., 1993; Erhard et al., 1994) and (3) a pre-determined clinical prediction rule (CPR) for spinal manipulation (Childs et al., 2004; Cleland et al., 2009; Hallegraeff et al., 2009; Sutlive et al., 2009) developed by Flynn et al. (2002). The CPR included five variables predicting a positive response to spinal manipulation comprising: duration of symptoms <16 days, Fear Avoidance Beliefs Questionnaire work subscale score <19 , at least one hip with $>35^\circ$ internal rotation, hypomobility in the lumbar spine and no symptoms distal to the knee.

3.1.2. Interventions

There was significant between trial heterogeneity regarding the type of manual therapy \pm co-intervention provided. In six trials manipulation was provided with a variety of co-interventions

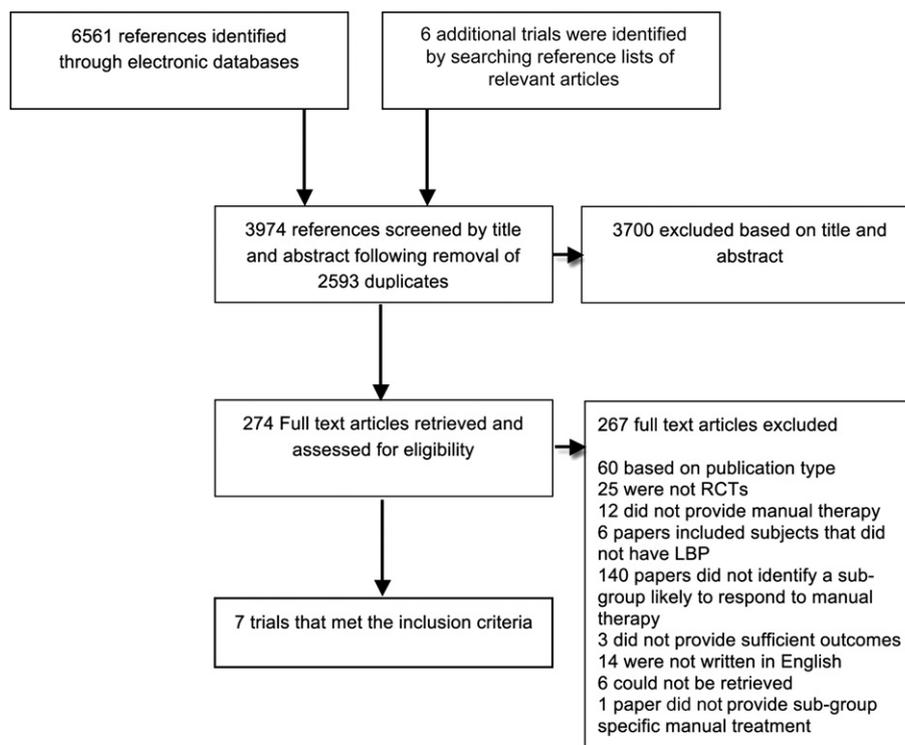


Fig. 1. Flow chart showing progression of Randomized Controlled Trials (RCTs) 4 through the selection process.

Table 3
Characteristics of included trials.

RCT	Participant characteristics, sample size, symptom duration, inclusion criteria and sub-group	Intervention	Comparison	Outcome
Browder et al. (2007) PEDro: 6	<p>Consecutive patients attending 9 PT clinics, mostly Department of Defence centers (US) N = 48 Mean age = 39 y Median symptom duration: = 59.5 d Inclusion Criteria: LBP of any duration with symptoms extending distal to the buttock on at least one lower extremity, 18–60 y, a modified ODQ score of $\geq 30\%$, Sub-group: Participants had to present with the centralization phenomenon</p>	<p>Extension Orientated Treatment Approach (EOTA) which involved:</p> <ol style="list-style-type: none"> 1. Repeated extension exercise in prone and standing. In addition 1 set of ten reps of the prone press up or extension in standing were completed every 2–3 waking hours during the four weeks. 2. 10–20 postero-anterior mobilizations of grade I–IV oscillations. Therapist selected treatment grade and segmental level of mobilization. <p>Participants were advised to maintain lumbar lordosis in sitting, avoid sitting for >20–30 min and to avoid activities that caused symptoms to peripheralize. All participants were provided with exercise instruction booklet outlining correct technique and frequency of each exercise, and progression. 6 sessions of PT over 4 weeks</p>	<p>Strengthening program. Exercise program to improve isolated contractions of the deep abdominal muscle and strengthen the primary stabilizers (TA, LM, QL, ES, IO EO). All participants were provided with exercise instruction booklet outlining correct technique and frequency of each exercise, and progression. 6 sessions of PT over 4 weeks</p>	<p>Pain (NPRS) and activity (ODQ) at 1 week, 4 weeks and 6 months</p>
Childs et al. (2004) PEDro: 8	<p>Consecutive patients referred for PT at participating clinics. Mostly US Air force healthcare facilities N = 131 Mean age = 33.9 y Median symptom duration = 27 d Inclusion criteria: LBP with or without referral into the lower extremity, 18–60 y, ODQ score of $\geq 30\%$ Sub-group: Participants positive on Flynn's CPR (2002) satisfying 4/5 criteria:</p> <ul style="list-style-type: none"> • Duration of symptoms <16 days • FABQ work subscale score <19 • At least one hip with >35° internal rotation • Hypomobility in the lumbar spine • No symptoms distal to the knee 	<p>Spinal manipulation plus exercise program: First 2 sessions included a high velocity thrust manipulation and a range of motion exercise only (Flynn et al., 2002). After two sessions the participants received a low stress aerobic and lumbar spine strengthening program identical to the control. All participants received an exercise instruction booklet outlining correct technique and frequency of each exercise and instructing them to complete once daily 5 Sessions of PT in 4 weeks</p>	<p>Exercise program alone Low stress aerobic and lumbar spine strengthening program. Lumbar spine strengthening program aimed to target stabilizers of the spine Low stress aerobic exercises started with 10 minutes of exercise on a stationary bike or treadmill at a self-selected pace. All participants received an exercise instruction booklet outlining correct technique and frequency of each exercise and instructing them to complete once daily 5 sessions of PT in 4 weeks</p>	<p>Pain (NPRS) and activity (ODQ) at 1 week, 4 weeks and 6 months</p>
Cleland et al. (2009) PEDro: 8	<p>Patients attending out patient PT clinics (US). N = 112 Mean age = 40.3 y Median symptom duration = 45 d Inclusion criteria: LBP with or without referral into the lower extremity, 18–60 y, ODQ score of >25%, Sub-group: Participants positive on Flynn's CPR (Flynn et al., 2002) satisfying 4/5 criteria.</p>	<p>Supine thrust manipulation First 2 sessions included a high velocity thrust manipulation and a range of motion exercise only (Flynn et al., 2002) After two sessions the participants received a low stress aerobic and lumbar spine strengthening program identical to the control (Childs et al., 2004). All participants received an exercise instruction booklet outlining correct technique and frequency of each exercise and instructing them to complete once daily 5 sessions of PT in 4 weeks</p>	<p>Comparison 1: side-lying thrust manipulation group Minimum of 1 manipulation per side. Maximum of two attempts at manipulation per side Comparison 2: non-thrust spinal manipulation Central lumbar posterior–anterior non-thrust manipulation directed at L4 and L5. This included 2 sets of 60 seconds of low velocity, high amplitude oscillatory force As per the intervention group all participants in both comparisons performed a spinal range of motion exercise following manipulation After two sessions all participants in both comparisons received low stress aerobic and lumbar spine strengthening program (Childs et al., 2004)</p>	<p>Pain (NPRS) and activity (ODQ) at 1 week, 4 weeks and 6 months</p>

Delitto et al. (1993) PEDro: 4	<p>Patients referred to PT for the management of LBP (US) N = 24 Mean age = 37 y (intervention), 27 y (control) Symptom duration = 6 d (intervention), 11 d (Control) Inclusion criteria: Either acute (<7 days) or subacute (7 days to 7 weeks) LBP Sub-group: Participants categorized as belonging to an extension-mobilization sub-group based on centralization of symptoms and symptom reproduction in three out of four provocative tests for sacroiliac joint pain</p>	<p>Manipulation technique for the SIJ This was followed by extension oriented treatment regime including prone press up exercises, discouragement of flexed posture and use of a lumbar roll All participants received a handout illustrating their exercises and were encouraged to complete this x3/day 3 sessions of PT in one week</p>	<p>Flexion oriented exercise regime All participants received a handout illustrating their exercises and were encouraged to complete this x3/day 3 sessions of PT in one week</p>	<p>Activity (ODQ) at day 3 and day 5 (after initiation of treatment)</p>
Erhard et al. (1994) PEDro: 5	<p>Patients presenting at PT clinics with LBP (US) N = 24 Mean age = 47 y (intervention), 41 y (control) Mean Symptom duration = 20 d (intervention), 22 d control Inclusion criteria: LBP >3 months duration Sub-group: Participants categorized as belonging to an extension-mobilization sub-group based on centralization of symptoms and symptom reproduction in three out of four provocative tests for sacroiliac joint pain</p>	<p>Manipulation technique for the SIJ SIJ manipulation immediately followed by completion of the hand-heel rock. No postural instruction. Participants received a handout illustrating their exercises If at follow-up assessment 3 or more of the SIJ signs were present, a second manipulation was administered. 3 sessions of PT in one week</p>	<p>Extension Group. Extension oriented treatment regime including press-ups, use of a lumbar roll and postural instruction. Flexed postures were discouraged. Participants were provided with a lumbar roll and written information on completing the extension exercises. 3 sessions of PT in one week</p>	<p>Activity (ODQ) at day 3, day 5 and one month (after initiation of treatment)</p>
Hallegraef et al. (2009) PEDro: 6	<p>3 Primary healthcare centers for Physical Therapy (Netherlands) N = 64 Mean age = 39 y Mean symptom duration = not given Inclusion criteria: Acute NSLBP, 20–55 y Sub-group: Participants needed to satisfy 2/5 criteria on Flynn's CPR (Flynn et al., 2002) • Duration of symptoms <16 days • No symptoms distal to the knee</p>	<p>Manual therapy and PT 4 sessions provided by manual therapists. Treatment focused on the restricted lumbar-sacral and SIJ, involving high velocity low amplitude thrust manipulations. No other technique was applied. Participants also received PT as described in the control section</p>	<p>PT alone. Low load endurance exercises, training abdominal muscles and stretching lumbar extensor muscles, (5 min × 2 per day). Participants were also provided with information on back pain, posture and advice to stay active.</p>	<p>Pain (VAS) and activity (ODQ) after treatment 4 (2 ½ weeks)</p>
Sutlive et al. (2009) PEDro: 8	<p>Military healthcare beneficiaries with LBP (US) N = 60 Mean age = 25.5 y Mean symptom duration = 13.5 d Inclusion criteria: LBP ±leg symptoms, 18–65 y, ODQ score ≥30%. Sub-group: Participants needed to satisfy 3/5 criteria on Flynn's CPR (Flynn et al., 2002). At least 1 of 3 had to be duration of symptoms ≤16 days or no pain radiating distal to the knee.</p>	<p>Lumbopelvic manipulation (LP) Participants were treated with the manipulation for which Flynn's CPR (Flynn et al., 2002) was originally developed. Minimum of 1 manipulation per side. Maximum of two attempts at manipulation per side Following manipulation all participants received an exercises booklet detailing correct performance of pelvic tilt exercises</p>	<p>Lumbar neutral gap manipulation (NG) Minimum of 1 manipulation per side. Maximum of two attempts at manipulation per side Following manipulation all participants received an exercises booklet detailing correct performance of pelvic tilt exercises</p>	<p>Pain (NPRS) and activity (ODQ) at 48 h post treatment.</p>

PT physical therapy, NSLBP non-specific low back pain, ODQ Modified Oswestry Low Back Pain Disability Questionnaire, NPRS numerical pain-rating scale, SIJ Sacroiliac joint, TA transversus abdominis, LM lumbar multifidus, QL quadratus lumborum, ES erector spinae, IO internal oblique, EO external oblique FABQ Fear Avoidance Beliefs Questionnaire, CPR clinical prediction rule. PT physical therapy, NSLBP non-specific low back pain, ODQ Modified Oswestry Low Back Pain Disability Questionnaire, NPRS numerical pain-rating scale, SIJ Sacroiliac joint, TA transversus abdominis, LM lumbar multifidus, QL quadratus lumborum, ES erector spinae, IO internal oblique, EO external oblique FABQ Fear Avoidance Beliefs Questionnaire.

including trunk muscle training (Childs et al., 2004; Cleland et al., 2009; Hallegraef et al., 2009), an extension oriented program (Delitto et al., 1993), flexion and extension exercises (Erhard et al., 1994) and a range of movement exercise (Sutlive et al., 2009). The final trial (Browder et al., 2007) provided physiotherapy mobilization as part of an extension oriented treatment approach including extension exercises and postural education.

There was also significant between trial heterogeneity in the comparison interventions provided which comprised of a flexion oriented exercise program (Delitto et al., 1993), an extension oriented program (Erhard et al., 1994), and trunk muscle training (Childs et al., 2004; Browder et al., 2007; Hallegraef et al., 2009). The trunk muscle training involved stabilization exercise, and, in two of these three trials, an additional low stress aerobic exercise program (Childs et al., 2004; Hallegraef et al., 2009). Two trials included alternative manual therapy as the comparison, defined as a manipulation technique not used in the development of Flynn

et al.'s CPR for spinal manipulation (Cleland et al., 2009; Sutlive et al., 2009).

3.1.3. Outcomes

All trials assessed activity using a version of the Oswestry Disability Questionnaire (ODQ). Five trials assessed pain using the VAS or NPRS (Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009; Hallegraef et al., 2009; Sutlive et al., 2009). All seven trials assessed short term follow-up at less than three months and three trials assessed intermediate term follow-up between three months and one year (Fig. 2) (Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009). In some trials, follow-up data were collected at multiple time points within the pre-determined time periods. In this event, data were reported from the follow-up time point less than and closest to pre-determined thresholds of six weeks (short term), six months (intermediate term) and 12 months (long-term) (Appendix 2).

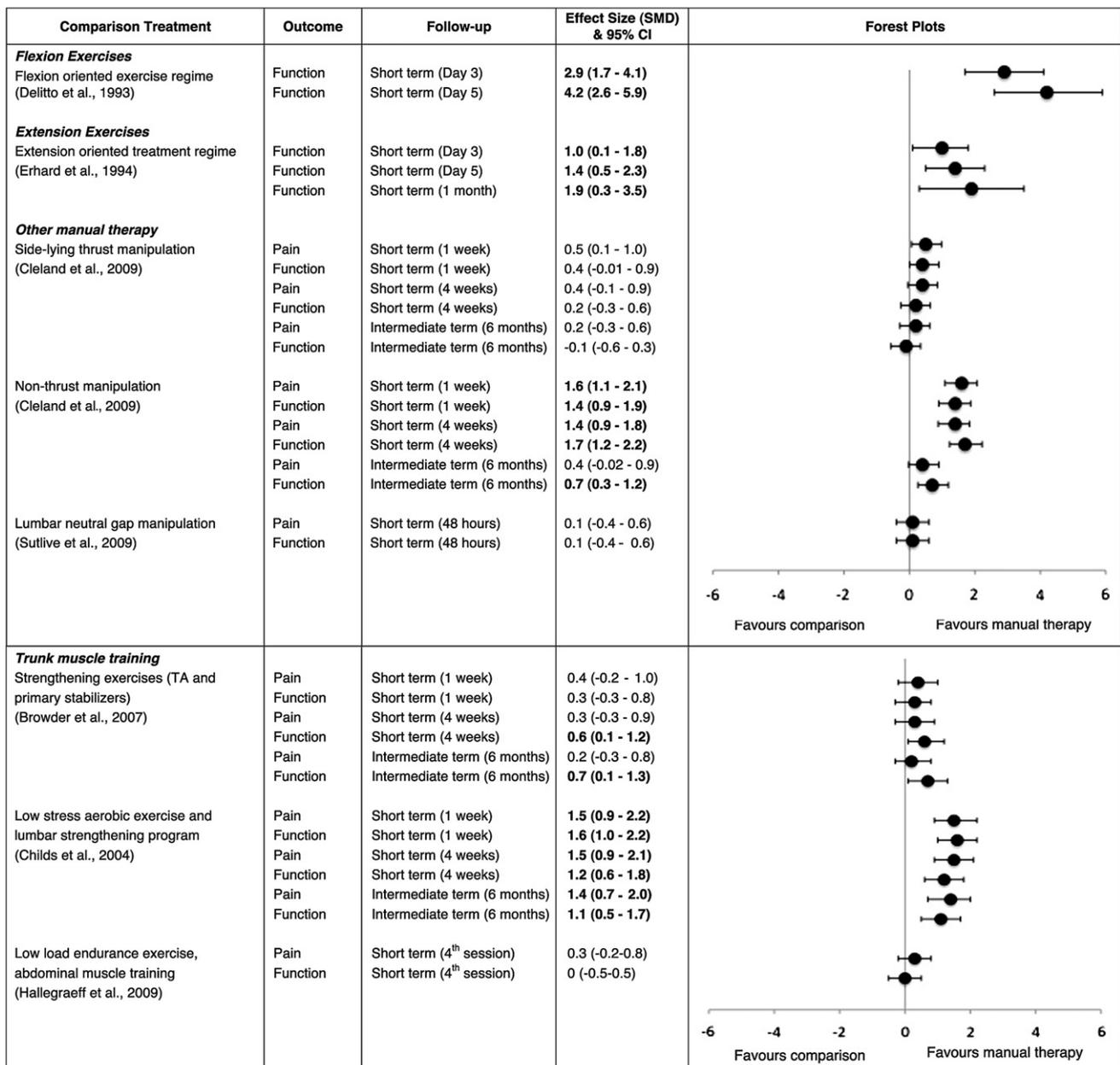


Fig. 2. Results of treatment effects for specific manual therapy versus comparison treatments for all included trials.

3.2. Assessment of risk of bias

The mean PEDro score for the trials was 6/10 (Table 2). Five of the seven trials scored six or more on the PEDro scale and were considered to be at low risk of bias (Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009; Hallegraef et al., 2009; Sutlive et al., 2009). In the two trials (Delitto et al., 1993; Erhard et al., 1994) scoring less than six, limitations included inadequate concealment of treatment allocation, insufficient blinding of assessors and failure to apply intention to treat principles. All trials lost points for failure to blind participants and therapists, however in RCTs of physiotherapy interventions, this is extremely difficult (Bhogal et al., 2005).

Five of the trials (Delitto et al., 1993; Erhard et al., 1994; Browder et al., 2007; Cleland et al., 2009; Sutlive et al., 2009) were considered to be clinically heterogeneous. Two trials (Childs et al., 2004; Hallegraef et al., 2009) were considered to be clinically homogenous, based on sub-group, intervention and comparison treatment, but were assessed to be statistically heterogeneous ($I^2 = 92\%$). Meta-analysis was therefore not appropriate for any combinations of trials (Dagenais et al., 2007; Staal et al., 2008). The overall quality of the evidence was rated according to the five GRADE domains and downgraded for reasons including limitations in study design based on the PEDro score, inconsistency due to conflicting results, indirectness due to clinical heterogeneity and imprecision due to sparse data (Furlan et al., 2009; BMJ Clinical Evidence, 2012) (Appendices 2 and 3).

3.3. Specific manual therapy versus a flexion oriented exercise program

One trial with high risk of bias ($n = 24$) compared specific manual therapy (and the co-intervention of an extension oriented program) to flexion exercises (Delitto et al., 1993). This trial found a large and significant treatment effect for activity at short term follow-up in favour of manual therapy (SMD: 4.2 95% CI: 2.6–5.9) but provided very low quality evidence that specific manual therapy was better than flexion exercises at short term follow-up (trial limitations, indirectness and imprecision).

3.4. Specific manual therapy versus an extension oriented program

One trial with high risk of bias ($n = 24$) compared specific manual therapy (and the co-intervention of flexion exercises) with an extension oriented program (Erhard et al., 1994). This trial found a large and significant effect size for activity at short term follow-up favouring manual therapy (SMD: 1.9 95% CI: 0.3–3.5) but provided very low quality evidence that specific manual therapy was better than an extension oriented program at short term follow-up (trial limitations, indirectness and imprecision).

3.5. Specific manual therapy versus other manual therapy

Two trials with low risk of bias compared specific manual therapy (manipulation matched to Flynn et al.'s CPR) to either spinal manipulation (unmatched to the CPR) (Cleland et al., 2009; Sutlive et al., 2009) or mobilization techniques (Cleland et al., 2009). One of these trials ($n = 112$) (Cleland et al., 2009) showed a large and significant treatment effect on pain and activity at short term follow-up (short term pain SMD: 1.4 95% CI: 0.9–1.8, short term activity SMD: 1.7 95% CI: 1.2–2.2) and a moderate and significant effect on activity at intermediate term follow-up (SMD: 0.7 95% CI: 0.3–1.2) in favour of the specific manipulation over mobilization. Comparisons between manipulation matched to the CPR and unmatched manipulation showed no significant difference (Cleland et al., 2009; Sutlive et al., 2009). Overall, there was very low quality evidence that specific manual therapy was better than

other manual therapy at reducing pain and increasing activity at any follow-up period (inconsistency, indirectness, imprecision).

3.6. Specific manual therapy versus trunk muscle training

Three trials with low risk of bias compared specific manual therapy to trunk muscle training (Childs et al., 2004; Browder et al., 2007; Hallegraef et al., 2009). Two trials (Childs et al., 2004; Hallegraef et al., 2009) compared spinal manipulation with trunk muscle training to trunk muscle training alone in a group of participants positive on Flynn et al.'s CPR. One of these trials ($n = 131$) (Childs et al., 2004) demonstrated a large and significant effect on pain and activity in the short term and intermediate term (short term pain SMD: 1.5 95% CI: 0.9–2.1, intermediate term pain SMD: 1.4 95% CI: 0.7–2.0, short term activity SMD: 1.2 95% CI: 0.6–1.8 and intermediate activity SMD: 1.1 95% CI: 0.5–1.7) in favour of the specific manipulation group. The second trial ($n = 64$) (Hallegraef et al., 2009) showed no significant differences. Another trial ($n = 48$), compared manual therapy as part of an extension oriented treatment approach to trunk muscle training (Browder et al., 2007). This trial demonstrated a moderate and significant effect on improving activity at both short and intermediate follow-up (short term SMD: 0.6 95% CI: 0.1–1.2; intermediate term SMD: 0.7 95% CI: 0.1–1.3) in favour of the group receiving manual therapy. In summary, there was very low quality evidence that specific manual therapy was better than trunk muscle training at reducing pain and increasing activity at any follow-up period (inconsistency, indirectness, imprecision).

4. Discussion

Researchers have proposed that RCTs evaluating sub-group specific treatment for LBP should demonstrate larger effect sizes than those published to date. Given the emphasis in the LBP literature on classification and sub-grouping (Ford et al., 2007; Fritz et al., 2007), it was surprising that only seven RCTs were located during this review that reported data on sub-groups likely to respond to specific manual therapy. Due to the clinical and statistical heterogeneity between the included trials, a meta-analysis was not appropriate. However, statistically significant treatment effects favouring manual therapy in individual trials provided preliminary evidence that when manual therapy is targeted to a sub-group likely to respond it may produce a greater reduction in pain and greater increase in activity than comparison treatments.

Numerous systematic reviews have evaluated the effectiveness of manual therapy for LBP, however none have focused solely on sub-group specific manual therapy (Assendelft et al., 2004; Dagenais et al., 2010; Rubinstein et al., 2011). There is considerable potential for treatment effects to be masked in such reviews because of clinical heterogeneity (Delitto, 2005; Ford et al., 2007). A recent systematic review attempted to address the heterogeneity issue by evaluating targeted manual therapy and/or exercise to non-targeted treatment (Kent et al., 2010). This review only included trials that evaluated the effectiveness of manual therapy on sub-groups developed using a 'treatment effect modifier' approach, which considers response to a specific treatment based on a collection of signs and symptoms (Kent et al., 2010). Therefore only one of the seven trials identified by the current review was included by these authors (Childs et al., 2004).

In the current review, six of the seven included trials originated or had authors from a common research group (Delitto et al., 1993; Erhard et al., 1994; Childs et al., 2004; Browder et al., 2007; Cleland et al., 2009; Sutlive et al., 2009) that predominantly used statistical methods to develop a CPR. Whilst a statistical approach to developing sub-group criteria is becoming more common (Stanton et al., 2010) and has been recommended by some (Hancock et al., 2009;

Kent et al., 2010), other approaches can be appropriate in complex domains such as LBP (George & Delitto, 2005; Ford et al., 2007; Reitsma et al., 2009; Ford et al., 2011).

Despite the overall GRADE quality of the evidence being very low, two trials with low risk of bias (Childs et al., 2004; Cleland et al., 2009) and two with high risk of bias (Delitto et al., 1993; Erhard et al., 1994) produced large treatment effects in favour of specific manual therapy. Researchers have recommended replication of RCTs that demonstrate large effect sizes (Underwood, 2005; Bogduk, 2010). It is of note that all of these trials provided a manipulation technique developed for the sacroiliac joint (Delitto et al., 1993; Flynn et al., 2002) for participants scoring four out of five on the Flynn CPR (2002) or being positive for provocative sacroiliac joint testing. Support for replication of these trials is strengthened by the fact that large effects were demonstrated compared to other types of physiotherapy treatment. This contrasts to an overwhelming number of systematic reviews that show minimal clinically meaningful effects when comparing one physiotherapy treatment to another (Assendelft et al., 2004; Rubinstein et al., 2011).

4.1. Limitations

We were unable to conduct a meta-analysis in the review because of heterogeneity between trials (Dagenais et al., 2007; Kent et al., 2010). Meta-analysis when clinical heterogeneity is present can lead to bias within the results (Bogduk et al., 2009). We therefore chose to conduct a qualitative GRADE analysis based on the recommendations of previous reviews (Staal et al., 2008; Kroeling et al., 2009). A second potential limitation of this review was the inclusion of RCTs providing co-interventions alongside sub-group specific manual therapy. However, this approach has been used in other systematic reviews in LBP research (Ferreira et al., 2006) and attempts to better reflect clinical practice (Delitto, 2005). We chose to exclude papers that were not written in English. However previous research has shown that this has little impact on treatment effect estimates (Juni et al., 2002; Moher et al., 2003).

Sub-grouping of LBP can be a complex exercise (Ford et al., 2007) and a potential limitation of this review was the eligibility criterion requiring a sub-group of LBP being identified by the trial authors as likely to respond to manual therapy. In order to ensure that all included trials were consistent with a specific manual therapy approach a number of confirmatory eligibility criteria were employed. One of the papers excluded on the basis of these criteria was Hancock et al. (2008) due to the treatment provided not matching that upon which the Flynn CPR was developed. We believe this approach is consistent with identifying RCTs with an explicit aim of evaluating the effectiveness of sub-group specific manual therapy.

5. Conclusion

This review found preliminary evidence that sub-group specific manual therapy may produce a greater reduction in pain and increase in activity in people with LBP when compared with other treatments. Individual trials with low risk of bias found large and significant effect sizes in favour of specific manual therapy. However, because of downgrading using the GRADE domains, the overall quality of evidence was found to be very low. Further research is required with a particular focus on evaluating the effect of specific manual therapy on sub-groups with acceptable validity.

Acknowledgements

Supported by a La Trobe University Postgraduate Research Scholarship.

Appendix 1

Search strategy: EMBASE (Ovid, 1988–October 2010)

Phase 1: Back pain type terms

1. exp backache/
2. exp low back pain/
3. exp lumbar spine/
4. lumbosacral spine/
5. exp ischialgia/
6. backache\$.ti,ab.
7. back pain\$.ti,ab.
8. low back syndrom\$.ti,ab.
9. LBP.ti,ab.
10. dorsalgia.ti,ab.
11. sciatic\$.ti,ab.
12. coccy\$.ti,ab.
13. spondy\$.ti,ab.
14. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13

Phase 2: Manual therapy type terms

15. exp manipulative medicine/
16. exp orthopedic manipulation/
17. exp osteopathic medicine/
18. exp kinesiology/
19. manip\$.ti,ab.
20. manual therap\$.ti,ab.
21. Mobil\$.ti,ab.
22. 21 or 18 or 19 or 16 or 17 or 20 or 15

Phase 3: Trial type terms

23. randomized controlled trial/
24. exp controlled clinical trial/
25. clinical trial/
26. controlled study/
27. exp randomization/
28. single blind procedure/
29. double blind procedure/
30. Major Clinical study/
31. Crossover procedure/
32. Multicenter study/
33. Placebo/
34. Phase 3 Clinical Trial/
35. Phase 4 Clinical Trial/
36. random\$.ti,ab.
37. RCT.ti,ab.
38. placebo\$.ti,ab.
39. trial\$.ti,ab.
40. group\$.ti,ab.
41. 35 or 33 or 32 or 26 or 30 or 23 or 29 or 25 or 27 or 28 or 39 or 40 or 36 or 38 or 34 or 24 or 37 or 31

Phase 4: Limit to humans

42. limit 41 to humans
43. 42 and 22 and 14

Appendix 2

Tables of GRADE Quality Assessment and Summary of Findings for all Outcomes and Comparisons

Specific manual therapy (and co-interventions) versus a flexion oriented exercise program (Delitto et al., 1993)

Quality of evidence	Summary of findings							GRADE quality of evidence
	Outcome (number of trials)	Trial limitations	Inconsistency	Indirectness	Imprecision	Manual therapy (N)	Flexion program (N)	
Function – short term follow-up (1)	–1 serious ^a	Non-applicable ^f	–1 indirectness ^c	–1 imprecision ^d	14	10	2.9 (1.7 to 4.1) 4.2 (2.6 to 5.9)	Very low

Specific manual therapy (and co-interventions) versus an extension oriented program (Erhard et al., 1994)

Quality of evidence	Summary of findings							GRADE quality of evidence
	Outcome (number of trials)	Trial limitations	Inconsistency	Indirectness	Imprecision	Manual therapy (N)	Extension program (N)	
Function – short term follow-up (1)	–1 serious ^a	Non-applicable ^f	–1 indirectness ^c	–1 imprecision ^d	6	6	1.9 (0.3 to 3.5)	Very low

Specific manual therapy (and co-interventions) versus other manual therapy (and co-interventions) (Cleland et al., 2009; Sutlive et al., 2009)

Quality of evidence	Summary of findings							GRADE quality of evidence
	Outcome (number of trials)	Trial limitations	Inconsistency	Indirectness	Imprecision	manual therapy (N)	Other normal therapy (N)	
Pain – short term follow-up (3) (Cleland et al., 2009; Sutlive et al., 2009)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	37	38	0.4 (–0.1 to 0.9)	Very low
					37	37	1.4 (0.9 to 1.8)	
					30	30	0.1 (–0.4 to 0.6)	
Function – short term follow-up (3) (Cleland et al., 2009; Sutlive et al., 2009)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	37	38	0.2 (–0.3 to 0.6)	Very low
					37	37	1.7 (1.2 to 2.2)	
					30	30	0.1 (–0.4 to 0.6)	
Pain – intermediate term follow-up (2) (Cleland et al., 2009) ^g	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	37	38	0.2 (–0.3 to 0.6)	Very low
					37	37	0.4 (–0.02 to 0.9)	
Function – intermediate term follow-up (2) (Cleland et al., 2009) ^g	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	37	38	–0.1 (–0.6 to 0.3)	Very low
					37	37	0.7 (0.3 to 1.2)	

Specific manual therapy (and co-interventions) versus lumbar trunk muscle training (Childs et al., 2004; Browder et al., 2007; Hallegraeff et al., 2009)

Quality of evidence	Summary of findings							GRADE quality of evidence
	Outcome (number of trials)	Trial limitations	Inconsistency	Indirectness	Imprecision	Manual therapy (N)	Trunk muscle training (N)	
Pain – short term follow-up (3) (Childs et al., 2004; Browder et al., 2007; Hallegraeff et al., 2009)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	26	22	0.3 (–0.3 to 0.9)	Very low
					23	24	1.5 (0.9 to 1.2)	
					31	33	0.3 (–0.2 to 0.8)	
							0.7 (0.3 to 1.2)	

(continued on next page)

(continued)

Quality of evidence	Summary of findings				Manual therapy (N)	Trunk muscle training (N)	Effect size (SMD) with CI	GRADE quality of evidence
	Trial limitations	Inconsistency	Indirectness	Imprecision				
Function – short term follow-up (3) (Childs et al., 2004; Browder et al., 2007; Hallegraef et al., 2009)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	26	22	0.6 (0.1 to 1.2) (Browder et al., 2007)	Very low
					23	24	1.2 (0.6 to 1.8) (Childs et al., 2004)	
					31	33	0 (–0.5 to 0.5) (Hallegraef et al., 2009)	
Pain – intermediate follow-up (2) (Childs et al., 2004; Browder et al., 2007)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	26	22	0.2 (–0.3 to 0.8) (Browder et al., 2007)	Very low
					23	24	1.4 (0.7 to 2.0) (Childs et al., 2004)	
Function – Intermediate term follow-up (2) (Childs et al., 2004; Browder et al., 2007)	Nil serious	–1 inconsistency ^b	–1 indirectness ^c	–1 imprecision ^e	26	22	0.7 (0.1 to 1.3) (Browder et al., 2007)	Very low
					23	24	1.1 (0.5 to 1.7) (Childs et al., 2004)	

Treatment effects favouring specific manual therapy were assigned positive SMD values.

Results in bold represent statistically significant effects favouring manual therapy versus comparisons based on the 95% confidence intervals of the SMD.

N = number of participants, SMD = standardized mean difference.

Where SMD values are given, these are for individual outcomes in single trials. Meta-analysis was not performed due to the clinical and statistical heterogeneity of included trials.

^a Quality point deducted for trial with high risk of bias (rating using the PEDro scale: <6/10).

^b Quality point deducted for inconsistency due to conflicting results.

^c Quality point deducted for indirectness due clinical heterogeneity between trials or for issues limiting generalisability of findings, including, but not limited to, the inclusion of co-interventions, restricted sample population, unclear definition of measurement and outcomes, as recommended by BMJ Clinical Evidence (BMJ Clinical Evidence, 2012).

^d Quality point deducted for imprecision due to data from single trial (as recommended by the Cochrane Back Review Group (Furlan et al., 2009)).

^e Quality point deducted for imprecision due to sparse data as recommended by GRADE working group (Atkins et al., 2004) and BMJ Clinical Evidence (BMJ Clinical Evidence, 2012).

^f Inconsistency is non-applicable with only one trial.

^g One trial, two comparison interventions.

Appendix 3

GRADE assessment of the overall quality of the body of evidence in systematic reviews

The GRADE approach evaluates the overall quality of evidence for individual outcomes based on domains including limitations of study design and risk of bias, consistency, directness, precision of results and publication bias (Atkins et al., 2004; Furlan et al., 2009; BMJ Clinical Evidence, 2012).

In this review, the following definitions for assessing the quality of evidence were applied to each outcome for each comparison (Atkins et al., 2004; Furlan et al., 2009):

- High quality evidence: further research is very unlikely to change our confidence in the estimate of effect. At least 75% of the RCTs with no trial limitations, have consistent findings, direct and precise data, and no known or suspected publication biases
- Moderate quality evidence: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. One of the domains is not met
- Low quality evidence: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Two of the domains are not met
- Very low quality of evidence: Any estimate of effect is very uncertain. Three of the domains are not met
- No evidence: No RCTs were identified that addressed this outcome

The following domain definitions were used in this review to assess the overall quality of the evidence.

Limitations of trial design

Refers to the risk of bias of included studies. This was rated using the PEDro scale which has demonstrated reliability (Maher et al.,

2003; Bhogal et al., 2005). Trials with a score of six or more were considered low risk of bias.

If 75% or more of the included trials scored 6 or more then this domain was marked as “no limitations” (Atkins et al., 2004; Furlan et al., 2009; Schaafsma et al., 2010). A quality point was deducted on this domain when less than 75% of trials included for each outcome scored 6 or more on the PEDro scale.

Consistency

This domain refers to the similarity of treatment effect estimates for each outcome across the trials (Atkins et al., 2004). Trial results were considered consistent when directions, effect size and statistical significance were considered similar enough to draw the same conclusion (Furlan et al., 2009).

“Consistency in direction was defined as 75% of more of the included trials showing either benefit or no benefit, and consistency of effect when 75% or more of the trials showing a clinically important or unimportant treatment effect” (based on the minimum clinically important difference for the outcome measures considered in the review) as recommended by the CBRG. A quality point was deducted for inconsistent or conflicting results (BMJ Clinical Evidence, 2012).

When all studies are included in the meta-analysis, ‘consistency’ is defined as absence of statistical heterogeneity (Gross et al., 2010).

This domain was not applicable when there was only one trial per outcome.

Directness

Refers to the generalizability of findings (Furlan et al., 2009). Indirectness occurs when the population, interventions and outcomes are sufficiently different to the population of interest in the review (Atkins et al., 2004). A quality point was deducted for clinical heterogeneity between trials (BMJ Clinical Evidence, 2012) or for issues limiting generalizability of findings including, but not limited to, use of co-interventions, restricted sample population,

the comparison group receiving treatment expected to be less effective than standard treatment, unclear definition and measurements of outcomes

Precision

Refers to number of studies, the population and the events for each outcome (Furlan et al., 2009; Kroeling et al., 2009). Data were considered imprecise and a quality point deducted when:

- Only one study reported a particular outcome for the chosen comparison (Furlan et al., 2009)
- There were sparse data for each comparison, with less than 200 participants per comparison (BMJ Clinical Evidence, 2012)
- Fewer than 75% of RCTs presented data that could be included in a meta-analysis (Furlan et al., 2009)
- In studies where a meta-analysis had been performed, the confidence intervals were wide enough that the estimate of treatment effect provides conflicting recommendations (Atkins et al., 2004)

Publication bias

This was only considered when there was evidence of reporting bias (Higgins & Green, 2011).

References

- Assendelft WJ, Morton SC, Yu EI, Suttrop MJ, Shekelle PG. Spinal manipulative therapy for low-back pain. *Cochrane Database of Systematic Reviews*; 2004 (1).
- Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *British Medical Journal (Clinical Research Edition)* 2004; 328:1490.
- Bhogal SK, Teasell RW, Foley NC, Speechley MR. The PEDro scale provides a more comprehensive measure of methodological quality than the Jada Scale in stroke rehabilitation literature. *Journal of Clinical Epidemiology* 2005;58:668–73.
- BMJ Clinical Evidence. What is GRADE?, <http://clinicalevidence.bmj.com/ceweb/about/about-grade.jsp>; 2012.
- Bogduk N. A cure for back pain? *Pain* 2010;149:7–8.
- Bogduk N, Dreyfuss P, Govind J. A narrative review of lumbar medial branch neurotomy for the treatment of back pain. *Pain Medicine* 2009;10(6): 1035–45.
- Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 2000;25(24):3100–3.
- Browder DA, Childs JD, Cleland JA, Fritz JM. Effectiveness of an extension-orientated treatment approach in a subgroup of subjects with low back pain: a randomized clinical trial. *Physical Therapy* 2007;87(12):1608–18.
- Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Annals of Internal Medicine* 2004;141(12):920–8.
- Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Annals of Internal Medicine* 2007;147(7):478–91.
- Cleland JA, Fritz JM, Kulig K, et al. Comparison of the effectiveness of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule: a randomized clinical trial. *Spine* 2009;34(25):2720–9.
- Cochrane Back Review Group. Updated search strategies for CBG; 2008.
- Cohen J. *Statistical power analysis in the behavioural sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1988.
- Coste J, Delecoeuillerie G, Cohen de Lara A, Le Parc JM, Paolaggi JB. Clinical course and prognostic factors in acute low back pain: inception cohort study in primary care practice. *British Medical Journal* 1994;308:577–80.
- Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine Journal* 2008;8:8–20.
- Dagenais S, Gay E, Tricco AC, Freeman MD, Mayer JMNASS. Contemporary concepts in spine care: spinal manipulation therapy for acute low back pain. *Spine Journal* 2010;10:918–40.
- Dagenais S, Yelland MJ, Del Mar C, Schoene ML. Prolotherapy injections for chronic low-back pain. *Cochrane Database of Systematic Reviews*; 2007. doi:10.1002/14651858.CD004059.pub3 (2). Art. No.: CD004059.
- Delitto A. Research in low back pain: time to stop seeking the elusive "magic bullet". *Physical Therapy* 2005;85(3):206–8.
- Delitto A, Cibulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Physical Therapy* 1993;73(4):216–22.
- Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23(18):2003–13.
- Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from the U.S. National Surveys, 2002. *Spine* 2006;31(23):2724–7.
- Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Physical Therapy* 1994;74(12): 1093–100.
- Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Maher C. Efficacy of spinal manipulation for low back pain of less than three months' duration. *Journal of Manipulative and Physiological Therapeutics* 2003;26(9):593–601.
- Ferreira PH, Ferreira ML, Maher CG, Herbert RD, Refshauge K. Specific stabilisation exercise for spinal and pelvic pain: a systematic review. *Australian Journal of Physiotherapy* 2006;52(2):79–88.
- Flynn T, Fritz JM, Whitman JM, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine* 2002;27(24):2835–43.
- Ford J, Story I, O'Sullivan P, McMeeken J. Classification systems for low back pain: a review of the methodology for development and validation. *Physical Therapy Reviews* 2007;12:33–42.
- Ford JJ, Thompson SL, Hahne AJ. A classification and treatment protocol for low back disorders part 1- specific manual therapy. *Physical Therapy Reviews* 2011;16(3): 168–77.
- Fritz J, Cleland J, Childs J. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. *Journal of Orthopaedic and Sports Physical Therapy* 2007;37(6):290–302.
- Fritz JM, Cleland JA, Speckman M, Brennan GP, Hunter SJ. Physical therapy for acute low back pain: associations with subsequent healthcare costs. *Spine* 2008; 33(16):1800–5.
- Furlan A, Pennick V, Bombardier C, van Tulder M. Updated method guidelines for systematic reviews in the Cochrane back review group. *Spine* 2009;34(18):1929–41.
- George SZ, Delitto A. Clinical examination variables discriminate among treatment-based classification groups: a study of construct validity in patients with acute low back pain. *Physical Therapy* 2005;85(4):306–14.
- Gross A, Miller J, D'Sylva J, et al. Manipulation or mobilisation for neck pain: a Cochrane review. *Manual Therapy* 2010;15:315–33.
- Hahne A, Ford J, McMeeken J. Conservative management of lumbar disc herniation with associated radiculopathy. *Spine* 2010;35(11):E488–504.
- Hallegraff HJM, de Greef M, Winters JC, Lucas C. Manipulative therapy and clinical prediction rule criteria in treatment of acute nonspecific low back pain. *Perceptual and Motor Skills* 2009;108:196–208.
- Hancock MJ, Herbert RD, Maher CG. A guide to interpretation of studies investigating subgroups of responders to physical therapy interventions. *Physical Therapy* 2009;89:698–704.
- Hancock MJ, Maher CG, Latimer J, Herbert RD, McAuley JH. Independent evaluation of a clinical prediction rule for spinal manipulative therapy: a randomised controlled trial. *European Spine Journal* 2008;17:936–43.
- Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. *European Spine Journal* 2003;12:149–65.
- Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions* version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from, www.cochrane-handbook.org; 2011.
- Hopley C, Stengel D, Ekkernkamp A, Wich M. Primary total hip arthroplasty versus hemiarthroplasty for displaced intracapsular hip fractures in older patients: systematic review. *British Medical Journal*; 2010.
- Juni P, Hohenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology* 2002;31(1):115–23.
- Kent P, Keating J. Do primary-care clinicians think that non-specific low back pain is one condition? *Spine* 2004;29(9):1022–31.
- Kent P, Keating J. The epidemiology of low back pain in primary care. *Chiropractic and Osteopathy*; 2005:13.
- Kent P, Marks D, Pearson W, Keating J. Does clinician treatment choice improve the outcomes of manual therapy for non specific low back pain? A metaanalysis. *Journal of Manipulative and Physiological Therapeutics* 2005;28(5):312–22.
- Kent P, Mjosund H, Petersen D. Does targeting manual therapy/and or exercise improve patient outcomes on nonspecific low back pain? A systematic review. *BMC Medicine* 2010;8(22).
- Kent PM, Keating JL, Buchbinder R. Searching for a conceptual framework for nonspecific low back pain. *Manual Therapy* 2009;14:387–96.
- Koes B, van Tulder M, Thomas S. Diagnosis and treatment of low back pain. *British Medical Journal* 2006;332:1430–4.
- Kroeling P, Gross A, Goldsmith CH, et al. Electrotherapy for neck pain. *Cochrane Database of Systematic Reviews*; 2009. doi:10.1002/14651858.CD004251.pub4. Issue 4. Art. No.: CD004251.
- Li L, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. *Physical Therapy* 2001;81:1018–28.
- Liddle D, Baxter D, Gracey J. Physiotherapists' use of advice and exercise for the management of chronic low back pain: a national survey. *Manual Therapy* 2009;14:189–96.
- Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical Therapy* 2003;83:713–21.
- Maitland G. *Maitland's vertebral manipulation*. 7th ed. Philadelphia: Elsevier; 2005.
- May S, Johnson R. Stabilisation exercises for low back pain: a systematic review. *Physiotherapy* 2008;94:179–89.

- Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment* 2003;7(41):1–90.
- de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *The Australian Journal of Physiotherapy* 2009;55(2):129–33.
- National Collaborating Centre for Primary Care. *Low back pain: early management of persistent non-specific low back pain*. London: National Institute for Health and Clinical Excellence; 2009.
- Reitsma JB, Rutjes AWS, Khan K, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology* 2009;62:797–806.
- Version 5.0 Review Manager. (RevMan) [Computer program]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2008.
- Riddle D. Classification and low back pain: a review of the literature and critical analysis of selected systems. *Physical Therapy* 1998;78(7):708–37.
- Rubinstein SM, van Middelkoop M, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for chronic low-back pain. *Cochrane Database of Systematic Reviews*; 2011. doi:10.1002/14651858.CD008112.pub2. Issue 2. Art. No.: CD008112.
- Schaafsma F, Schonstein E, Whelan KM, et al. Physical conditioning programs for improving work outcomes in workers with back pain. *Cochrane Database of Systematic Reviews*; 2010. doi:10.1002/14651858.CD001822.pub2. Issue 1. Art. No.: CD001822.
- Slade SC, Keating J. Trunk-strengthening exercises for chronic low back pain: a systematic review. *Journal of Manipulative and Physiological Therapeutics* 2006;29(2):163–73.
- Staal JB, de Bie R, de Vet HCW, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low-back pain. *Cochrane Database of Systematic Reviews*; 2008. doi:10.1002/14651858.CD001824.pub3 (3): Art. No.: CD001824.
- Stanton TR, Hancock MJ, Maher CG, Koes BW. Critical appraisal of clinical prediction rules that aim to optimise treatment selection for musculoskeletal conditions. *Physical Therapy* 2010;90(6):843–54.
- Sutlive TG, Mabry LM, Easterling EJ, et al. Comparison of short-term response to two spinal manipulation techniques for patients with low back pain in a military beneficiary population. *Military Medicine* 2009;174(7):750–6.
- van Tulder M, Becker A, Bekkering T, et al. European guidelines for the management of acute nonspecific low back pain in primary care. *European Spine Journal* 2004;15(Suppl. 2):s169–91.
- Underwood M. A clinical prediction rule predicted outcome in patients with low back pain having spinal manipulation and exercise treatment. *Evidence-Based Medicine* 2005;10:125.
- Verhagen AP, de Vet HC, de Bie RA, et al. The Delphi list: a criteria list for quality assessment of randomised clinical trial for conducting systematic reviews developed by Delphi consensus. *Journal Clinical Epidemiology* 1998;51(12):1235–41.
- Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21(24):2820–5.