

Conservative Management of Lumbar Disc Herniation With Associated Radiculopathy

A Systematic Review

Andrew J. Hahne, BPhysio,* Jon J. Ford, PhD,* and Joan M. McMeeken, MSct

Study Design. A systematic review of randomized controlled trials.

Objective. To determine the efficacy and adverse effects of conservative treatments for people who have lumbar disc herniation with associated radiculopathy (LDHR).

Summary of Background Data. Although conservative management is commonly used for people who have LDHR, the efficacy and adverse effects of conservative treatments for this condition are unclear.

Methods. We searched 10 computer databases for trials published in English between 1971 and 2008. Trials focusing on people with referred leg symptoms and radiologic confirmation of a lumbar disc herniation were included if at least 1 group received a conservative and noninjection treatment.

Results. Eighteen trials involving 1671 participants were included. Seven (39%) trials were considered of high quality. Meta-analysis on 2 high-quality trials revealed that advice is less effective than microdiscectomy surgery at short-term follow-up, but equally effective at long-term follow-up. Individual high-quality trials provided moderate evidence that stabilization exercises are more effective than no treatment, that manipulation is more effective than sham manipulation for people with acute symptoms and an intact annulus, and that no difference exists among traction, laser, and ultrasound. One trial showed some additional benefit from adding mechanical traction to medication and electrotherapy methods. Adverse events were associated with traction (pain, anxiety, lower limb weakness, and fainting) and ibuprofen (gastrointestinal events).

Conclusion. Advice is less effective than microdiscectomy in the short term but equally effective in the long term for people who have LDHR. Moderate evidence favors stabilization exercises over no treatment, manipulation over sham manipulation, and the addition of mechanical traction to medication and electrotherapy. There was no difference among traction, laser, and ultrasound. Adverse events were associated with traction and ibuprofen. Additional high-quality trials would allow firmer conclusions regarding adverse effects and efficacy.

Key words: systematic review, sciatica, radiculopathy, disc herniation, conservative treatment. **Spine 2010;35:E488–E504**

Lumbosacral radiculopathy is a condition that results from compression of 1 or more spinal nerve roots, and is characterized by radiating leg pain and parasthesia, as well as clinical signs of neurologic impairment.¹ Lumbar disc herniation, defined as the localized displacement of disc material beyond the margins of the intervertebral disc space,² is considered to be the most common cause of lumbosacral radiculopathy.^{3–5} Systematic reviews have evaluated the efficacy of surgical⁵ and injection therapies^{6–8} for the management of lumbar disc herniation with associated radiculopathy (LDHR). Although discectomy surgery has been shown to provide rapid reduction in leg pain and good overall treatment satisfaction for 65% to 90% of people,⁵ and injections have been shown to provide at least short-term pain relief for 50% to 85% of patients,^{6–8} they are costly. Data from the United States showed that \$5 billion was spent on inpatient laminectomy and discectomy surgeries for LDHR in 2004,⁹ with a further \$450 million spent annually on costs associated with epidural injections.¹⁰ In addition to their cost, these treatments carry small risks of significant adverse events.^{11,12}

Although surgical and injection therapies remain a useful option for people with LDHR who wish to hasten their recovery,⁵ the higher costs and potential risks associated with these treatments may explain why noninjection conservative management remains the preferred initial treatment option for most people with LDHR. Relevant clinical guidelines by the Health Council of the Netherlands¹³ and the North American Spine Society¹⁴ support the use of conservative interventions as the primary approach to managing the early stages of LDHR in the absence of cauda equine syndrome. Data obtained from several countries suggest that this recommendation is followed in clinical practice. A report from the Australian Institute of Health and Welfare indicated that general practitioners in Australia were more likely to refer people with LDHR to a physical therapist than to an orthopedic surgeon.¹⁵ A survey of general practitioners in the Netherlands revealed that they referred 49% of their patients with LDHR to physical therapy.¹⁶ In the United States, over 60% of the participants with LDHR in the Spine Patients Outcomes Research Trial had received physical therapy before enrollment, whereas around 40% had received chiropractic treatment.¹⁷

Despite being the preferred initial management method, the efficacy of many conservative treatments for LDHR remains unclear.¹⁸ Although systematic reviews have collated the published evidence on conservative

From the *Musculoskeletal Research Centre, School of Physiotherapy, LA Trobe University, Bundoora, Australia; and †Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Australia. Acknowledgment date: May 13, 2009. Revision date: October 25, 2009. Acceptance date: October 26, 2009.

The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.

No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Supported by a National Health and Medical Research Council (Australia) Public Health Postgraduate Scholarship (to A.J.H.).

Address correspondence and reprint requests to Andrew John Hahne, School of Physiotherapy, LA Trobe University, Bundoora VIC 3086, Australia; E-mail: andrewhahne@gmail.com

treatments for sciatica up until 1998,¹⁹ for lumbosacral radicular syndrome up until 2004,²⁰ and for herniated lumbar discs up until 2006,²¹ they failed to identify any treatments that were consistently efficacious. One potential explanation for the failure of systematic reviews and clinical trials to consistently identify treatment effects relates to clinical heterogeneity among included trials and their respective participants.^{22,23} None of the systematic reviews cited above required imaging to confirm the cause of the participant's symptoms; hence, it is possible that some trials may have included people with conditions other than LDHR, such as spinal stenosis or spondylolisthesis. Because these other potential causes of radiculopathy are considered to respond less favorably to treatment than LDHR,^{9,24} it is plausible that the inclusion of participants with different pathologies may have diluted the obtained treatment effects in both the systematic reviews and the original trials. One way of reducing the confounding influences of clinical heterogeneity is to conduct a systematic review involving only those trials that recruited participants with both clinical signs and radiologic evidence of LDHR.

A radiologic diagnosis of LDH is best made with computed tomography (CT) or magnetic resonance imaging (MRI).²⁵ Although these imaging techniques are not considered necessary for all people with low back pain, they are appropriate for people who have radiculopathy that does not resolve after around 6 weeks.^{25,26} An advantage of these imaging methods is their ability to visualize structural disorders that may potentially be causing the symptoms and signs of radiculopathy. The reliability of CT and MRI for detecting a LDH has been shown to be high.²⁵⁻²⁷ Depictions of a LDH on CT or MRI have been shown to correlate well with surgical exploration,²⁵ although no true gold standard exists for diagnosing this condition.²⁸ Although false-positive findings of LDHs are common with imaging techniques,^{25,26} the relevance of the findings are increased if they are related to the patient's presenting signs and symptoms.²⁵

Our aim was to conduct a systematic review evaluating the efficacy of conservative treatments for people with clinical and radiologic evidence of LDHR. A secondary aim was to determine any adverse effects reported in the randomized controlled trials (RCT) that were to be reviewed.

Materials and Methods

The methodology in this review was guided via published guidelines by the Cochrane Collaboration^{29,30} and the QUORUM statement.³¹

Criteria for Selecting Trials in This Review

Types of Trials. All full reports of RCTs were eligible to be included if they were published in English between January 1, 1971, and August 31, 2008. Trials published earlier than 1971 could not have been eligible because CT and MRI were not in use.³²

Types of Participants. Trials were included if they involved participants aged >18 years with referred leg symptoms, with or without low back pain, where at least 75% of the participants had confirmation of a LDH *via* CT or MRI. Trials verifying LDH only with myelography were excluded because disc herniations are not directly visualized with this technique.³³ The term "herniation" was defined as a localized displacement of disc material beyond the margins of the intervertebral disc space.² This included synonymous terms such as prolapse, protrusion, and sequestration, but disc bulging was not sufficient.² Trials including >25% of participants who had previously undergone surgery or who had symptoms likely attributable to bony or ligamentous spinal stenosis were excluded. Participants with symptoms of any duration could be included, but for subgrouping purposes, we categorized participants according to duration of symptoms, with acute symptoms defined as <6 weeks, subacute as 6 to 12 weeks, and chronic as >12 weeks.²⁹

Types of Interventions. Trials were included if at least 1 group of participants received a conservative intervention. For the purposes of this review, we defined a conservative intervention as one that did not involve penetration through the deep skin layers. Trials in which all groups received injection therapy or any type of surgical intervention were, therefore, excluded; however, acupuncture was considered a conservative treatment. This definition was chosen because previous reviews have already evaluated the literature relating to surgical⁵ and injection therapies⁶⁻⁸ for people with LDHR, and the majority of the trials in these reviews seemed to use radiologic imaging to determine the potential cause of the participants' symptoms.

Types of Outcome Measures. We only included trials that reported data relating to treatment efficacy or adverse events. The outcomes of interest in this review were (1) back specific functional status (*e.g.*, Oswestry, Roland-Morris), (2) pain intensity (*e.g.*, visual analog scale, numerical rating scale), (3) global measures of improvement (*e.g.*, percentage of participants recovering, overall improvement ratings), and (4) adverse events or complications potentially attributable to the interventions. The length of follow-up of outcomes were categorized as short-term (<3 months after randomization), intermediate (between 3 months and 1 year), or long-term (1 year or more).²⁹

Search Methods

The following methods were used for identifying trials meeting our inclusion criteria:

- a. Computer database search was undertaken for the period between 1971 and August 31, 2008, using Medline (Ovid), CINAHL (Ovid), EMBASE (Elsevier 1971-1987 and Ovid 1988-2008), PEDro, Current Contents (Ovid), Cochrane central register of controlled trials, Cochrane database of systematic reviews, AMED, ISI Web of Science, and Australasian Medical Index (Informit). The database search strategy used key words for the condition of interest combined with a sensitive search strategy for locating RCTs as recommended by the Cochrane Collaboration²⁹ and empirical studies investigating sensitive search strategies.³⁴⁻³⁶ Intervention-specific terms were not used to avoid biasing the search results toward particular types of treatments. The search terms used for Medline are shown in Appendix, and these were adapted where necessary for other database interfaces.

- b. Citation tracking was performed *via* the ISI Web of Science for all included trials.
- c. The reference lists of all included trials were searched.
- d. Past relevant systematic reviews were searched.^{19–21,37}

Trial Selection

Two reviewers (A.H. and J.F.) independently evaluated the eligibility of trials to be included in the review. Trials were firstly eliminated if both reviewers independently excluded them based on initial review of the title and abstract. Full-text copies of remaining articles were then obtained and evaluated independently. Subsequently, disagreement was resolved *via* discussion until consensus was achieved. A third reviewer (J.M.) was to be consulted for unresolved disagreements, but was not required for this purpose. The level of agreement between the 2 independent reviewers (before discussion) was estimated using the kappa statistic.³⁸

Methodologic Quality Assessment

The 2 reviewers independently assessed the methodologic quality of included trials using the PEDro scale.³⁹ This scale rates RCTs on 10 key methodologic criteria that were identified by research experts involved in a consensus study using the Delphi method.⁴⁰ It has demonstrated adequate reliability,³⁹ and has been shown to be a valid indicator of trial methodologic quality.⁴¹ In addition, 3 items on the PEDro scale have been shown to be capable of influencing the outcomes of trials: randomization,^{29,42,43} concealed allocation,^{44–46} and blinding of participants and outcome assessors.^{29,42–45} The items on the PEDro scale are shown in Table 1.

Assessment of Clinical Relevance

The 2 reviewers independently evaluated the clinical relevance of included trials using the 5 criteria recommended by the Cochrane Back Review Group.²⁹

1. Are the patients described in detail so that you can decide whether they are comparable to those that you see in your practice?
2. Are the interventions and treatment settings described well enough so that you can provide the same for your patients?

3. Were all clinically relevant outcomes measured and reported?
4. Is the size of the effect clinically important?
5. Are the likely treatment benefits worth the potential harms?

We extended the third criterion to require papers to comment on the reliability and validity of the outcome measures used because reporting on such properties is recommended in the revised CONSORT statement.⁴⁷

Data Extraction and Analysis

Information from each trial regarding the type and number of participants and the interventions used were independently extracted by the 2 reviewers and entered into standardized computer spreadsheets. For continuous data, treatment effects and 95% confidence intervals were calculated using the Hedges adjusted-g standardized mean difference (SMD).^{48,49} The SMD was chosen because it allows comparison of effect sizes between trials that use different outcome measures.^{30,50} SMDs were calculated using group mean scores and pooled standard deviations (SDs) at the follow-up time of interest. When these values were not reported, they were estimated from mean change scores, baseline SDs, median values,³⁰ and SDs derived from the standard error or range.⁵¹ Positive SMD values were used to indicate treatment effects favoring the primary conservative intervention group. SMD values of 0.2, 0.5, and 0.8 were considered to represent small, moderate, and large effect sizes, respectively.⁵² For dichotomous data, the relative risk (RR) and 95% confidence intervals were calculated.⁴⁹ RRs were standardized so that RRs >1 indicated an increased risk of the event occurring in the primary conservative intervention group relative to the comparison group.

Data extraction, methodologic quality ratings, and clinical relevance ratings were piloted by the 2 reviewers on 2 ineligible trials^{53,54} before commencement of the review.²⁹

Data Synthesis

Pooling of data *via* meta-analysis was planned in cases where at least 2 trials contained sufficiently similar participants (diagnostic criteria, baseline pain and function, and duration of symptoms), treatment methods, comparison interventions, outcome measures, methodologic quality, and length of follow-up. When clinically homogenous trials were identified, they were assessed for statistical heterogeneity, which was considered likely if *P*-values of <0.1 were obtained on the χ^2 test, or if the *I*² statistic was >25%.^{30,55} Trials that were deemed to be both clinically and statistically homogenous were subjected to a fixed-effects model meta-analysis³⁰ using RevMan 4.2.⁵⁶

Where statistical pooling was deemed inappropriate because of clinical or statistical heterogeneity, effect sizes and 95% confidence intervals were reported for individual trials, and collation of results was limited to a narrative analysis using the levels of evidence criteria previously proposed by the Cochrane Collaboration²⁹:

- Strong evidence: consistent findings among multiple high-quality RCTs
- Moderate evidence: consistent findings among multiple low-quality RCTs or 1 high-quality RCT
- Limited evidence: one low-quality RCT
- Conflicting evidence: inconsistent findings among multiple RCTs
- No evidence from trials: no RCTs

Table 1. Methodologic Criteria Rated on the PEDro Scale

Item	Description
1	Were eligibility criteria specified?
2	Were participants randomly allocated to groups?
3	Was allocation concealed?
4	Were the groups similar at baseline regarding the most important prognostic indicators?
5	Were all participants blinded?
6	Was there blinding of all therapists who administered the therapy?
7	Was there blinding of all assessors who measured at least 1 key outcome?
8	Were measures of at least 1 key outcome obtained from >85% of the participants initially allocated to groups?
9	Did all subjects for whom outcome measures were available receive the treatment or control condition as allocated or, where this was not the case, were data for at least 1 key outcome analyzed by "intention to treat"?
10	Were the results of between-group statistical comparisons reported for at least 1 key outcome?
11	Did the study provide both point measures and measures of variability for at least 1 key outcome?

Note: Only items 2 to 11 are included in the calculation of the PEDro score.

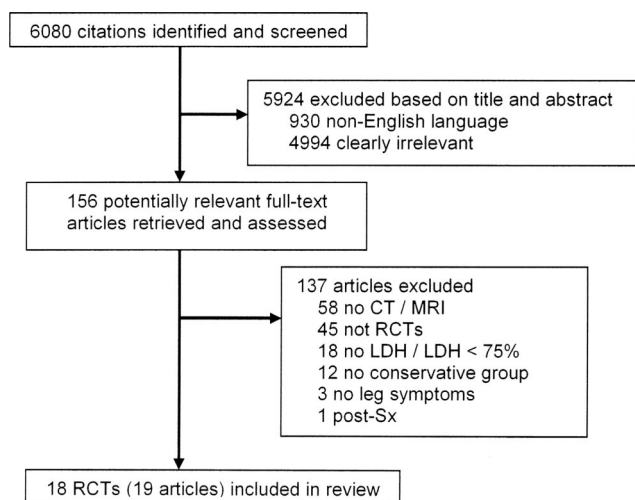


Figure 1. Flow chart showing progression of trials through the selection process.

The levels of evidence approach has been widely used to allow collation of results in reviews where meta-analysis is not possible.^{20,57–61} We noted that the Cochrane Back Review Group recently recommended the phasing in of the GRADE approach for collating bodies of evidence in systematic reviews.⁶² There is considerable similarity between these 2 approaches, because they both use information about the number of trials, the methodologic quality of the trials, the significance of the treatment effect, and the consistency of results across studies, to develop summary statements for different comparisons. We chose to continue using the levels of evidence approach in this review rather than switching to the newly recommended GRADE approach because our review process had already commenced, and we felt that it was unlikely that different conclusions would result if the GRADE approach were used. We defined contradictory evidence as <75% of trials agreeing.^{57,59} We considered a PEDro methodologic score of 6 or more out of 10 to represent high quality, in accordance with previous reviews.^{63,64}

■ Results

Selection of Trials

The number of trials considered at each stage of the review is outlined in Figure 1. The agreement between reviewers for selection of trials for stage 1 (title and abstract review) was 98.2%, with a kappa value of 0.62 (95% CI: 0.54–0.69), and 91.7% with a kappa value of 0.69 (95% CI: 0.53–0.84) for stage 2 (review of full text). This indicated substantial agreement during both stages.³⁸

Description of Trials

A total of 19 articles reporting on 18 RCTs randomizing 1671 participants to groups were ultimately included in the review.^{65–83} One trial included only participants with acute symptoms,⁷⁹ 2 trials included only subacute participants,^{75,78} 3 trials included only chronic participants,^{66,80,83} and all other trials included participants with mixed symptom durations.^{65,67–74,76,81,82}

Seventeen trials presented short-term follow-up data,^{65–76,78–82} 8 reported intermediate term follow-up

data,^{67,70,71,75,78,79,82,83} and 4 trials contained long-term follow-up data.^{68,75,78,83} A wide range of outcome measures were used, but all trials included at least 1 measure of pain or global change, and 10 included measures of function.^{66–68,70,72,75,78,80,82,83} The most common measure of pain intensity was the visual analogue scale (13 trials),^{65–67,70,72,73,75,76,78–80,82,83} whereas the most common measure of function was the Oswestry (4 trials).^{70,75,82,83} In 6 trials, a conservative intervention was compared to surgery or injections^{67,68,70,75,78,83}; hence, the group receiving conservative treatment in these trials may have been considered as a control group rather than a primary intervention. The characteristics of the included trials are outlined in Table 2.

Methodologic Quality and Clinical Relevance of Trials

Ratings for all trials on each item of the PEDro scale and clinical relevance scale are presented in Table 3. The agreement between reviewers on PEDro scale items was 85.4%, with a kappa value of 0.70 (95% CI: 0.60–0.80), indicating substantial agreement. For ratings on the clinical significance scale, the inter-rater agreement was 60.0%, kappa = 0.43 (95% CI: 0.27–0.59), indicating moderate agreement.

The mean PEDro methodologic score was 5.4 out of 10, with a range of 4 to 8. Seven (39%) trials were considered high quality based on achieving a PEDro score of 6 or more.^{65,67,75,76,78,79,82} Common methodologic limitations were failure to blind treating therapists (all 18 trials), failure to blind participants (17 trials^{65–76,78,80–83}), failure to report an intention to treat analyses (13 trials^{66,68–72,74–76,80–83}), and inadequate concealment of treatment allocation (13 trials^{66,68–74,76,80–83}). The mean clinical relevance score was 2.2, with all trials failing to comment on the validity or reliability of their chosen outcome measures.

Evidence for Efficacy of Interventions

Inter-rater agreement for extraction of means and SDs was 98.2%. The treatment effect sizes and associated 95% confidence intervals are presented in Table 4 and Figure 2. The levels of evidence summaries are presented in Table 5.

Advice

Two high-quality trials compared advice with microdiscectomy surgery in participants with subacute LDHR.^{75,78} These trials were deemed clinically homogenous, and tests for statistical heterogeneity were negative for outcomes measured at 6 to 8 weeks, 6 months, and 12 months. Meta-analysis was, therefore, performed for these follow-up periods using a fixed-effect model. The pooled SMD for back pain intensity was –0.4 (95% CI: –0.6 to –0.2) at short-term follow-up, indicating a statistically significant effect favoring surgery over advice. However, the pooled SMD values for intermediate and long-term back pain intensity were –0.12 (95% CI: –0.3 to 0.1) and –0.1 (95% CI: –0.3 to 0.2), respec-

Table 2. Characteristics of Included Trials

Trial	Participants	Primary Conservative Intervention	Comparison Intervention
Bakhtiyari <i>et al</i> ⁶⁵ (PEDro = 7, clinical = 4)	Iran, n = 60 (mean age 32.8, mean symptom duration 4 mo), referred to a physical therapy department with >2 mo of LBP, sciatic pain, and reduced functional performance. MRI or CT showing L4–L5 or L5–S1 disc herniation	Four-week lumbar stabilizing exercise program (1 time per week supervised by physical therapist, 2 times per day at home)	Four-week no treatment period before cross-over to intervention
Bonaiuti <i>et al</i> ⁶⁶ (PEDro = 4, clinical = 1)	Italy, n = 64, (mean age 48.9, mean symptom duration 14 mo), presenting to a hospital with >4 mo of LBP or “sciatalgia” (85% had referred leg pain), failed standard care, objective signs of benign LBP or sciatalgia, normal neurologic findings, CT or MRI showing disc herniation or protrusion	Manual autotractor for 45 min, 3 times per week for 2 wk	Natchev autotractor for 45 min, 3 times per week for 2 wk
Buchner <i>et al</i> ⁶⁷ (PEDro = 7, clinical = 1)	Germany, n = 36 (mean age 34.4, median symptom duration 8 wk), admitted to hospital with clinical symptoms of nerve root compression such as radicular pain below the knee, positive SLR <60°. MRI confirming a “lumbar nucleus pulposus prolapse” of at least 5 mm	Bed rest, analgesics, NSAIDs, hydrotherapy, electrotherapy, postural exercise classes (back school), soft tissue massage, joint mobilization, stabilization program, dynamic and static strengthening exercises. Dosage not reported	Primary conservative intervention plus 3 times lumbar epidural injections during the first 14 days in hospital
Burton <i>et al</i> ⁶⁸ (PEDro = 4, clinical = 2)	England, n = 40 (mean age 41.9, mean symptom duration 31 wk), from the orthopedic department of a hospital. Unilateral, unremitting sciatica (leg pain worse than back pain), positive SLR with positive tension signs, radiculopathy limited to a single nerve root, unequivocal evidence of a nonsequestered LDH on CT or MRI	Osteopathic sessions, including soft tissue techniques, mobilization, and high-velocity thrust manipulations. Mean of 11 (range 6–18) sessions × 15 min per session	Chemoneucleolysis
Deli ⁶⁹ (PEDro = 4, clinical = 1)	China, n = 147 (age range 17–53, duration of symptoms range 6 d to 17 yr), with back and leg symptoms (pain, weakness, or heaviness), prolapsed intervertebral disc confirmed <i>via</i> CT or MRI	Oral herbal medications once per day for 30 days plus the control group treatment	“Tuina massotherapy”: mobilization, massage, and mechanical traction, 1–2 times per week for a total of 30 sessions
Dincer <i>et al</i> ⁷⁰ (PEDro = 5, clinical = 2)	Turkey, n = 64 (mean age 28.4, symptom duration 1–3 mo) presenting to a military hospital with LBP and radicular pain below the knee, symptoms 1–3 mo duration, at least one nerve root compression sign (radicular pain with SLR, distal parasthesia, sensory deficit, motor deficit, reflex deficit), VAS pain score >4/10, MRI showing lumbar disc protrusion contained by the anulus and posterior longitudinal ligament	Diclofenac sodium 75 mg, prescribed to be taken orally twice per day for 14 days, also advised to perform lumbopelvic mobilization and lumbar stabilization exercises daily	Single caudal epidural injection, also advised to perform lumbopelvic mobilization and lumbar stabilization exercises daily
Guvencol <i>et al</i> ⁷¹ (PEDro = 4, clinical = 1)	Turkey, n = 30 (mean age 36.7, mean symptom duration 33.9 mo), low back pain and lower extremity pain, >1 mo duration, LDH diagnosed <i>via</i> CT	Inverted traction daily for 10 days, 5–10 min traction, 15 min infrared, isometric abdominal and gluteal exercises, bed rest	Mechanical traction 5–10 min daily for 10 days, plus 15 min infrared, isometric abdominal and gluteal exercises, bed rest
He <i>et al</i> ⁷² (PEDro = 5, clinical = 2)	China, n = 60 (mean age 42.6, symptom duration range 2 days to 12 yr) outpatients and inpatients of a hospital, aged 18–70, first attack or acute stage of a recurrent attack, 2 of the following features: (1) LBP with radicular pain aggravated with increased abdominal pressure, (2) local tenderness on vertebral palpation, pain radiating to the leg or foot, or scoliosis, (3) limited lumbar flexion and positive SLR or femoral nerve stretch test, (4) 2 of 4 neurologic signs: muscular atrophy, reduced myotome, abnormal reflex, sensory disturbance. LDH shown on CT or MRI	Herbal magnetic corset (worn all day, and lay on at night) plus the control group treatment, for 4 wk	Traction, electrotherapy, and massage daily for 4 wk
Kanayama <i>et al</i> ⁷³ (PEDro = 5, clinical = 1)	Japan, n = 40 (mean age 32.7), presenting to a hospital, with low back and sciatic symptoms, MRI revealing herniated nucleus pulposus	Sarpogrelate hydroxychloride orally 300 mg/day for 2 wk	Diclofenac sodium orally 75 mg/day for 2 wk
Liu and Zhang ⁷⁴ (PEDro = 4, clinical = 4)	China, n = 112 (“majority” aged 30–50, symptom duration range 2 hr to 20 yr), age 18–75, lumbar and radiating buttock/leg pain increased by coughing or sneezing, history of trauma or chronic muscle strain before development of LDH, positive Lasague sign, location and degree of protrusion shown by CT scan	“Pulling and turning manipulations”: massage, mobilization, and manipulation. Dosage not stated	Mechanical traction daily for 40–50 min, for 4 wk

(Continued)

Table 2. Continued

Trial	Participants	Primary Conservative Intervention	Comparison Intervention
Osterman <i>et al</i> ⁷⁵ (PEDro = 7, clinical = 2)	Finland, n = 56 (mean age 37.5, mean symptom duration 68.5 days), referred for orthopedic consultation at 1 of 4 hospitals, aged 20–50 with 6–12 wk of radicular pain below the knee, and 1 of the following findings: positive SLR <70°, muscle weakness, altered reflexes, or dermatomal sensory change. Disc extrusion or sequester on CT	Physical therapy instructions to stretch, bend, and perform isometric strengthening exercises (at baseline assessment), encouragement of activity at the 3 follow-up visits	Microdiscectomy within 2 wk of randomization. Isometric exercises before and after surgery. Postoperative physical therapy included active instructions to stretch and bend, strength exercises
Ozturk <i>et al</i> ⁷⁶ (PEDro = 6, clinical = 4)	Turkey, n = 46 (mean age 46.2) hospitalized because of <6 mo of low back pain or sciatica (91% had sciatica), L3–S1 radiculopathy, consistency between pain patterns, neurologic examination and radiologic findings, no history of previous physical therapy, LDH verified by CT	Continuous mechanical lumbar traction for 15 min (15 sessions, 1 session per week for 3 wk) plus the electrotherapy and medication that the control group received	Electrotherapy modalities: 15 min hot pack, 5 min ultrasound, 10 min diadynamic currents (15 sessions during 3 wk). Medication: ibuprofen 400 mg 3 times per day, mephenoxalone 200 mg 3 times per day, paracetamol 450 mg 3 times per day
Peul <i>et al</i> ^{77,78} (PEDro = 7, clinical = 1)	Netherlands, n = 283 (mean age 42.6, mean duration of symptoms 9.5 mo) referred from GP's to 1 of 9 hospitals, aged 18–85 yr with incapacitating lumbosacral radicular syndrome for 6–12 wk. All had MRI confirming a LDH	General practitioner care: advice regarding good prognosis, encouragement to return to daily activities and remain active, medication if required, physical therapy if fearful of moving	Microdiscectomy within 2 wk of randomization, followed by postoperative home-based rehabilitation supervised by a physical therapist
Santilli <i>et al</i> ⁷⁹ (PEDro = 8, clinical = 4)	Italy, n = 102 consecutive patients presenting to 2 medical rehabilitation centers with acute LBP (<10 days), >5/10 on a VAS (as evoked by palpation), radiating pain >5/10 (during SLR or femoral stretch). MRI showing disc protrusion with intact anulus	Active manipulation: soft tissue manipulations and brisk rotational thrust. Up to 20 sessions over 30 days, 5 min per session	Simulated manipulation: soft tissue muscle pressing with no thrusting. Up to 20 sessions over 30 days, 5 min per session
Sherry <i>et al</i> ⁸⁰ (PEDro = 5, clinical = 4)	Australia, n = 44 (mean age 42.0, mean duration of symptoms 7.3 yr), responding to newspaper advertising, aged 18–65, chronic LBP (>3 mo) and leg pain with minimum VAS of 2/10, living within 45 min of a treatment clinic, able to follow protocol, disc protrusion or herniation on CT or MRI	Vertebral axial decompression (VAX-D) therapy: 5 times per week for 4 wk, then weekly for 4 wk, 30 min per session	TENS: 30 min daily for 20 days, then weekly for 4 wk
Tesio and Merlo ⁸¹ (PEDro = 4, clinical = 4)	Italy, n = 44 (mean age 44.6, median symptom duration 1 yr), selected from outpatient department of a hospital rehabilitation unit with unremitting LBP ± radiating pain (75%) along a lumbosacral root distribution, duration >1 mo, failure of conservative approaches, consistency between pain pattern, neurologic findings and radiologic findings, disc herniation or protrusion on CT or MRI	Autotractor: 3 sessions of traction with patient generating own traction force, 30–60 min sessions every 2–3 days	Passive mechanical traction: 5 sessions on a daily basis, lasting 45 min per session
Unlu <i>et al</i> ⁸² (PEDro = 6, clinical = 2)	Turkey, n = 60 (mean age = 44.5, mean symptom duration = 44.6 days) consecutive patients presenting to a medical facility aged 20–60, with acute low back and leg pain (sciatica or femoral neuralgia) of <3 mo duration. MRI showing herniation of 1 or more discs consistent with the pain complaints and neurologic examination findings	Traction (mechanical): 5 days per week for 3 wk, 15 min sessions	(1) Ultrasound: 5 days per week for 3 wk, 8 min at 1.5 W/cm ² ; (2) Low-powered laser: 5 sessions per week for 3 wk, 50 mV, wavelength 830 nm, 4 min per point
Veihelmann <i>et al</i> ⁸³ (PEDro = 5, clinical = 0)	Germany, n = 99 (mean age = 43.6), with chronic LBP and radicular leg pain. MRI confirmed nerve root compression by either a LDH (87%) or scar tissue from previous surgery (13%)	“Conservative treatment with physical therapy” not described further	Epidural neuroplasty: local anesthetic, steroid, and saline injected into the LDH or scar tissue adjacent to the nerve

MRI indicates magnetic resonance imaging; CT, computed tomography; LDH, lumbar disc herniation; LBP, low back pain; VAS, visual analog scale; TENS, transcutaneous electric nerve stimulation; NSAIDs, nonsteroidal anti-inflammatory drugs; mg, milligrams; W/cm², watts per square centimeter; mV, millivolts; nm, nanometer; SLR, straight leg raise; GP, general medical practitioner.

Table 3. Ratings of Trials on the PEDro Methodologic Quality and Clinical Relevance Scales

Study	PEDro Score											Total	Clinical Relevance Score					Total
	1	2	3	4	5	6	7	8	9	10	11		1	2	3	4	5	
Bakhtiari <i>et al</i> ⁶⁵	✓	✓	✓	—	—	—	✓	✓	✓	✓	✓	7	✓	✓	—	✓	✓	4
Bonaiuti <i>et al</i> ⁶⁶	✓	✓	—	✓	—	—	—	—	—	✓	✓	4	✓	✓	—	✓	✓	1
Buchner <i>et al</i> ⁶⁷	✓	✓	✓	✓	—	—	—	✓	✓	✓	✓	7	✓	—	—	—	—	1
Burton <i>et al</i> ⁶⁸	✓	—	—	—	—	—	✓	✓	—	✓	✓	4	✓	✓	—	—	—	2
Deli ⁶⁹	—	✓	—	—	—	—	—	✓	—	✓	✓	4	—	—	—	✓	—	1
Dincer <i>et al</i> ⁷⁰	✓	✓	—	—	—	—	✓	✓	—	✓	✓	5	—	—	—	✓	✓	2
Guvenol <i>et al</i> ⁷¹	✓	✓	—	—	—	—	—	✓	—	✓	✓	4	✓	—	—	—	—	1
He <i>et al</i> ⁷²	✓	✓	—	✓	—	—	—	✓	—	✓	✓	5	✓	✓	—	—	—	2
Kanayama <i>et al</i> ⁷³	✓	✓	—	—	—	—	—	✓	✓	✓	✓	5	✓	—	—	—	—	1
Liu and Zhang ⁷⁴	✓	✓	—	—	—	—	—	—	—	✓	✓	4	✓	✓	—	✓	✓	4
Osterman <i>et al</i> ⁷⁵	✓	✓	✓	✓	—	—	✓	✓	—	✓	✓	7	✓	—	—	—	—	1
Ozturk <i>et al</i> ⁷⁶	✓	✓	✓	✓	—	—	✓	✓	—	✓	✓	6	✓	✓	—	✓	✓	4
Peul <i>et al</i> ^{77,78}	✓	✓	✓	✓	—	—	—	✓	✓	✓	✓	7	✓	—	—	—	—	1
Santilli <i>et al</i> ⁷⁹	✓	✓	✓	—	✓	—	✓	✓	✓	✓	✓	8	✓	✓	—	✓	✓	4
Sherry <i>et al</i> ⁸⁰	✓	✓	—	✓	—	—	—	✓	—	✓	✓	5	✓	✓	—	✓	✓	4
Tesio and Merlo ⁸¹	✓	✓	—	—	—	—	—	✓	—	✓	✓	4	✓	✓	—	✓	✓	4
Unlu <i>et al</i> ⁸²	✓	✓	—	✓	—	—	✓	✓	—	✓	✓	6	✓	✓	—	—	—	2
Veihelmann <i>et al</i> ⁸³	✓	✓	—	—	—	—	—	✓	—	✓	✓	5	—	—	—	—	—	0
Total	17	17	5	8	1	0	8	17	5	18	18		15	9	0	8	8	

tively, indicating no statistically significant difference between groups at these later follow-ups.

For leg pain intensity, pooled SMDs of -0.7 (95% CI: -1.0 to -0.5) and -0.3 (95% CI: -0.5 to -0.1) were obtained for short-term and intermediate-term follow-ups respectively, indicating significant effects in favor of surgery over advice. The long-term outcomes revealed no statistically significant difference between groups, with a pooled SMD of 0.0 (95% CI: -0.3 to 0.2). The actual mean leg pain scores at 12-month follow-up were 9/100 for the advice group and 6/100 for the surgery group in the Osterman *et al*⁷⁵ study, and 11/100 for both groups in the Peul *et al*⁷⁸ study.

For function, there was a statistically significant effect favoring surgery compared with advice at short-term follow-up, with a fixed effect model revealing a pooled SMD of -0.5 (95% CI: -0.7 to -0.3). Intermediate and long-term outcomes revealed no statistically significant differences between surgery and advice, with pooled SMD values of -0.2 (95% CI: -0.4 to 0.1) and -0.1 (95% CI: -0.3 to 0.2), respectively.

For global recovery, a statistically significant pooled RR of 0.4 (95% CI: 0.4 – 0.6) was obtained at short-term follow-up, indicating that participants receiving advice were significantly less likely than the surgical group to achieve a full or almost-full recovery at short-term follow-up. The difference between groups on global recovery was not maintained at intermediate and long-term follow-ups, with pooled RR values of 0.9 (95% CI: 0.8 – 1.1) and 1.0 (95% CI: 0.9 – 1.1), respectively. In terms of overall global outcomes at 12 months, the Peul *et al*⁷⁸ study reported that 83% of the advice group and 86% of the surgery group described a full or almost-full recovery at the 12-month follow-up, whereas in the Osterman *et al*⁷⁵ study, a full recovery was reported by 18% of the advice group and 25% of the surgery group at 12 months.

Follow-up data at 24 months failed at least 1 statistical test of heterogeneity for each outcome; hence, results were pooled *via* a narrative analysis. This provided strong evidence (2 high-quality trials,^{75,77} N = 316) that there is no difference between advice and microdiscectomy surgery for the long-term (24-month) outcomes of back pain intensity, leg pain intensity, or function, in people with subacute LDHR.

Medication

Two trials contained a group of participants who received oral diclofenac.^{70,73} The results of 1 of these trials provided limited evidence (1 low-quality trial,⁷⁰ N = 64) that diclofenac is less effective than caudal epidural injection for reducing pain intensity at short- and intermediate-term follow-ups, and for reducing function at short-term follow-up. The same trial provided limited evidence that there is no difference in function measured at intermediate-term follow-up between diclofenac and caudal epidural injection. The other trial provided limited evidence (1 low-quality trial,⁷³ N = 40) that there is no difference between diclofenac and sarpogrelate hydrochloride for short-term outcomes of back pain intensity or leg pain intensity. There was also limited evidence (1 low-quality trial,⁶⁹ N = 147) that the addition of oral herbal medication to “tunia massotherapy” (massage, mobilization, and mechanical traction) provided additional benefit in short-term global improvement.

Traction

Although 9 trials included traction as a component of treatment in at least 1 treatment group,^{66,69,71,72,74,76,80–82} only 7 of these trials allowed the effect of traction to be determined.^{66,71,74,76,80–82} Three trials compared 1 type of traction with another type.^{66,71,81} One low-quality trial⁸¹ (N = 44) provided limited evidence that manual autotraction is more effective than passive mechanical traction for providing global patient-perceived improve-

Table 4. Results of Comparisons of Treatment Effects for All Included Trials

Comparison	Back Pain/Overall Pain, SMD (95% CI)	Leg Pain, SMD (95% CI)	Function, SMD (95% CI)	Global Effect, RR (95% CI)
Stabilization vs. waiting list (Bakhtiaryet al ⁶⁵)	VAS			
4 wk	2.7 (2.0 to 3.3)			
Manual autotractor vs. Natchev autotractor (Bonaiuti et al ⁶⁶)	VAS		Backill	
2 wk	0.5 (−0.1 to 1.0)		0.0 (−0.6 to 0.5)	
Multimodal program vs. epidural plus multimodal program (Buchner et al ⁶⁷)	VAS		Hannover	Rating “good” or “very good”
2 wk	−0.4 (−1.1 to 0.2)		−0.4 (−1.0 to 0.3)	
6 wk	−0.2 (−0.9 to 0.4)		−0.2 (−0.8 to 0.5)	0.8 (0.6 to 1.2)
6 mo	−0.3 (−0.9 to 0.4)		−0.3 (−1.0 to 0.4)	
Manipulation (osteopathic) vs. chemonucleolysis (Burton et al ⁶⁸)	7-point thermometer	7-point thermometer	Roland Morris	
2 wk	0.7 (0.0 to 1.3)	0.0 (−0.6 to 0.7)	0.7 (0.1 to 1.3)	
6 wk	0.7 (0.0 to 1.3)	0.0 (−0.6 to 0.7)	0.5 (−0.2 to 1.2)	
12 mo	0.4 (−0.3 to 1.1)	0.1 (−0.6 to 0.8)	0.2 (−0.5 to 0.9)	
Oral herbal medication added to mobilization, massage, and mechanical traction (Deli ⁶⁹)				Rating “any improvement” vs. “failed treatment”
? 30 days				1.1 (1.0 to 1.3)
Diclofenac sodium orally vs. epidural injection (Dincer et al ⁷⁰)	VAS		Oswestry	
2 wk	−2.2 (−2.8 to −1.6)		−1.5 (−2.0 to −1.0)	
1 mo	−1.0 (−1.5 to −0.5)		−0.6 (−1.2 to −0.2)	
3 mo	−0.6 (−1.1 to −0.1)		−0.4 (−1.9 to 0.1)	
Inverted traction vs. mechanical traction (Guvencol et al ⁷¹)	0 to 10 scale			
10 d	Data not reported			
3 mo	Data not reported			
Herbal magnetic corset added to traction, electrotherapy, and massage (He et al ⁷²)	VAS		Lumbar disease grade	
1 wk	0.5 (0.0 to 1.0)		0.3 (−0.2 to 0.8)	
2 wk	0.7 (0.1 to 1.2)		0.8 (0.2 to 1.8)	
4 wk	0.8 (0.2 to 1.3)		1.0 (0.4 to 1.5)	
Sarpogrelate hydroxychloride vs. diclofenac sodium (Kanayama et al ⁷³)	VAS	VAS		
2 wk	0.0 (−0.6 to 0.7)	0.5 (−0.1 to 1.1)		
Manipulation vs. mechanical traction (Liu and Zhang ⁷⁴)				Rating “cured or improved” vs. “ineffective”
? 5 wk				1.3 (1.1 to 1.6)
Advice vs. microdiscectomy (Osterman et al ⁷⁵)	VAS	VAS	Oswestry	Rating “full recovery” vs. “not full recovery”
6 wk	−0.3 (−0.8 to 0.2)	−0.5 (−1.1 to −0.0)	−0.4 (−0.9 to 0.2)	0.1 (0.0 to 1.6)
3 mo	−0.3 (−0.9 to 0.2)	−0.3 (−0.9 to 0.2)	−0.5 (−1.0 to 0.1)	0.8 (0.2 to 2.7)
6 mo	−0.3 (−0.8 to 0.3)	−0.4 (−0.9 to 0.2)	−0.3 (−0.8 to 0.2)	
12 mo	0.1 (−0.4 to 0.6)	−0.2 (−0.7 to 0.3)	−0.1 (−0.6 to 0.5)	0.7 (0.3 to 2.0)
24 mo	−0.4 (−1.0 to 0.1)	−0.5 (−1.0 to 0.0)	−0.4 (−0.9 to 0.2)	
Mechanical traction added to electrotherapy modalities and medication (Ozturk et al ⁷⁶)	VAS			Back pain “present” vs. “absent”
? 3 wk	0.5 (−0.1 to 1.1)			0.8 (0.3 to 1.8)
				Sciatica “present” vs. “absent”
				0.5 (0.2 to 0.9)
Advice vs. microdiscectomy (Peul et al ^{77,78})	VAS	VAS	Roland	7-point global rating
2 wk	−0.1 (−0.3 to 0.2)	−0.7 (−0.9 to −0.5)	0.2 (0.0 to 0.5)	−0.3 (−0.6 to −0.1)
8 wk	−0.5 (−0.7 to −0.2)	−0.8 (−1.0 to −0.6)	−0.5 (−0.8 to −0.3)	
6 mo	−0.1 (−0.3 to 0.1)	−0.3 (−0.5 to 0.0)	−0.1 (−0.4 to 0.1)	−0.8 (−1.0 to −0.5)
12 mo	−0.1 (−0.3 to 0.1)	0.0 (−0.2 to 0.2)	−0.1 (−0.3 to 0.2)	−0.2 (−0.4 to 0.1)
24 mo	−0.1 (−0.3 to 0.2)	0.1 (−0.2 to 0.3)	−0.1 (−0.2 to 0.3)	−0.2 (−0.4 to 0.1)
				Rating “full or near-full” recovery vs. “not full recovery”
2 wk				0.5 (0.4 to 0.6)
6 mo				0.9 (0.8 to 1.1)
12 mo				1.0 (0.9 to 1.1)
24 mo				1.0 (0.9 to 1.1)

(Continued)

Table 4. Continued

Comparison	Back Pain/Overall Pain, SMD (95% CI)	Leg Pain, SMD (95% CI)	Function, SMD (95% CI)	Global Effect, RR (95% CI)
Manipulation vs. sham manipulation (Santilli <i>et al</i> ⁷⁹)				
	VAS	VAS		Free of back pain
2 wk	0.5 (0.1 to 0.9)	0.4 (0.0 to 0.8)		N/A
4 wk	0.5 (0.1 to 0.9)	0.7 (0.3 to 1.1)		6.5 (0.3 to 122.4)
6 wk	0.8 (0.4 to 1.2)	0.7 (0.3 to 1.1)		3.0 (0.9 to 10.4)
3 mo	0.9 (0.5 to 1.3)	0.9 (0.5 to 1.3)		4.3 (1.3 to 14.2)
6 mo	0.7 (0.3 to 1.2)	0.7 (0.3 to 1.1)		5.0 (1.5 to 16.2)
				Free of leg pain
2 wk				3.4 (0.7 to 15.4)
4 wk				2.0 (0.8 to 4.9)
6 wk				2.6 (1.3 to 5.2)
3 mo				4.8 (2.2 to 10.6)
6 mo				2.9 (1.6 to 5.3)
Vertebral axial decompression vs. transcutaneous electric nerve stimulation (Sherry <i>et al</i> ⁸⁰)				
	VAS		4 self-selected items on 4-point scale	Rating "successful case"
? 8 wk	2.6 (0.4 to 1.8)		1.5 (0.8 to 2.2)	29.7 (1.9 to 467.9)
Autotraction vs. mechanical traction (Tesio and Merlo ⁸¹)				
				Rating "improved" vs. "not improved"
1 to 2 wk				4.3 (1.7 to 10.6)
Traction vs. ultrasound (Unlu <i>et al</i> ⁸²)*				
	Back pain (VAS)	VAS	Roland Morris	
3 wk	0.0 (-0.6 to 0.6)	0.1 (-0.5 to 0.7)	-0.1 (-0.7 to 0.5)	
2 mo	0.1 (-0.5 to 0.7)	0.3 (-0.3 to 0.9)	-0.1 (-0.7 to 0.6)	
4 mo	-0.3 (-0.9 to 0.4)	-0.3 (-0.9 to 0.4)	-0.1 (-0.7 to 0.6)	
			Oswestry	
3 wk			0.0 (-0.7 to 0.6)	
2 mo			0.1 (-0.5 to 0.8)	
4 mo			-0.1 (-0.7 to 0.5)	
Physical therapy vs. epidural neuroplasty (Veihelmann <i>et al</i> ⁸³)				
	VAS	VAS	Oswestry	
3 mo	-1.4 (-1.9 to -1.0)	-1.4 (-1.8 to -0.9)	-1.0 (-1.5 to -0.6)	
6 mo	-1.5 (-2.0 to -1.0)	-1.6 (-2.1 to -1.1)	-1.5 (-2.0 to -1.0)	
12 mo	-1.1 (-1.6 to -0.7)	-1.2 (-1.7 to -0.7)	-1.0 (-1.5 to -0.6)	

Results in bold represent statistically significant comparisons based on the 95% confidence interval of the SMD or relative risk.

*Comparisons for traction vs. laser and ultrasound vs. laser not presented, as SMDs and confidence intervals were similar for all groups.

SMD indicates hedges standardized mean difference; RR, relative risk; CI, confidence interval; VAX-D, vertebral axial decompression; TENS, transcutaneous electric nerve stimulation.

ment at short-term follow-up. One low-quality trial⁶⁶ (N = 64) provided limited evidence that there is no difference between manual autotraction and Nachev autotraction for short-term pain intensity and function in people with chronic LDHR. Another trial that compared inverted traction with mechanical traction did not report sufficient primary outcome data to allow SMDs and their associated confidence intervals to be calculated.⁷¹

Three trials compared traction with other treatments.^{74,80,82} One high-quality trial⁸² (N = 60) provided moderate evidence that there is no difference between mechanical traction and either ultrasound or laser for back pain intensity, leg pain intensity, or function at short- or intermediate-term follow-ups. One low-quality trial⁸⁰ (N = 44) provided limited evidence that vertebral axial decompression traction therapy is more effective than transcutaneous electrical nerve stimulation for intermediate-term outcomes of pain intensity, function, and risk of global treatment success, for people with chronic LDHR. One low-quality trial⁷⁴ (N = 112) provided limited evidence that mechanical traction is less

effective than "pulling and turning manipulations" for global ratings of improvement at short-term follow-up.

One high-quality trial⁷⁶ (N = 46) provided moderate evidence that the addition of mechanical traction to electrotherapy methods (hot pack, ultrasound, and diadynamic currents) and medication (ibuprofen, mephenoxalone, and paracetamol) reduces the risk of having sciatica at short-term follow-up, but provides no additional short-term benefit for pain intensity or risk of having low back pain.

Stabilization Exercises

One high-quality trial⁶⁵ (N = 60) provided moderate evidence that a stabilizing exercise program is more effective than no treatment for reducing pain intensity at short-term follow-up.

Physical Therapy

There is limited evidence (1 low-quality trial,⁸³ N = 99) that physical therapy is less effective than epidural neuroplasty for the intermediate- and long-term outcomes of

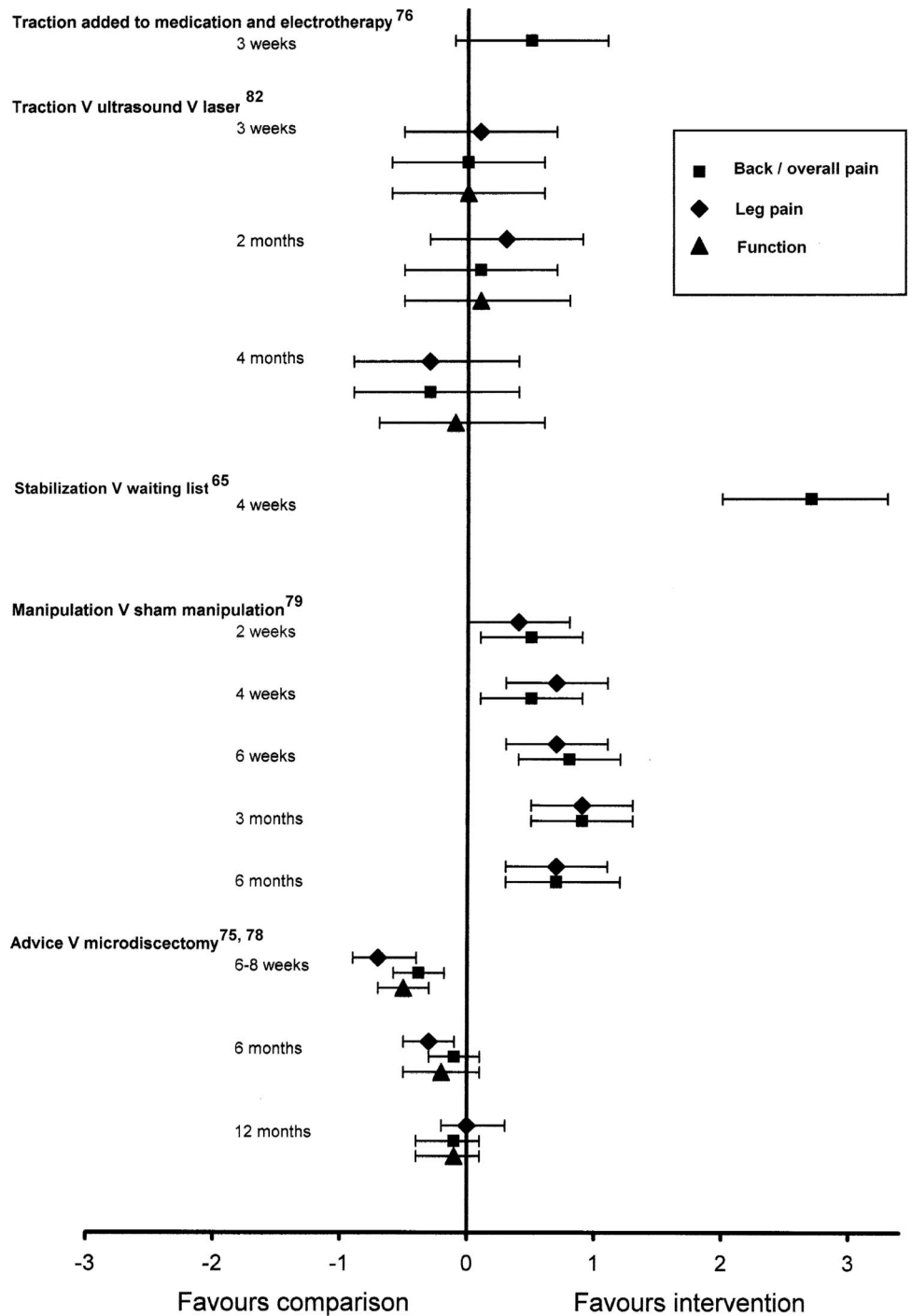


Figure 2. Forrest plot of standardized mean differences with 95% confidence intervals for all high-quality trials.

leg pain intensity, back pain intensity, and function, for people with chronic LDHR.

Manipulation

Three trials investigated the effect of manipulation, but all used different comparison interventions.^{68,74,79} In all of these trials, manipulation consisted of soft tissue manipulation or massage along with high-velocity rotational thrusts. One high-quality trial⁷⁹ (N = 102) provided moderate evidence that, in people with acute LDHR and an intact annulus, manipulation is more effective than simulated manipulation for the outcomes of

back pain intensity (short and intermediate follow-up), leg pain intensity (4 week, 6 week, and intermediate follow-up), risk of becoming free of back pain (intermediate-term follow-up), and risk of becoming free of leg pain (6 week and intermediate follow-ups). The same trial provided moderate evidence that there is no difference between active and simulated manipulation for leg pain intensity (2-week follow-up), risk of becoming free of back pain (short-term follow-up), and risk of becoming free of leg pain (2-week and 4-week follow-ups).

Table 5. Summary of Results According to the Levels of Evidence Criteria

Comparison	Symptom Duration	Level of Evidence	Global Rating	Pain (Back or Overall)	Pain (Leg)	Function
Advice						
Vs. microdiscectomy ^{75,77,78}	Subacute	High	-(S), =(I, L)	-(S), =(I, L)	-(S, I), =(L)	-(S), =(I, L)
Medication						
Diclofenac vs. epidural ⁷⁰	Mixed	Low		-(S, I)		-(S), =(I)
Diclofenac vs. sarpogrelate hydroxychloride ⁷³	Mixed	Low		=(S)	=(S)	
Herbal added to massage, mobilization, and traction ⁶⁹	Mixed	Low	+(S)			
Traction						
Auto vs. mechanical ⁸¹	Mixed	Low	+(S)			
Auto vs. Natchev ⁶⁶	Chronic	Low		=(S)		=(S)
Inverted vs. mechanical ⁷¹	Mixed	Low				
Mechanical vs. ultrasound vs. laser ⁸²	Mixed	Moderate		=(S, I)	=(S, I)	=(S, I)
VAX-D vs. TENS ⁸⁰	Chronic	Low	+(I)	+(I)		+(I)
Mechanical vs. manipulation ⁷⁴	Mixed	Low	-(S)			
Mechanical added to medication and electrotherapy modalities ⁷⁶	Mixed	Moderate		=(S)	+(S)	
Stabilization exercises						
Vs. waiting list control ⁶⁵	Mixed	Moderate		+(S)		
Physical therapy						
Vs. epidural neuroplasty ⁸³	Chronic	Low		-(I, L)	-(I, L)	-(I, L)
Manipulation						
Vs. sham manipulation ⁷⁹	Acute	Moderate		+(S, I)	+/+(S)* +(I)	
Vs. chemonucleolysis ⁶⁸	Mixed	Low		+(S), =(L)	=(S, L)	+/=(S)† =(L)
Vs. mechanical traction ⁷⁴	Mixed	Low	+(S)			
Laser and ultrasound						
Laser vs. ultrasound vs. traction ⁸²	Mixed	Moderate		=(S, I)	=(S, I)	=(S, I)
Corset						
Added to traction, electrotherapy, and massage ⁷²	Mixed	Low		+(S)		+(S)

+ indicates an effect in favor of the intervention over the comparison; -, an effect in favor of the comparison intervention; =, no significant difference between interventions; S, short-term follow-up; I, intermediate-term follow-up; L, long-term follow-up; VAX-D, vertebral axial decompression; TENS, transcutaneous electric nerve stimulation.

*There was no significant difference at the 2-week follow-up, but significant differences in favor of manipulation at the 4-week and 6-week follow-ups.

†There was a significant difference in favor of manipulation at the 2-week follow-up, and no significant difference at 6-week follow-up.

One low-quality trial⁶⁸ (N = 40) provided limited evidence that manipulation was more effective than chemonucleolysis for the outcomes of back pain intensity (short-term follow-up) and function (2-week follow-up), but that no difference exists between these treatments for the outcomes of leg pain intensity (short- and long-term follow-ups), back pain intensity (long-term follow-up), and function (6-week and long-term follow-ups). The other trial provided limited evidence (1 low-quality trial,⁷⁴ N = 112) that manipulation is more effective than mechanical traction in terms of short-term global improvement ratings.

Laser and Ultrasound

One high-quality trial⁸² (N = 60) provided moderate evidence that there is no difference between laser and mechanical traction, no difference between ultrasound and mechanical traction, and no difference between laser and ultrasound, for back pain intensity, leg pain intensity, or function at short- and intermediate-term follow-ups.

Corsets

One low-quality trial⁷² (N = 60) provided limited evidence that the addition of a herbal magnetic corset to a program of traction, electrotherapy, and massage pro-

vides additional benefits in short-term pain intensity and lumbar function compared with traction, electrotherapy, and massage alone.

Multimodal Inpatient Program

One high-quality trial⁶⁷ investigated the effect of adding 3 epidural injections to a multimodal inpatient treatment program consisting of bed rest, hydrotherapy, electrotherapy, back school, massage, mobilization, and exercises. There were no significant differences between groups on short- or intermediate-term outcomes of pain intensity, function, and subjective rating of outcome (Table 4). Because both groups received the same multimodal treatment program, it was not possible to determine the relative effectiveness of the conservative component of the treatment.

Adverse Events

Three trials reported at least 1 adverse event in conservative treatment groups.^{71,74,76} More than 16 adverse events were associated with traction treatment. In 1 trial comparing inverted traction with mechanical traction,⁷¹ 11 of the 16 participants in the inverted traction group reported pain associated with treatment compared with 2 of the 15 participants in the mechanical traction group;

RR = 5.2 (95% CI: 1.4–19.5). The same trial reported that “almost all” of the participants in the inverted traction group reported anxiety during treatment, and 1 developed lower limb muscle weakness after treatment, compared with no such events in the mechanical traction group. In 1 trial, 2 of the 50 participants receiving mechanical traction fainted.⁷⁴ In another trial, 2 of the 46 participants needed to cease ibuprofen medication because of gastrointestinal effects.⁷⁶

A further 4 trials reported that there were no adverse events associated with conservative treatment.^{68,70,72,73} Four trials reported adverse events associated with surgery or injections.^{70,75,78,83} Six trials made no mention of adverse events.^{65,66,69,79–82}

■ Discussion

Efficacy of Conservative Treatments for LDHR

This review of conservative management for LDHR revealed that a wide variety of different treatments have been evaluated in RCTs focusing on people with this condition. Conservative interventions included advice, medication, traction, manipulation, stabilization exercises, physical therapy, laser, ultrasound, and corsets. A wide range of comparison interventions were used in these trials, including sham manipulation, waiting lists, other conservative treatments, surgery, and injections. Most trials included participants with mixed symptom durations. A variety of outcome measurement tools have been used, and the length of follow-up varied from 1 week to 24 months. The wide variation in all of these variables made it difficult to collate the findings of multiple trials *via* meta-analysis or even by using the levels of evidence approach. Most of the evidence summaries were therefore derived from individual trials.

The only strong evidence to emerge from this review was obtained by collating the results of 2 clinically and statistically homogenous trials that compared advice with microdiscectomy in people with subacute LDHR.^{75,78} Our analysis indicated that advice is less effective than surgery for producing short-term improvements in back pain intensity, leg pain intensity, function, and global improvement. These differences were maintained at intermediate-term follow-up for leg pain intensity, but not for back pain intensity, function, or global change. There was strong evidence that no difference existed on any of these outcome measures at long-term follow-up. We noted that leg pain scores at 12-month follow-up were quite low in both groups for both trials (ranging from 6/100 to 11/100), indicating that the long-term prognosis was good regardless of the intervention received. In both trials, the advice group was a control intervention that was compared with the primary intervention of microdiscectomy. We found no trials that compared advice with other conservative interventions; thus, the relative efficacy of advice compared with other conservative interventions remains unclear. Other reviews of advice for the management of nonspecific low back pain (NSLBP) suggest that advice may be more

effective than several other conservative treatments.^{84,85} Further research comparing advice with other conservative interventions for people with LDHR, therefore, seems warranted.

Moderate strength evidence relating to several interventions was derived from individual high-quality trials. There was moderate evidence that stabilization exercises are more effective than no treatment for short-term improvement in pain intensity.⁶⁵ The evidence relating to stabilization exercises for LDHR is consistent with another recent review of the efficacy of motor control exercises for NSLBP.⁸⁶ It remains unclear whether stabilization exercises are more effective than other treatments for people with LDHR because no trials have investigated this. There was also moderate evidence that manipulation is more effective than sham manipulation for back and leg pain outcomes in people with acute LDHR with an intact anulus,⁷⁹ although the results may not be generalizable to people who have a ruptured anulus. Other systematic reviews on NSLBP have also shown that manipulation may be more effective than placebo.^{87,88} The other moderate strength evidence to emerge from this review showed that there is no difference between laser, mechanical traction, and ultrasound, and that the addition of mechanical traction to a treatment program involving electrotherapy methods and medication adds some benefits in terms of reducing the likelihood of sciatica being present at short-term follow-up. The evidence relating to all other trials in this review was rated as limited because of their low methodologic quality scores.

A high number of trials in this review showed no difference in several outcomes between groups. This may indicate that many of the intervention and comparison treatments were truly equivalent, or it may be that low statistical power to detect differences between groups resulted in type II errors. The latter explanation may be plausible in some cases because sample sizes were small in several trials included in this review. Caution should, therefore, be used when interpreting evidence summaries in this review that conclude that there is no difference between 2 interventions, as an alternative explanation may be that significant effects were missed because of low statistical power in the original trials.⁸⁹

Methodologic Limitations of Included Trials

We chose to use the PEDro scale to evaluate the methodologic quality, or risk of bias, of trials in this review because of its documented reliability and validity for measuring trial quality.^{39–41} In its most recent update on methodologic guidelines for systematic reviews, the Cochrane Back Review Group recommend a domain-based “risk of bias” assessment of trials, which considers potential sources of bias individually without a sum score.⁶² The items on the PEDro scale are very similar to those recommended by the Cochrane Back Review Group. Some concerns have been expressed about the validity of assigning equal weights to individual items on

scales such as PEDro and then summing the scores to achieve an overall method score for each trial,⁹⁰ but a study has recently validated this practice for the PEDro scale.⁴¹ When assessing the quality of trials in this review, however, we recommend that the individual criteria be considered for each trial (Table 3) in addition to the total PEDro score. We have also chosen to discuss some specific aspects of methodologic quality below to discourage sole reliance on the PEDro score when assessing the quality of trials in this review.

The most common methodologic limitation of trials included in this review was failure to blind participants and therapists. Only 1 trial in this review attempted to blind participants by comparing manipulation with sham manipulation,⁷⁹ and no trials attempted to blind therapists. Achieving adequate blinding of participants and therapists in trials of physical treatments is very difficult, and is more easily achieved in drug trials.⁹¹ Interestingly, the 1 trial in this review that compared 2 drugs⁷³ also failed to blind participants or therapists. Although it was more common for outcome assessments to be undertaken in a blinded manner, 10 of the 18 trials still failed to achieve this. Future trials of conservative treatments for LDHR should aim to blind outcome assessors at a minimum, even if blinding of therapists or participants is not possible.

Although the PEDro scale includes random allocation as a methodologic quality criterion, it only requires trials to mention random allocation to satisfy this criterion. However, only 9 trials described robust randomization methods,^{65–67,72,73,75,78–80,83} 7 failed to describe the method used,^{69–71,74,76,81,82} and in 1 trial, the planned randomization process was compromised for 15 of the 40 subjects because of an administrative error.⁶⁸ Failing to report the precise method of randomization raises questions as to whether a truly random method was applied in these trials, and this may have implications for the validity of their results.^{47,92}

Several trials that were considered high quality based on achieving a PEDro score of 6 or more still contained considerable methodologic flaws. Two trials in our review^{67,78} that both scored 7/10 on the PEDro scale failed to use blinded outcome assessment, which introduces significant potential for bias.⁴² Another high-quality trial failed to use concealed allocation,⁷⁶ which has been shown to lead to inflated estimates of treatment effects.^{43–46}

Clinical Significance of Included Trials

Although each of the items on the clinical significance rating scale seems to have merit, we found the scale to be lacking in standardized decision criteria. This may explain the lower levels of inter-rater agreement obtained for these ratings. Features of trials that may influence clinical relevance have been suggested elsewhere,^{47,93} and these may serve as a useful starting point to provide further standardization to the clinical significance scale. Despite the limitations of the existing scale, we noted

that no trials in our review mentioned the validity or reliability of their outcome measurement tools. Although some authors using common low back pain measurement tools, such as visual analog scales, may have assumed that these properties were widely accepted, the specific focus of the trials on people with LDHR warranted clarification of the validity and reliability of the tools for this population.

Adverse Events

No trials in our review described an intention or a methodology for detecting adverse events. This raises the possibility that other adverse events may have been overlooked or dismissed by the authors of the trials, particularly if some authors defined an adverse event differently from others. Authors of future RCTs should take note of the revised CONSORT statement, which recommends that adverse events be operationally defined and reported in all RCTs.⁴⁷

Adverse events reported by trials in this review were rare, but they were most commonly attributed to traction. Pain, anxiety, lower limb weakness, and fainting were all reported in trials using traction, although only pain associated with inverted traction produced a statistically significant RR.⁷¹ Gastrointestinal side effects are commonly associated with the use of nonsteroidal anti-inflammatory drugs.⁵⁸ Although 1 trial in this review reported gastrointestinal side effects associated with the use of ibuprofen,⁷⁶ the 2 trials that used diclofenac as a sole treatment reported no “serious” side effects,^{70,73} although it is not clear how they defined a serious side effect.

Although this review of RCTs allowed some indication of the frequency and nature of adverse effects, the limited sample sizes of the included studies reduces the precision of these estimates. When adverse events are rare, data from other sources, such as case series and case reports, can add to the body of evidence regarding adverse effects that may be associated with treatments.³⁰ Other reviews have collated data from case series and case reports relating to adverse events associated with manipulation.^{21,94,95} These reviews suggest that some concerns remain over the potential for manipulation to cause or exacerbate a LDH,^{21,94} although no adverse events related to manipulation were reported by the trials in our review.

Comparison With Other Reviews

This review differed from previous reviews that have been conducted on the efficacy of interventions for people with sciatica,¹⁹ lumbosacral radicular syndrome,²⁰ and herniated lumbar discs.²¹ Our focus was on a specific diagnostic group with clinical and radiologic evidence of LDHR, and we searched the literature up until August 31, 2008. None of the previous reviews had specific inclusion criteria for clarifying the presence of LDHR in participants. It is, therefore, likely that these other reviews included trials that contained participants with a more heterogeneous array of pathologies. One

example of this is seen from a trial that compared 4 different physiotherapy treatments for people with sciatic symptoms.⁹⁶ That trial was included in all previous reviews, but was excluded from ours. To be included in the trial, participants needed only to have “sciatic symptoms” as far as the gluteal fold, and no imaging was undertaken to confirm the potential source of the symptoms. These broad inclusion criteria would likely have resulted in a heterogeneous group of participants, and only a minority of them may have actually had LDHR as defined *via* typical diagnostic criteria.¹ The conclusions of our review might, therefore, be more specifically applicable to people with LDHR because of our tighter diagnostic inclusion criteria.

The different inclusion criteria and search strategies between our review and previous reviews seemed to result in a considerably different collection of trials. For example, only 2^{67,68} of the 18 trials in our review were included in the Luijsterburg *et al*²⁰ review. This seemed to be attributable to 12 of the trials in our review being published after the Luijsterburg *et al* review. Four other trials that we included in our review^{71,74,80,81} were not present in the Luijsterburg *et al* review, possibly because of the different search strategies used. Conversely, of the 30 trials included in the Luijsterburg *et al* review, 28 did not meet our inclusion criteria (in 13 trials all groups received injections, 11 did not use CT or MRI to confirm a LDH, 2 contained <75% of subjects with a confirmed LDH, and 2 were published in a non-English language).

The different collection of trials in the various reviews led to some differences in evidence summaries. Although our review found moderate evidence for the effectiveness of manipulation over sham manipulation for acute LDHR, and another review concluded that manipulation was likely to be beneficial,²¹ other reviews have found limited^{19,37} or no²⁰ evidence to support the use of manipulation for LDHR. Our review found mixed results among trials using traction, with moderate evidence of efficacy on 1 of the 3 outcome measures in 1 trial,⁷⁶ limited evidence of efficacy in 2 trials,^{80,81} no difference in 2 trials,^{66,82} and evidence of ineffectiveness in 1 trial.⁷⁴ The other reviews did not recommend traction for LDHR.^{19–21} Our review did not find any significant effects in trials using nonsteroidal anti-inflammatory drugs, and this finding was consistent with all other reviews.^{19–21} Only our review included the trials investigating advice, stabilization exercises, herbal magnetic corsets, herbal medication, laser, and ultrasound; hence, the other reviews did not comment on the efficacy of these treatments.

Limitations of This Review

Although we attempted to focus on participants with a specific pathology, we were unable to control all potential sources of variability among participants. We acknowledge that our minimum diagnostic criteria of leg pain with LDH on CT or MRI still leaves room for potential variation in diagnostic subgroup. In addition, al-

though we used a specific radiologic definition of LDH,² authors of the included trials may have used terms such as herniation or protrusion differently. Variability in the severity and nature of symptoms was also apparent among the participants of the included trials. For example, 2 trials excluded people with motor or sensory neurologic deficits,^{65,66} other trials contained a large proportion of participants with positive neurologic findings,^{67,75–77} and many other trials failed to report the proportion of participants with such features.^{68–70,72,74,79,80,83} It has been proposed that clinical heterogeneity within RCTs and systematic reviews may account for some of the null and inconclusive results that are prevalent in back pain research.^{22,23} Although reviews such as ours that attempt to focus on specific subgroups of low back pain are, therefore, important, authors of RCTs can assist these efforts by thoroughly reporting the key diagnostic features and measures of condition severity of their participants.

Another limitation of our review was the exclusion of trials published in languages other than English because of funding and resource limitations. Although there is some evidence to suggest that excluding non-English articles does not typically have an effect on systematic review results,⁹⁷ we are unable to determine this definitively. Publication bias is also a possibility in this review. However, because most interventions that we included showed only limited or moderate evidence of efficacy, including more trials with null results is unlikely to have significantly changed the conclusions.

Reviewers' Conclusion

This systematic review of RCTs involving people with clinical and radiologic evidence of LDHR provides strong evidence that advice is less effective than microdiscectomy at short-term follow-up, but equally effective at long-term follow-up, for people with subacute LDHR. There is moderate evidence that stabilization exercises are better than no treatment at short-term follow-up, that manipulation is better than sham manipulation at short- and intermediate-term follow-ups for people with acute LDHR and an intact anulus, and that no difference exists between traction, laser, and ultrasound at short- and intermediate-term follow-ups. Moderate evidence was found that the addition of mechanical traction to medication and electrotherapy methods reduces the risk of sciatica being present at short-term follow-up, but not the risk of back pain being present or mean pain intensity. There was either limited or no evidence to support the efficacy of manipulation compared with other treatments, traction compared with other treatments, physical therapy compared with neuroplasty, or for herbal medication, herbal magnetic corsets, or nonsteroidal anti-inflammatory medication. Two trials reported adverse events associated with traction (pain, anxiety, lower limb weakness, and fainting), whereas 1 trial reported gastrointestinal events associated with ibuprofen. Additional high-quality trials are required to determine which con-

servative treatments are the safest and most effective for people with LDHR.

■ Key Points

- There is strong evidence that advice is less effective than microdiscectomy surgery at short-term follow-up, but that no difference exists at long-term follow-up, for people with subacute lumbar disc herniation with associated radiculopathy (LDHR).
- For people with LDHR, there is moderate evidence that stabilization exercises are more effective than no treatment at short-term follow-up, moderate evidence that manipulation is more effective than sham manipulation at short and intermediate follow-ups for people with acute symptoms and an intact anulus, and moderate evidence that no difference exists between traction, laser, and ultrasound at short and intermediate follow-ups.
- There is moderate evidence that the addition of mechanical traction to medication and electrotherapy provides some additional short-term benefits for people with LDHR.
- Adverse events were associated with traction and ibuprofen in the treatment of LDHR.

References

- Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ* 2007;334:1313–7.
- Fardon DF, Milette PC. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined Task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. *Spine* 2001;26:E93–113.
- Tarulli AW, Raynor EM. Lumbosacral radiculopathy. *Neurol Clin* 2007;25:387–405.
- Frymoyer JW. Back pain and sciatica. *N Engl J Med* 1988;318:291–300.
- Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse: updated Cochrane Review. *Spine* 2007;32:1735–47.
- Armon C, Argoff CE, Samuels J, et al. Assessment: use of epidural steroid injections to treat radicular lumbosacral pain: report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology* 2007;68:723–9.
- DePalma MJ, Bhargava A, Slipman CW. A critical appraisal of the evidence for selective nerve root injection in the treatment of lumbosacral radiculopathy. *Arch Phys Med Rehabil* 2005;86:1477–83.
- Abdi S, Datta S, Trescot AM, et al. Epidural steroids in the management of chronic spinal pain: a systematic review. *Pain Physician* 2007;10:185–212.
- Deyo RA. Back surgery—who needs it? *N Engl J Med* 2007;356:2239–43.
- Friedly J, Chan L, Deyo R. Increases in lumbosacral injections in the Medicare population: 1994 to 2001. *Spine* 2007;32:1754–60.
- Hoffman RM, Wheeler KJ, Deyo RA. Surgery for herniated lumbar discs: a literature synthesis. *J Gen Intern Med* 1993;8:487–96.
- Leonardi M, Pfirrmann CW, Boos N. Injection studies in spinal disorders. *Clin Orthop Relat Res* 2006;443:168–82.
- Health Council of the Netherlands. *Management of the Lumbosacral Radicular Syndrome (Sciatica)*. The Hague: Health Council of the Netherlands; 1999.
- Wong DA, Mayer TG, Errico TJ, et al. *Herniated Disc: Phase 3 Clinical Guidelines for Multidisciplinary Spine Care Specialists*. LaGrange, IL: North American Spine Society; 2000.
- Penm E, Dixon T, Bhatia K. *Health Expenditure for Arthritis and Musculoskeletal Conditions in Australia, 2000–01*. Canberra, Australia: Australian Institute of Health and Welfare; 2006.
- Luijsterburg PAJ, Verhagen AP, Braak S, et al. General practitioners' management of lumbosacral radicular syndrome compared with a clinical guideline. *Eur J Gen Pract* 2005;11:113–8.
- Cummins J, Lurie JD, Tosteson TD, et al. Descriptive epidemiology and prior healthcare utilization of patients in the Spine Patient Outcomes Research Trial's (SPORT) three observational cohorts: disc herniation, spinal stenosis, and degenerative spondylolisthesis. *Spine* 2006;31:806–14.
- Rhee JM, Schaufele M, Abdu WA. Radiculopathy and the herniated lumbar disc. Controversies regarding pathophysiology and management. *J Bone Joint Surg Am* 2006;88:2070–80.
- Vroomen PC, de Krom MC, Slofstra PD, et al. Conservative treatment of sciatica: a systematic review. *J Spinal Disord* 2000;13:463–9.
- Luijsterburg PAJ, Verhagen AP, Ostelo RWJG, et al. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: a systematic review. *Eur Spine J* 2007;16:881–99.
- Jordon J, Tamara SM, Weinstein J, et al. Herniated lumbar disc. *Clin Evid (Online)* 2007;1–22.
- Ford JJ, Story I, O'Sullivan P, et al. Classification systems for low back pain: a review of the methodology for development and validation. *Phys Ther Rev* 2007;12:33–42.
- Kent P, Keating J. Do primary-care clinicians think that nonspecific low back pain is one condition? *Spine* 2004;29:1022–31.
- Bejia I, Younes M, Zrouf S, et al. Factors predicting outcomes of mechanical sciatica: a review of 1092 cases. *Joint Bone Spine* 2004;71:567–71.
- Herzog RJ. The radiologic assessment for a lumbar disc herniation. *Spine* 1996;21:195–385.
- Jarvik JG, Deyo RA. Diagnostic evaluation of low back pain with emphasis on imaging. *Ann Intern Med* 2002;137:586–97.
- Lurie JD, Tosteson AN, Tosteson TD, et al. Reliability of magnetic resonance imaging readings for lumbar disc herniation in the Spine Patient Outcomes Research Trial (SPORT). *Spine* 2008;33:991–8.
- Andersson GB, Deyo RA. History and physical examination in patients with herniated lumbar discs. *Spine* 1996;21:105–185.
- van Tulder M, Furlan A, Bombardier C, et al. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine* 2003;28:1290–9.
- Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 [Updated September 2006]*. The Cochrane Library, Issue 4. Chichester, UK: John Wiley & Sons, Ltd. 2006.
- Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354:1896–900.
- Beckmann EC. CT scanning the early days. *Br J Radiol* 2006;79:5–8.
- Fagerlund MK, Thelander UE. Comparison of myelography and computed tomography in establishing lumbar disc herniation. *Acta Radiol* 1989;30:241–6.
- Wong SSL, Wilczynski NL, Haynes RB. Optimal CINAHL search strategies for identifying therapy studies and review articles. *J Nurs Scholarsh* 2006;38:194–9.
- Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc* 2006;94:451–5.
- Zhang L, Ajiferuke I, Sampson M. Optimizing search strategies to identify randomized controlled trials in MEDLINE. *BMC Med Res Methodol* 2006;6:23.
- Snelling NJ. Spinal manipulation in patients with disc herniation: a critical review of risk and benefit. *Int J Osteopath Med* 2006;9:77–84.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- Maher CG, Sherrington C, Herbert RD, et al. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003;83:713–21.
- Verhagen AP, de Vet HC, de Bie RA, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 1998;51:1235–41.
- de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother* 2009;55:129–33.
- Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: medical. *Stat Med* 1989;8:441–54.
- Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001;323:42–6.
- Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
- Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601–5.

46. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;352:609–13.
47. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94.
48. Hedges L, Olkin I. *Statistical Methods for Meta-Analysis*. 1st ed. Orlando, FL: Academic Press; 1985.
49. Deeks J, Higgins J. Statistical algorithms in Review Manager 5 [on behalf of the Statistical Methods Group of the Cochrane Collaboration]. 2007. Available at: www.cochrane.org/resources/handbook/Statistical_Methods_in_RevMan5.pdf. Accessed August 27, 2008.
50. White IR, Thomas J. Standardized mean differences in individually-randomized and cluster-randomized trials, with applications to meta-analysis. *Clin Trials* 2005;2:141–51.
51. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005;5:13.
52. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1988.
53. Cleland JA, Childs JD, Palmer JA, et al. Slump stretching in the management of non-radicular low back pain: a pilot clinical trial. *Man Ther* 2006;11:279–86.
54. Carette S, Leclaire R, Marcoux S, et al. Epidural corticosteroid injections for sciatica due to herniated nucleus pulposus. *N Engl J Med* 1997;336:1634–40.
55. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
56. *Review Manager (RevMan) [Computer Program]*. Version 4.2 for Windows. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2003.
57. Maher CG. A systematic review of workplace interventions to prevent low back pain. *Aust J Physiother* 2000;46:259–69.
58. Roelofs PD, Deyo RA, Koes BW, et al. Nonsteroidal anti-inflammatory drugs for low back pain: an updated Cochrane review. *Spine* 2008;33:1766–74.
59. Staal JB, de Bie RA, de Vet HC, et al. Injection therapy for subacute and chronic low back pain: an updated Cochrane review. *Spine* 2009;34:49–59.
60. Trinh K, Graham N, Gross A, et al. Acupuncture for neck disorders. *Spine* 2007;32:236–43.
61. van Tulder MW, Koes B, Malmivaara A. Outcome of non-invasive treatment modalities on back pain: an evidence-based review. *Eur Spine J* 2006;15(suppl 1):S64–81.
62. Furlan AD, Pennick V, Bombardier C, et al. Updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;34:1929–41.
63. Bleakley CM, McDonough SM, MacAuley DC. Some conservative strategies are effective when added to controlled mobilisation with external support after acute ankle sprain: a systematic review. *Aust J Physiother* 2008;54:7–20.
64. Slade SC, Keating JL. Trunk-strengthening exercises for chronic low back pain: a systematic review. *J Manipulative Physiol Ther* 2006;29:163–73.
65. Bakhtiari AH, Safavi-Farokhi Z, Rezasoltani A. Lumbar stabilizing exercises improve activities of daily living in patients with lumbar disc herniation. *J Back Musculoskeletal Rehabil* 2005;18:55–60.
66. Bonaiuti D, Gatti R, Raschi A, et al. Manual autotractor: preliminary study on the effectiveness of a new device for back pain treatment. *Eura Medicophys* 2004;40:75–81.
67. Buchner M, Zeifang F, Brocai DR, et al. Epidural corticosteroid injection in the conservative management of sciatica. *Clin Orthop Relat Res* 2000;375:149–56.
68. Burton AK, Tillotson KM, Cleary J. Single-blind randomised controlled trial of chemonucleolysis and manipulation in the treatment of symptomatic lumbar disc herniation. *Eur Spine J* 2000;9:202–7.
69. Deli D. Treatment of prolapse of lumbar intervertebral disc by Tuina massage combined with oral administration of Buyang Huanwu Tang—a report of 75 cases. *J Trad Chin Med* 2007;27:43–5.
70. Dincer U, Kiralp MZ, Cakar E, et al. Caudal epidural injection versus nonsteroidal anti-inflammatory drugs in the treatment of low back pain accompanied with radicular pain. *Joint Bone Spine* 2007;74:467–71.
71. Guvenol K, Tuzun C, Peker O, et al. A comparison of inverted spinal traction and conventional traction in the treatment of lumbar disc herniations. *Physiother Theory Pract* 2000;16:151–60.
72. He C, Chen P, Wang X, et al. The clinical effect of herbal magnetic corsets on lumbar disc herniation. *Clin Rehabil* 2006;20:1058–65.
73. Kanayama M, Hashimoto T, Shigenobu K, et al. New treatment of lumbar disc herniation involving 5-hydroxytryptamine_{2A} receptor inhibitor: a randomized controlled trial. *J Neurosurg Spine* 2005;2:441–6.
74. Liu J, Zhang S. Treatment of protrusion of lumbar intervertebral disc by pulling and turning manipulations. *J Trad Chin Med* 2000;20:195–7.
75. Osterman H, Seitsalo S, Karppinen J, et al. Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. *Spine* 2006;31:2409–14.
76. Ozturk B, Gunduz OH, Ozoran K, et al. Effect of continuous lumbar traction on the size of herniated disc material in lumbar disc herniation. *Rheumatol Int* 2006;26:622–6.
77. Peul WC, van den Hout WB, Brand R, et al. Prolonged conservative care versus early surgery in patients with sciatica caused by lumbar disc herniation: two year results of a randomised controlled trial. *BMJ* 2008;336:1355–8.
78. Peul WC, van Houwelingen HC, van den Hout WB, et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 2007;356:2245–56.
79. Santilli V, Beghi E, Finucci S. Chiropractic manipulation in the treatment of acute back pain and sciatica with disc protrusion: a randomized double-blind clinical trial of active and simulated spinal manipulations. *Spine J* 2006;6:131–7.
80. Sherry E, Kitchener P, Smart R. A prospective randomized controlled study of VAX-D and TENS for the treatment of chronic low back pain. *Neurol Res* 2001;23:780–4.
81. Tesio L, Merlo A. Autotractor versus passive traction: an open controlled study in lumbar disc herniation. *Arch Phys Med Rehabil* 1993;74:871–6.
82. Unlu Z, Tasci S, Tarhan S, et al. Comparison of 3 physical therapy modalities for acute pain in lumbar disc herniation measured by clinical evaluation and magnetic resonance imaging. *J Manipulative Physiol Ther* 2008;31:191–8.
83. Veihelmann A, Devens C, Trouillier H, et al. Epidural neuroplasty versus physiotherapy to relieve pain in patients with sciatica: a prospective randomized blinded clinical trial. *J Orthop Sci* 2006;11:365–9.
84. Liddle SD, Gracey JH, Baxter GD. Advice for the management of low back pain: a systematic review of randomised controlled trials. *Man Ther* 2007;12:310–27.
85. Pengel HM, Maher CG, Refshauge KM. Systematic review of conservative interventions for subacute low back pain. *Clin Rehabil* 2002;16:811–20.
86. Macedo LG, Maher CG, Latimer J, et al. Motor control exercise for persistent, nonspecific low back pain: a systematic review. *Phys Ther* 2009;89:9–25.
87. Assendelft WJJ, Morton SC, Yu Emily I, et al. Spinal manipulative therapy for low-back pain. *Cochrane Database Syst Rev* 2004;1:CD000447.
88. Ferreira ML, Ferreira PH, Latimer J, et al. Efficacy of spinal manipulative therapy for low back pain of less than three months' duration. *J Manipulative Physiol Ther* 2003;26:593–601.
89. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995;311:485.
90. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [Updated September 2009]*. The Cochrane Collaboration; 2008. Available at: www.cochrane-handbook.org.
91. Deyo RA, Walsh NE, Schoenfeld LS, et al. Can trials of physical treatments be blinded? The example of transcutaneous electrical nerve stimulation for chronic pain. *Am J Phys Med Rehabil* 1990;69:6–10.
92. Altman DG, Bland JM. How to randomise. *BMJ* 1999;319:703–4.
93. Malmivaara A, Koes BW, Bouter LM, et al. Applicability and clinical relevance of results in randomized controlled trials: the Cochrane review on exercise therapy for low back pain as an example. *Spine* 2006;31:1405–9.
94. Oppenheim JS, Spitzer DE, Segal DH. Nonvascular complications following spinal manipulation. *Spine J* 2005;5:660–6.
95. Oliphant D. Safety of spinal manipulation in the treatment of lumbar disc herniations: a systematic review and risk assessment. *J Manipulative Physiol Ther* 2004;27:197–210.
96. Coxhead CE, Inskip H, Meade TW, et al. Multicentre trial of physiotherapy in the management of sciatic symptoms. *Lancet* 1981;1:1065–8.
97. Moher D, Pham B, Lawson ML, et al. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess* 2003;7:1–90.

■ Appendix: Search Strategy for Ovid Medline

Phase 1: Condition Terms

1. Sciatica/
2. Sciatic\$.ti,ab
3. Radiculopathy/
4. Intervertebral disc displacement/

5. Spinal nerve roots/
6. Pain, referred/
7. ((nerv\$ or root\$ or neuro\$ or neural) ADJ5 (compress\$ or involv\$ or displac\$ or imping\$ or irritat\$ or entrap\$ or compromi\$)).ti,ab
8. (neurologic ADJ5 signs).ti,ab
9. ((refer\$ or radiat\$) ADJ5 (pain or symptoms)).ti,ab
10. (parasthesia or numbness).ti,ab
11. radicul\$.ti,ab
12. ((disc or disc or discs or discs or pulposus) ADJ5 (sequest\$ or protrus\$ or extrus\$ or prolaps\$ or slipped or displac\$ or ruptur\$ or herniat\$ or de-range\$)).ti,ab
13. (lumbar or back or lumbo\$ or L1 or L2 or L3 or L4 or L5 or S1).mp
14. (3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12) AND 13
15. 1 OR 2 OR 14

Phase 2: Trial Type Terms

16. Randomized controlled trial (pub type)
17. Controlled clinical trial (pub type)
18. Clinical trial (pub type)
19. Randomized controlled trials/
20. Exp Clinical trials/
21. Double-blind method/
22. Single-blind method/
23. Random allocation/
24. Placebos/
25. Research design/
26. ((singl\$ or doubl\$ or tripl\$ or trebl\$) adj25 (blind\$ or mask\$)).ti,ab
27. placebo\$.ti,ab
28. random\$.ti,ab
29. (clin\$ adj25 trial\$).ti,ab
30. *versus*.ti,ab
31. (latin adj square).ti,ab
32. Cross-over studies/
33. 16–33/OR

Phase 3: Combine Disorder and Trial Type, Limit to 1971 to 2008, Limit to Human

34. 33 AND 15
35. Limit 34 to yr = “1971–2008”
36. Limit 35 to humans