

## The evidence on surgical interventions for low back disorders, an overview of systematic reviews

Wilco C. H. Jacobs · Sidney M. Rubinstein · Paul C. Willems ·  
Wouter A. Moojen · Ferran Pellisé · Cumhur F. Oner ·  
Wilco C. Peul · Maurits W. van Tulder

Received: 28 August 2012/Revised: 21 March 2013/Accepted: 6 May 2013/Published online: 17 May 2013  
© Springer-Verlag Berlin Heidelberg 2013

### Abstract

**Purpose** Many systematic reviews have been published on surgical interventions for low back disorders. The objective of this overview was to evaluate the available evidence from systematic reviews on the effectiveness of surgical interventions for disc herniation, spondylolisthesis, stenosis, and degenerative disc disease (DDD). An earlier version of this review was published in 2006 and since then, many new, better quality reviews have been published.

**Methods** A comprehensive search was performed in the Cochrane database of systematic reviews (CDSR), database of reviews of effectiveness (DARE) and Pubmed. Two reviewers independently performed the selection of studies, risk of bias assessment, and data extraction. Included are Cochrane reviews and non-Cochrane systematic reviews published in peer-reviewed journals. The following conditions were included: disc herniation, spondylolisthesis, and DDD with or without spinal stenosis. The following comparisons were evaluated: (1) surgery

vs. conservative care, and (2) different surgical techniques compared to one another. The methodological quality of the systematic reviews was evaluated using AMSTAR. We report (pooled) analyses from the individual reviews.

**Results** Thirteen systematic reviews on surgical interventions for low back disorders were included for disc herniation ( $n = 6$ ), spondylolisthesis ( $n = 2$ ), spinal stenosis ( $n = 4$ ), and DDD ( $n = 4$ ). Nine (69 %) were of high quality. Five reviews provided a meta-analysis of which two showed a significant difference. For the treatment of spinal stenosis, intervertebral process devices showed more favorable results compared to conservative treatment on the Zurich Claudication Questionnaire [mean difference (MD) 23.2 95 % CI 18.5–27.8]. For degenerative spondylolisthesis, fusion showed more favorable results compared to decompression for a mixed aggregation of clinical outcome measures (RR 1.40 95 % CI 1.04–1.89) and fusion rate favored instrumented fusion over non-instrumented fusion (RR 1.37 95 % CI 1.07–1.75).

W. C. H. Jacobs (✉) · W. A. Moojen · W. C. Peul  
Department of Neurosurgery, Leiden University Medical Center,  
Albinusdreef 2, PO Box 9600, 2300, RC, Leiden,  
The Netherlands  
e-mail: w.c.h.jacobs@lumc.nl

S. M. Rubinstein · M. W. van Tulder  
Department of Epidemiology and Biostatistics, EMGO Institute  
for Health and Care Research, VU University Medical Center,  
Amsterdam, The Netherlands

P. C. Willems  
Department of Orthopedic Surgery, Maastricht University  
Medical Center, Maastricht, The Netherlands

F. Pellisé  
Department of Orthopedic Surgery, Hospital Vall d'Hebron,  
Barcelona, Spain

C. F. Oner  
Department of Orthopedic Surgery, University Medical Center  
Utrecht, Utrecht, The Netherlands

W. C. Peul  
Department of Neurosurgery, Medical Center Haaglanden,  
The Hague, The Netherlands

M. W. van Tulder  
Department of Health Sciences, Faculty of Earth and Life  
Science and EMGO Institute for Health and Care Research,  
VU University, Amsterdam, The Netherlands

**Conclusions** For most of the comparisons, the included reviews were not significant and/or clinically relevant differences between interventions were identified. Although the quality of the reviews was quite acceptable, the quality of the included studies was poor. Future studies are likely to influence our assessment of these interventions.

**Keywords** Systematic review · Low back pain · Lumbar disc herniation · Spondylolisthesis · Spinal stenosis

## Introduction

Chronic low back pain is defined as a discomfort in the back which may radiate into the legs, hips, or buttocks, and which lasts more than 3 months [1, 35]. Low back pain has become a major health problem in the Western world with a point prevalence of 10.2 % [8], a 1-year prevalence ranging from 22 to 65 % and life-time prevalence of up to 84 % [44]. Approximately, 84 % of those who suffer from low back pain seek medical attention [8]. The economic burden to society is quite large and can be divided into direct costs associated with health care utilization (e.g., hospitalization, medication, diagnostic tests, and therapies) and indirect costs associated with lost productivity (e.g., due to work absenteeism and forced early retirement) [22].

Many episodes of low back pain are temporary and resolve spontaneously in about 25–58 % of the cases, even when specific causes, such as a herniated disc are present [14]. Therefore, conservative treatment is the first step in the management of low back pain. In case of persistent, chronic pain and a clear identification of the pain source, targeted injections or surgical intervention may be indicated.

Specific low back disorders are often discussed in terms of sciatica, spinal stenosis, leg pain, or axial (low) back pain, rather than the underlying disease. For example, spinal stenosis can have several underlying causes, such as spondylolisthesis, hypertrophy of the facet joints, or surrounding soft tissue from a bulging intervertebral disc or hypertrophy of the flaval ligament. Disc height loss caused by degeneration may be a direct cause of axial back pain, but may cause leg pain indirectly by bony compression of the nerve root through narrowing of the neural foramen.

During the last decades, the number of published reviews has dramatically increased. Over 2,000 publications are labeled as a “Review” in Medline each year [4]. A quick screening in Pubmed using ‘low back pain’ limited to ‘meta-analysis’, ‘systematic review’ or ‘review’ showed that in the years 2009–2011, there were on average 167 reviews and meta-analyses per year on low back pain. The quality of these reviews varies and, of course, the results may be biased. In addition, depending upon the methods

and quality of the review, inconsistent findings may be found.

Most systematic reviews assess the relative effectiveness of two interventions. However, usually several treatment options are available in clinical practice and the results of just one review of the two interventions may not help clinicians in their choice from multiple treatment options. The objective of this overview was to evaluate the available evidence from systematic reviews on the effectiveness of surgical interventions for disc herniation, spondylolisthesis, stenosis, and degenerative disc disease. An earlier version of this review was published in 2006 [43] and since then, many new, better quality reviews have been published.

## Methods

All methodological aspects were performed by two of the review authors (WJ, SR, WM, PW, FP) independent of one another. Discrepancies were discussed in a consensus meeting. The protocol of the study was registered in the Prospero database under registration number CRD42011 001819.

### Search

A comprehensive search was conducted in the following databases: the Cochrane database of systematic reviews (CDSR), the database of abstracts of reviews of effectiveness (DARE), as well as Pubmed. The search was conducted in November 2011 with disease specific search string combinations complemented with review specific search string combinations (Table 1) [38, 40]. The search was performed by an experienced librarian.

### Selection

The study had to be a systematic review on surgical interventions for low back disorders. Articles were included if they met the following criteria:

- The disorders examined were disc herniation with radiculopathy, spondylolisthesis, spinal stenosis, or degenerative disc disease. Studies examining fractures, malignancies, scoliosis, failed back surgery syndrome, ankylosing spondylitis and cauda equina syndrome were excluded. If a variety of indications were included, but not separately presented in the results, the article was excluded.
- The interventions studied were surgical interventions, including thermal coagulation therapies, radiofrequency denervation treatments, and surgery such as

**Table 1** Search strategy for finding evidence on surgical interventions for low back pain in Pubmed

Strings	Fields
Low back pain	("low back pain"[tiab] OR "back pain"[mesh] OR "back pain"[tiab] OR "intermittent neurogenic claudication"[tiab] OR "intermittent claudication"[mesh] OR "intermittent claudication"[tiab] OR "neurogenic claudication"[tiab] OR dorsalgia[tiab] OR backache[tiab] OR lumbago[tiab] OR (lumbar[tiab] AND ("pain"[mesh] OR "pain"[all fields])) OR "sciatica"[mesh] OR "sciatica"[all fields] OR sciatica[tiab] OR "Lumbar vertebrae"[mesh] OR "Spine"[mesh] OR "intervertebral disk"[mesh] OR "intervertebral disc"[tiab] OR spine[tiab] OR spinal[tiab] OR vertebra*[tiab] OR disc[tiab] OR discs[tiab] OR disk[tiab] OR disks[tiab])
Indications	("Spondylolisthesis"[mesh] OR spondylolisthesis[tw] OR isthmic[tw] OR lytic[tw] OR low-grade[tw] OR "lumbar stenosis"[tw] OR "spinal stenosis"[mesh] OR "spinal stenosis"[tw] OR stenosis[ti] OR scoliosis[ti] OR "degenerative disc disease"[ti] OR "spinal diseases"[mesh] OR "intervertebral disk displacement"[mesh] OR "intervertebral disk displacement"[tw] OR "discitis"[mesh] OR "discitis"[all fields] OR spondylosis[tiab] OR ((disc[tw] OR discs[tw] OR disk[tw] OR disks[tw]) AND degeneration[tw]) OR herniated[tw] OR hernia[tw])
Combination of clinical queries filter [40] and Shojinia and Bero [38] filter (combined with OR and adapted)	(systematic[sb] OR "systematic review" OR "systematic reviews" OR "systematic literature review" OR "systematic literature reviews" OR meta-analysis[pt] OR meta-analysis[tw] OR metaanalysis[tw] OR meta-analyses[tw] OR "evidence-based medicine" OR ("evidence-based" AND (guideline[tw] OR guidelines[tw] OR recommendation*)) OR ("evidenced-based" AND (guideline[tw] OR guidelines[tw] OR recommendation*)) OR consensus development conference[pt] OR health planning guideline OR "health planning guidelines" OR guideline[pt] OR cochrane database syst rev OR acp journal club OR health technol assess OR evid rep technol assess summ OR evid based dent OR evid based nurs OR evid based ment health OR clin evid OR ((systematic[tw] OR systematically OR critical[tw] OR (study[tiab] AND selection[tiab]) OR (predetermined OR inclusion AND criteri*[tw]) OR exclusion criteri* OR "main outcome measures" OR "standard of care")) AND (survey[tw] OR surveys[tw] OR overview* OR review[tw] OR reviews[tw] OR search* OR handsearch OR analysis[tw] OR critique[tw] OR appraisal OR (reduction AND risk AND (death OR recurrence))) AND (literature[tw] OR articles OR publications[tw] OR publication[tw] OR bibliography[tw] OR bibliographies OR published OR unpublished OR citation OR citations OR database OR internet[tw] OR textbooks[tw] OR references OR trials OR meta-analysis[mh] OR (clinical[tw] AND studies) OR treatment outcome)) OR ((review[pt] OR guideline[pt] OR consensus[ti] OR guideline*[ti] OR literature[ti] OR overview[ti] OR review[ti]) AND ((Cochrane[tw] OR Medline[tw] OR CINAHL[tw] OR CINAHL[tw] AND library[tw])) OR (handsearch*[tw] OR search*[tw] OR searching[tw]) AND (hand[tw] OR manual[tw] OR electronic[tw] OR bibliography*[tw] OR database* OR (Cochrane[tw] OR Medline[tw] OR CINAHL[tw] OR (national[tw] AND library[tw])))) OR ((synthesis[ti] OR overview[ti] OR review[ti] OR survey[ti]) AND (systematic[ti] OR critical[ti] OR methodologic[ti] OR quantitative[ti] OR qualitative[ti] OR literature[ti] OR evidence[ti] OR evidence-based[ti])) NOT (case report[ti] OR editorial[ti] OR editorial[pt] OR letter[pt] OR newspaper article[pt])

[ti] search field used in title

- decompression surgery or fusion surgery, compared to one another or compared to conservative care.
- Only systematic reviews with a comprehensive search (i.e., more than one database and extensive use of search string combinations) and risk of bias assessment (at least assessing selection bias) were included.
- Studies published prior to 2001 were excluded because interventions were thought not to be consistent with contemporary procedures.

#### Risk of bias assessment

Quality was assessed by the AMSTAR tool for systematic reviews [37]. Operationalization of items is presented in Table 2. Studies were considered to be of high quality if items 3, 6, 7, and 8 of the AMSTAR tool were met and the review had no critical errors (for example, deficiencies in

the meta-analysis). When one reviewer had been involved in an included study, the risk of bias of this study was assessed by one of the other reviewers.

#### Analysis

The effect of surgery for (1) lumbar radiculopathy secondary to disc herniation, (2) spondylolisthesis, (3) spinal stenosis, and (4) degenerative disc disease was compared to conservative care and different surgical techniques were compared to one another. Subjective outcome data, such as pain, functional status, recovery, as well as physiological/objective data (e.g., success of fusion) are reported when available. Data from individual studies were not pooled, instead we report on the meta-analyses from the individual reviews. We compared the effects of the interventions to the thresholds for minimal relevant differences as reported

**Table 2** AMSTAR [37] is a measurement tool created to assess the methodological quality of systematic reviews

1	Was an ‘a priori’ design provided? The research question and inclusion criteria should be established before the conduct of the review. Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a “yes.”
2	Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place. Two people do study selection, two people do data extraction, consensus process or one person checks the other’s work.
3	Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Keywords and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found. If at least two sources + one supplementary strategy used, select “yes” (Cochrane register/Central counts as two sources; a grey literature search counts as supplementary).
4	Was the status of publication (i.e., grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc. If review indicates that there was a search for “grey literature” or “unpublished literature,” indicate “yes.” SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.
5	Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided. Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select “no.”
6	Were the characteristics of the included studies provided? In an aggregated form, such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported. Acceptable if not in table format as long as they are described as above.
7	Was the scientific quality of the included studies assessed and documented? ‘A priori’ methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo-controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant. Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study (“low” or “high” is fine, as long as it is clear which studies scored “low” and which scored “high”; a summary score/range for all studies is not acceptable).

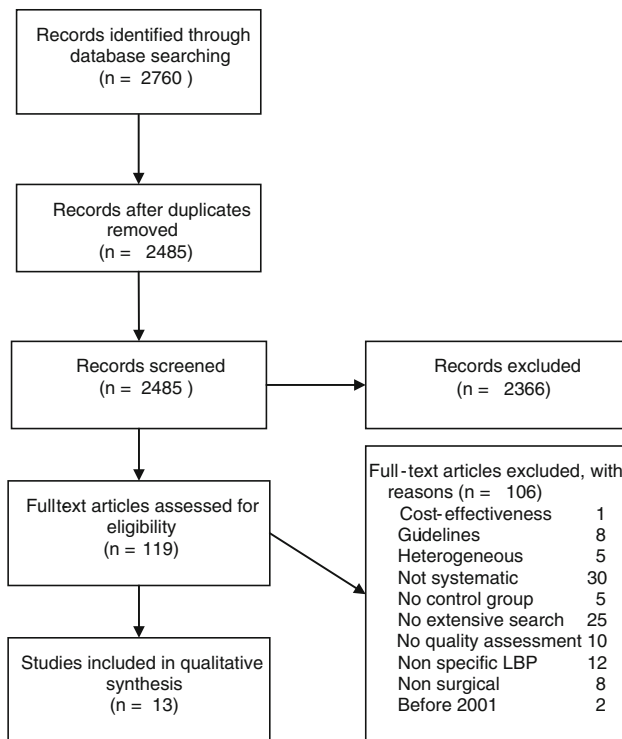
**Table 2** continued

8	Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations. Might say something such as “the results should be interpreted with caution due to poor quality of included studies.” Cannot score “yes” for this question if scored “no” for question 7.
9	Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, $I^2$ ). If heterogeneity exists, a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?). Indicate “yes” if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.
10	Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges–Olkin). If no test values or funnel plot included, score “no”. Score “yes” if mentions that publication bias could not be assessed, because there were fewer than ten included studies.
11	Was the conflict of interest included? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies. To get a “yes,” must indicate source of funding or support for the systematic review AND for each of the included studies.

by Ostelo et al. [30] being 15 for the Visual Analogue Scale, 5 for the Roland Disability Questionnaire, and 10 for the Oswestry Disability Index.

## Results

Electronic searches identified 2,485 references, excluding duplicates (Fig. 1). Thirteen systematic reviews on clinical effectiveness of several interventions were finally included. Nine reviews were classified as having a high quality based on the risk of bias assessment (Table 3). Seven of the reviews [7, 12, 13, 17, 20, 41, 48] included randomized studies only, while six reviews [15, 18, 23, 24, 28, 29] also included observational studies, including uncontrolled studies. Indications were disc herniation (six reviews [7, 13, 15, 17, 23, 29]), spondylolisthesis (two reviews [18, 24]), stenosis (four reviews [7, 12, 20, 28]), and degenerative disc disease (four reviews [7, 12, 41, 48]). Two reviews [7, 12] included a mix of indications (herniated disc, degenerative disc disease, and/or stenosis), but these allowed extraction of data and conclusions about the separate indications.



**Fig. 1** Flow chart

### Disc herniation with radiculopathy

There were six reviews (Table 4) dealing with low back pain due to disc herniation [7, 13, 15, 17, 23, 29].

#### *Conservative vs. surgical interventions*

Four high quality reviews [7, 13, 17, 23] studied conservative vs. surgical treatment for sciatica due to a herniated disc. In 2007, Gibson and Waddell [13] included three randomized studies plus one conference proceeding (815 patients). Chou et al. [7] included four randomized studies (767 patients) in 2009, and concluded that discectomy is moderately superior (3 points on a Roland Morris Disability questionnaire, where 5 is clinically relevant) to standard conservative treatment in terms of pain and function in the first 2–3 months postoperatively. In 2011, the same five randomized studies (1,235 patients), including those from the previous reviews were included by Jacobs et al. [17] and Lewis et al. [23]. The review of Jacobs et al. [17], including one low risk of bias study [31], demonstrated that surgery leads to faster pain relief after 3 months than conservative treatment with no differences at 1 or 2 years follow-up. However, another low risk of bias study [47] found no differences. No meta-analysis could be performed due to heterogeneity of the studies. The main conclusion from these four reviews was consistent that surgery appears to lead to predominantly short term

benefits, but that the scarcity of high quality studies does not support a definite choice for conservative or surgical treatment for disc herniation with sciatica. Lewis et al. [23] included five RCTs and nine observational studies on disc surgery vs. conservative interventions. Global effect and condition-specific outcome measures favored surgery, but not significantly at medium term (OR 1.66 95 % CI 0.98–2.18) and mean difference (MD  $-0.14$  95 % CI  $-0.50$  to  $0.23$ ), respectively. No overall conclusions could be drawn.

#### *Different surgical interventions vs. one another*

Four reviews (three high quality) [13, 15, 23, 29] evaluated the effect of different surgical techniques with one another. Two of these included all surgical techniques [13, 23], while two others focused on automated percutaneous discectomy [15] and transforaminal endoscopic surgery [29]. Gibson and Waddell [13] included 23 RCTs on five different comparisons in 2007. Results of meta-analyses were not reproduced. They concluded that the microscopic discectomy gives broadly the same results as open discectomy and that the evidence for minimally invasive discectomy remained unclear. Major design weaknesses of the included studies may have resulted in a considerable potential for bias. For example, there were many problems with sequence generation and allocation concealment, lack of outcome assessor blinding, and lack of proper outcome parameters.

Two reviews [15, 29] examined minimal invasive discectomies. Hirsch et al. [15] included one trial [34] (141 patients) and compared automated percutaneous lumbar discectomy (APLD) with chemonucleolysis. They concluded that with adequate patient selection (disc herniation) APLD is safe and provides appropriate relief. Several limitations were observed in the studies, such as including patients with an inappropriately (too large) sized disc herniation for APLD ( $>30\%$ ). Nellensteijn et al. [29] included one RCT (60 patients) and seven non-randomized controlled studies (1,822 patients). They compared transforaminal endoscopic surgery with open approaches to the disc herniation, such as open laminotomy and open microdiscectomy. No meta-analysis was performed. They concluded that there is still insufficient information to support this technique over open procedures. The quality of evidence was poor because of heterogeneity and study quality.

#### Low-grade isthmic spondylolisthesis (Type II)

One review [18] compared surgery with conservative treatment and surgical techniques to one another for low-grade adult isthmic spondylolisthesis (Table 5). Eight

**Table 3** Risk of bias of included systematic reviews

Main indication	Study Author	AMSTAR item											
		1	2	3	4	5	6	7	8	9	10	11	High quality
Herniated disc	Gibson and Waddell [13] (HNP)	Y	N	Y	Y	Y	Y	Y	Y	Y	N	N	Y
	Hirsch et al. [15]	N	N	Y	N	Y	N	Y	N	n/a	N	N	
	Jacobs et al. [17]	N	Y	Y	N	N	Y	Y	Y	?	Y	N	Y
	Lewis et al. [23]	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y
	Nellensteijn et al. [29]	N	Y	Y	?	N	Y	Y	Y	Y	N	N	Y
Spondylolisthesis	Jacobs et al. [18]	N	Y	Y	N	Y	Y	Y	N	n/a	N	N	
	Martin et al. [24]	N	N	Y	Y	N	Y	Y	Y	Y	N	N	Y
Stenosis	Kovacs et al. [20]	N	Y	Y	N	Y	Y	Y	Y	Y	N	Y	<sup>a</sup>
	Moojen et al. [28]	N	Y	Y	?	Y	Y	Y	Y	Y	N	Y	Y
DDD	van den Eerenbeemt et al. [41]	N	Y	Y	?	Y	Y	Y	Y	n/a	N	N	Y
	Yajun et al. [48]	N	Y	Y	N	N	Y	Y	Y	N	Y	N	
Combinations	Chou et al. [7] (HD, DDD, stenosis)	N	N	Y	N	Y	Y	Y	Y	n/a	N	Y	Y
	Gibson and Waddell [12] (Spl) (DDD, stenosis)	Y	N	Y	Y	Y	Y	Y	Y	Y	N	N	Y

Y yes, meets requirement; N no, does not meet requirement; n/a not applicable; DDD degenerative disc disease; HD herniated disc; SPL spondylolisthesis

<sup>a</sup> Quality was decreased because of inconsistencies in the evaluation

randomized studies (376 patients), four prospective studies (148 patients), and 17 retrospective case series (648 patients) were included. Seven randomized studies compared posterolateral fusion techniques to one another. Techniques that were compared include addition of instrumentation, such as screws and rods or plates, addition of decompression, and addition of anterior interbody fusion. No meta-analysis was performed and best evidence synthesis could not reveal a difference in fusion rate at 2 years between instrumented and non-instrumented fusion. One high-risk of bias study [27] showed superior results for clinical outcome (74 vs. 43 % good outcome) at 2 years of posterolateral fusion vs. exercise. Blinding and intention to treat were rarely used in the studies. The methodological quality of the included RCTs and observational studies was variable and did not allow pooling because of heterogeneity of the included populations.

**Degenerative spondylolisthesis (Type III)**

One high quality review [24] compared different surgical techniques to one another for adult degenerative spondylolisthesis (Table 5). The study, which included four RCTs (180 patients) and nine observational studies (405 patients), compared fusion vs. decompression alone, and instrumented fusion vs. non-instrumented fusion. Spinal fusion was found to lead to a better chance of improved clinical outcome than decompression with a pooled RR of 1.40 (95 % CI 1.04–1.89, see Table 5). Instrumentation led to improved fusion (pooled RR 1.37, 95 % CI 1.07–1.75), but

this was not related to better clinical outcome (pooled RR 1.19, ns). The quality of the included studies was limited because of unclear randomization techniques and limited use of patient centered outcomes in the randomized studies and because of unclear selection strategies for the interventions in the observational studies. A limitation of the review was the lack of duplicate selection for the whole selection process.

**Low back pain without stenosis in the presence of degenerative disc**

There were four reviews (Table 6) dealing with low back pain without stenosis in the presence of degenerative changes in the disc [7, 12, 41, 48].

*Conservative vs. surgical interventions*

Two high quality reviews [7, 12] compared surgery with conservative treatment for discogenic low back pain without stenosis. Gibson and Waddell [12] included three randomized studies (141 patients) on IDET vs. placebo. They could not draw any firm conclusions from these small studies. Chou et al. [7] included four higher quality studies (767 patients) on fusion vs. conservative interventions. They evaluated VAS pain, Oswestry Disability index, and SF-36 at one and 2 years, but no meta-analysis was performed. The results were inconsistent and were ascribed to the differences in rehabilitation intensity in the non-surgical intervention group, as fusion was no more effective

**Table 4** Characteristics of included reviews on disc herniation with results and conclusions

Author	Number of studies	Type of studies	Included diagnosis/es in controlled studies	Included comparisons in controlled studies	Results of meta-analyses of controlled studies	Conclusions
<b>Conservative vs. surgical interventions</b>						
Gibson and Waddell [13]	4	RCT	Symptomatic HNP	Surgical vs. conservative	No meta-analysis	Discectomy for selected patients with sciatica due to HNP provides faster relief than conservative management
Chou et al. [7]	4	RCT	HNP with radiculopathy	Discectomy vs. nonsurgical therapy	No meta-analysis	Surgery gives only short-term benefits for radiculopathy
Jacobs et al. [17]	5	RCT	HNP with sciatica	Surgery vs. conservative	No meta-analysis	Surgery leads to faster pain relief, with no differences at 1 year. difference in VAS pain at 8 weeks 17.7 mm
Lewis et al. [23]	5 9	RCT Other	Sciatica	Disc surgery vs. conservative	Global effect and CSOMs favor surgery (not significantly at medium term (OR 1.66 95 % CI 0.98–2.18) and (MD –0.14 95 % CI –0.50 to 0.23) resp	No overall conclusions
<b>Different surgical interventions</b>						
Gibson and Waddell [13]	23	RCT	Symptomatic HNP	Different forms of discectomy (9); automated percutaneous discectomy vs. chemonucleolysis (2) or microdiscectomy (2); laser discectomy (2); barrier membranes (8)	No results of meta-analysis used in manuscript	The evidence for minimally invasive techniques remains unclear
Hirsch et al. [15]	1 10	RCT PCS	HNP	Automated percutaneous lumbar discectomy vs. chemonucleolysis	No meta-analysis	APLD may provide appropriate relief in properly selected patients with contained disc herniation. APLD is a safe procedure with minimal complications
Nellensteijn et al. [29]	1 5 2 12 19	RCT QRCT RetroCT PCS RetroCS	Symptomatic disc herniation	Transforaminal endoscopic surgery	No meta-analysis	Current evidence is poor and does not provide valid information to support or refute transforaminal endoscopic surgery
Lewis et al. [23]	27 21	RCT Others	Sciatica	Disc surgery vs. chemonucleolysis (26), steroid injections (1), mixed treatments (4), surgical techniques (17)	Heterogeneous interventions, meta-analyses not reproduced	No overall conclusions

(Q/RCT (Quasi) randomized controlled trial, HNP herniated nucleus pulposus, MD mean difference, RR relative risk, CI confidence interval, VAS visual analogue scale, CSOM condition-specific outcome measures

**Table 5** Characteristics of included reviews on spondylolisthesis with results and conclusions

Author	Number of studies	Type of studies	Included diagnosis/-es in controlled studies	Included comparisons in controlled studies	Results of meta-analyses of controlled studies	Conclusions
Jacobs et al. [18]	8 2 4 16	RCT RetroCT PCS RetroCS	Low-grade, adult, isthmic spondylolisthesis (Type II)	Fusion techniques vs. nothing, conservative, or decompression	No meta-analysis	Best surgical technique could not be identified
Martin et al. [24]	4 1 8	RCT PCT RetroCT or unclear	Degenerative lumbar spondylolisthesis (Type III)	Fusion vs. decompression	Clinical outcome for fusion vs. decompression: RR 1.40 (95 % CI 1.04–1.89). Fusion for instrumentation vs. no instrumentation: RR 1.37 (95 % CI 1.07–1.75). Clinical outcome for instrumentation vs. no instrumentation: RR 1.19 (95 % CI 0.92–1.54)	Fusion better clinical outcome than decompression; no clinical effect of instrumentation, but better fusion

(*Q*)RCT (Quasi) randomized controlled trial, *PCT* prospective controlled trial, *RetroCT* retrospective controlled trial, *PCS* prospective case series/cohorts, *RetroCS* retrospective case series/cohorts, *RR* relative risk, *CI* confidence interval

than intensive rehabilitation, but fusion was associated with small-to-moderate benefits compared to standard non-surgical therapy.

#### *Different surgical interventions vs. one another*

There is no evidence from systematic reviews on the clinical effectiveness of different surgical techniques to achieve a fusion for discogenic low back pain without stenosis.

Four reviews [7, 12, 41, 48] (three high quality) compared fusion with disc replacement for discogenic low back pain without stenosis. In 2005, Gibson and Waddell [12] could not draw any firm conclusions due to the availability of preliminary data only. The review of Chou et al. [7] from 2009 included two randomized studies (596 patients) that showed that disc replacement and fusion result in comparable success and complication rates. In 2010, van den Eerenbeemt et al. [41] included three randomized studies (616 patients). No meta-analysis could be performed on Oswestry and VAS pain and the conclusion was that there was insufficient evidence to justify the widespread use of disc replacement. Yajun et al. [48] presented a review including four randomized studies (759 patients). They found a mean difference on VAS pain (0–100) of 4.75 (95 % CI 0.35–9.14) favoring disc replacement, which was not regarded clinically relevant according to Ostelo et al. [30] (15 mm). There is possible risk of bias in the included studies due to sponsoring.

#### Stenosis in the presence of degenerative changes

There were four reviews (Table 7) dealing with low back pain with stenosis in the presence of degenerative changes [7, 12, 20, 28].

#### *Conservative vs. surgical interventions*

Two reviews [7, 20] (one high quality, see Table 7) examined surgery vs. conservative treatment for spinal stenosis with symptoms of neurogenic claudication or sciatica. Both reviews included the same five randomized studies (918 patients). The reviews compared effectiveness of interspinous devices or decompressive surgery with conservative treatment. Kovacs et al. [20] found no difference in a pooled Oswestry Disability score (MD  $-1.57$  95 % CI  $-4.65$  to  $1.51$ ; ITT analysis, which is not clinically relevant according to Ostelo et al. [30]). Chou et al. [7] did not perform a meta-analysis, but concluded that there was good evidence that decompressive laminectomy (with or without fusion) is superior to non-surgical therapy for the first 2 years after surgery, but benefits appear to diminish afterwards. The studies included considerable variation in numbers of patients with spondylolisthesis (0–100 % or unknown) making the patient sample heterogeneous. The methodology of Kovacs et al. [20] appears to be of high quality, although some of the scores in the quality assessment do not match those of the quotes on methodology from included studies that are documented in the tables in the appendix. For example, they rated compliance as acceptable in five studies, although one study quoted “compliance was low”, another study had 50 % cross-over and a third study did not have any information on compliance. Two publications [2, 49] included by Chou et al. [7] were regarded to be one study as one publication is suspected to be a sub-sample from another included study.

Three high quality reviews [7, 12, 28] (Table 7) compared interspinous process distraction (IPD) devices with conservative treatment. Gibson and Waddell [12] included



**Table 6** Characteristics of included reviews on LBP without stenosis in the presence of degenerative changes in the disc with results and conclusions

Author	Number of studies	Type of studies	Included diagnosis/-es in controlled studies	Included comparisons in controlled studies	Results of meta-analyses of controlled studies	Conclusions
<b>Conservative vs. surgical interventions</b>						
Gibson and Waddell [12]	5	RCT	Back pain without neurological compromise, CLBP	Idet vs. placebo (3); fusion vs. conservative (2)	No meta-analysis	There are still open questions on clinical effectiveness of fusion
Chou et al. [7]	4	RCT	LBP with degenerative changes	Fusion vs. non-surgical therapy (4)	No meta-analysis	For non-radicular LBP with degenerative changes fusion not different from intensive rehabilitation, but more beneficial than standard conservative treatment
<b>Different surgical interventions</b>						
Gibson and Waddell [12]	3	RCT	Back pain without neurological compromise, CLBP	Total disc replacement vs. fusion (3) (comparisons on techniques excluded, unclear indications)	No meta-analysis	Only preliminary results on disc replacement, no conclusions permitted
Chou et al. [7]	2	RCT	LBP with degenerative changes	Total disc replacement vs. fusion (2)	No meta-analysis	No difference for total disc replacement discs and fusion with regard to effectiveness and complications, but long-term outcome is needed
Yajun et al. [48]	5	RCT	Degenerative disc disease	Total disc replacement vs. fusion (5)	Not reliable due to duplicate inclusion	From the existing outcomes, total artificial disc replacement does not show significant superiority for treatment of lumbar DDD compared with fusion More high quality RCTs with long-term outcome are needed
Van den Eerenbeemt et al. [41]	3	RCT	Symptomatic degenerative disc disease	Total disc replacement vs. fusion (3)	No meta-analysis	The existing evidence, specifically regarding long-term effectiveness and/or safety is considered insufficient to justify the widespread of total disc replacement over fusion

RCT randomized controlled trial, PCT prospective controlled trial, PCS prospective case series/cohorts, *RetrosCS* retrospective case series/cohorts, HNP herniated nucleus pulposus, LBP low back pain, ZCQ Zurich claudication questionnaire, ODI Oswestry disability index, ITT intention to treat

one RCT, Chou et al. [7] included two RCTs (142 patients), and Moojen et al. [28] additionally included one controlled observational study (60 patients), and seven non-controlled observational studies (391 patients). All three reviews concluded that the interspinous spacers, both statistically and clinically, significantly improved the Zurich Claudication Questionnaire total score more than conservative treatment did, although the quality of the evidence was low. The weighted mean difference calculated by Moojen et al. [28] on the Zurich Claudication Questionnaire (average of improvements of symptom severity and physical function scales compared to baseline) was 23.2 % (95 % CI 18.5–27.8) compared to the pooled baseline score of 5.2 points. This is larger than the minimally detectable change score of 15 % reported by Pratt et al. [33], but minimal clinically important differences for the averaged

change score are not known. From the observational studies a complication rate of 7 % could be calculated. Moojen et al. [28] also concluded that the cost for these techniques is high and more research (on cost-effectiveness) is necessary before worldwide implementation is justified. It should be noted that these reviews include duplicate publications of primary studies.

#### *Different surgical interventions vs. one another*

One high quality review [12] (Table 7) dealt with different surgical techniques and included four RCTs. One RCT [32] could not find any difference between laminectomy and multiple laminotomy, but the trial was small and there were considerable cross-over and co-interventions. Three randomized studies (139 patients) were available for

**Table 7** Characteristics of included reviews on LBP with stenosis in the presence of degenerative changes with results and conclusions

Author	Number of studies	Type of studies	Included diagnosis/es in controlled studies	Included comparisons in controlled studies	Results of meta-analyses of controlled studies	Conclusions
<b>Conservative vs. surgical interventions</b>						
Kovacs et al. [20]	5	RCT	Lumbar stenosis with sciatica or neurogenic claudication	Surgery (decompression or IPD) vs. conservative (5)	ODI for surgery vs. conservative: MD –1.57 (–4.65 to 1.51) (ITT analysis)	Surgery (IPD or instrumentation) is more effective than conservative treatment
Chou et al. [7]	6	RCT	(Discogenic?) with stenosis	Decompressive surgery vs. conservative (4); IPD vs. non-surgical treatment (2)	No meta-analysis	IPD superior to non-surgical therapy. Long-term outcome is needed.
Moojen et al. [28]	2 1 7	RCT PCT PCS	Intermittent neurogenic claudication with spinal stenosis	IPD vs. conservative treatment (2)	ZCQ: MD 23.2 (95 % CI 18.5–27.8; suspected duplicate inclusion)	Quality of evidence is low. More cost-effectiveness studies should be performed
Gibson and Waddell [12]	2	RCT	Spinal stenosis and/or nerve root compression	Spacer vs. non-operative (1); fusion vs. conservative (1)	No meta-analysis	No firm conclusions
<b>Different surgical interventions</b>						
Gibson and Waddell [12]	4	RCT	Spinal stenosis and/or nerve root compression	Decompression vs. decompression and fusion (3); different techniques for decompression (1)	Surgeon's rating (OR 0.44 95 % CI 0.13–1.148).	No firm conclusions

*RCT* randomized controlled trial, *PCT* prospective controlled trial, *PCS* prospective case series/cohorts, *RetroCS* retrospective case series/cohorts, *HNP* herniated nucleus pulposus, *LBP* low back pain, *ZCQ* Zurich claudication questionnaire, *IPD* interspinous process device, *MD* mean difference, *ODI* Oswestry disability index, *ITT* intention to treat

comparing additional fusion to decompression alone. No difference was found in surgeon's rating of outcomes (OR 0.44 95 % CI 0.13–1.15). Other outcomes (re-operation, spondylolisthesis progression, and improvement in walking distance) were not significant. Good result at 18–24 months was significant, but very imprecise (OR 4.41 95 % CI 1.09–17.8). The authors conclude that there is no clear evidence about the most effective decompressive technique or the extent of the decompression.

## Discussion

We included thirteen, principally good quality systematic reviews on the effectiveness of various surgical interventions for disc herniation with radiculopathy, spondylolisthesis, degenerative disc disease, and spinal stenosis. For most of the compared interventions in the included reviews, we were not able to identify significant and/or clinically relevant differences in effectiveness based on pooled estimates. The exceptions were (1) better fusion rate for additional instrumentation for degenerative lumbar spondylolisthesis, (2) better clinical outcome for fusion compared to decompression for degenerative lumbar

spondylolisthesis, (3) better Zurich claudication questionnaire scores for interspinous process device compared to conservative treatment, but long-term outcome and cost-effectiveness should be assessed, (4) based on individual studies, surgery for disc herniation with radiculopathy results in more short-term pain relief as compared to conservative treatment with no differences between the treatments at 1-year follow-up. However, the individual studies included in the reviews often showed a high-risk of bias and methodological flaws, such as poor reporting, heterogeneity, no validated outcomes, and only short- or mid-term follow-ups reported.

Spinal pathology that causes concordant symptoms should be the starting point for specific interventions. When using symptoms as a starting point, it is required that one knows the underlying pathology before one can decide on the optimal treatment. For example, taking sciatica or neurogenic claudication as a starting point for presenting evidence on treatment efficacy is only useful if the results are presented separately for causes, such as disc herniation, spondylolisthesis or nerve root compression in a narrowed foramen, as these pathologies may each require different interventions. Another example was found in the review of Andersson et al. [3], which compared fusion against

intradiscal electrothermal therapy (IDET) for intractable discogenic low back pain. They found no direct comparisons, but included uncontrolled studies reporting on one of both interventions. This makes the interpretation of such a review problematic, as the patient populations and specific indications in the separate series may be different, or are simply not reported in sufficient detail.

Cochrane reviews are known for their validated methodology; however, they mainly include RCTs, which might be considered a limitation. Especially for research on surgical interventions there are limitations associated with the randomized trial with regard to patient inclusions, performance bias, assessment of (long-term) complications, and limited external validity [5, 11, 16]. While there is debate about these possible limitations of randomization, observational studies should not be ignored as they might better reflect clinical practice. However, observational studies are likely to have a risk of selection bias, information bias, and confounding. Obviously, observational studies with a high-risk of selection bias do not reflect clinical practice either. We excluded reviews that did not use a comprehensive search strategy, because it was thought that a limited search strategy would give a biased view of the evidence. For example, two reviews were excluded because only three keywords were used which resulted in only 57 records [9] and 35 records [26].

Evaluation of the risk of bias in systematic reviews is essential. We excluded reviews if no formal risk of bias assessment was performed [9, 19, 25], or if only the study design was categorized as randomized, observational, or retrospective [39, 46]. Mirza and Deyo [26] evaluated the quality of studies with the Consort statement, which is rather a guideline for reporting a trial than a tool for evaluating their risk of bias. Overall, there is an inconsistent use of risk of bias tools in the included reviews. Future reviews need to adopt a “validated” or generally accepted tool. For spinal surgery, we would recommend the guidelines by the Cochrane Back Review Group [10].

The results that are found in any trial, but also in systematic reviews need to be evaluated against minimal clinically important changes as assessed by outcome measures [30]. A systematic review of many studies will lead to a precise and statistically significant effect estimate, but this effect may not be clinically relevant. On the other hand, systematic reviews with only a few studies, as is often the case in surgery, tend to produce no statistically significant effects. Also, in our overview, only one of the comparisons (surgical techniques for HNP) had more than ten RCTs, most of the RCTs were small, and only few significant differences were found. The conclusion for ‘no difference’ could lack precision or power, while the difference might be clinically relevant. This should not be

misinterpreted for equivalence of interventions, but regarded as “insufficient evidence for a difference”.

Two excluded review studies [25, 45] did not use an accepted pooling method. For example, McGirt et al. [25] included the comparative studies and the case series and analyzed both study designs in the same analysis. Such a pooling method should not be used for decision analysis because the different selection mechanisms in case series could result in bias in the estimates of effect.

#### Strengths and limitations

There are some limitations to our overview. (1) There is a possible conflict of interest with the included studies as studies from one or more of the members of our own group could have had a greater risk of being included into the review. Further, although assessment of risk of bias of a review was not performed by any of its own authors, evaluating a study of one of the other group members might have introduced a bias towards a lower risk of bias for our own studies. (2) We suspect that three reviews [7, 28, 48] included duplicate studies into their analyses. The inclusion of duplicate studies makes the pooled analyses unreliable, with a bias towards the results of the duplicate included studies. (3) We accepted degenerative disc disease (DDD) as a disease entity and included studies that reported on DDD. We suspect that some studies using the term ‘non-specific LBP’ deal with the same type of patients. (4) The quality of an overview relies on the quality of the included studies. Biases at the level of the primary studies as well as biases at the level of the included systematic reviews accumulate in an overview. For example, commercial interests related to the surgical techniques could result in sponsored studies with negative results failing to be submitted or published [6, 36], which could introduce publication bias, also in the overview. (5) The results of the present overview limit the possibility of drawing conclusions about indirect comparisons. Preferably, these indirect comparisons should be studied in an original study including multiple treatment arms, or alternatively, be addressed in a network meta-analysis. (6) When there was one study that had a strong influence on the results or conclusion of the included systematic review, we referred to their results and quality. We have to acknowledge that in those cases, we were dependent on the quality assessment of the authors of the original systematic reviews.

A strength of our overview is that we have adopted a systematic approach which included several steps to limit the potential for bias, including a systematic search, duplicate selection and quality appraisal.

## Recommendations for future research

In general, there is an abundance of reviews on innovative treatments for low back pain (IDET, disc replacement, interspinous devices) with relatively few studies. Evidence-based medicine requires new treatments to be compared to the current gold standard techniques, which would presumably be fusion or decompression surgery. However, from the current systematic literature reviews neither fusion nor decompression surgery could be regarded as the gold standard. More energy should be directed at establishing evidence or consensus for a current gold standard and its relative effectiveness compared to other treatment options. This would facilitate an evidence-based comparison for innovative treatments. We realize that there may be less interest in repeating previous studies than in studying innovative interventions, that funding for such studies may be difficult to obtain, and that there is a barrier to their publication due to the perceived lack of “novelty”. However, we hope that with the current trends in healthcare finance there will be a quicker elimination of (cost-)ineffective innovative interventions, improving the chances for funding and publication of research into established interventions.

Systematic literature reviews need to be updated regularly. Ideally, an accumulative review should be maintained, and be readily accessible after the addition of new evidence. Ideally, a review should be updated and published every 2–5 years, depending on the accumulation of evidence. This is also the goal of the Cochrane Collaboration.

We urge those preparing a report of a systematic review to address the items that are included in the PRISMA statement. The improved quality of the report will facilitate an assessment of risk of bias in the review process and, thus, a better understanding of the results and conclusions by the reader.

Finally, the results from an overview on surgical interventions for low back disorders need to be used together with an overview of conservative interventions, such as those from Kuijpers et al. and van Middelkoop et al. [21, 42]. Only then physicians will have a complete overview of the evidence for all possible treatment options for low back disorders. Such an overview could be implemented in clinical guidelines which could be used for counseling of the individual patient, but also for communication with insurers, policy makers, and other health care providers who are involved with the management of low back disorders.

**Acknowledgments** We would like to thank Jan Schoones, Waleus library, University Medical Center Leiden for his aid with the search.

**Conflict of interest** None.

## References

- Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15(Suppl 2):S192–S300
- Anderson PA, Tribus CB, Kitchel SH (2006) Treatment of neurogenic claudication by interspinous decompression: application of the X STOP device in patients with lumbar degenerative spondylolisthesis. *J Neurosurg Spine* 4(6):463–471
- Andersson GBJ, Mekhail NA, Block JE (2006) Treatment of intractable discogenic low back pain. A systematic review of spinal fusion and intradiscal electrothermal therapy (IDET). *Pain Physician* 9(3):237–248
- Bastian H, Glasziou P, Chalmers I (2010) Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 7(9):e1000326
- Bhandari M, Tornetta P, Ellis T, Audige L, Sprague S, Kuo JC, Swiontkowski MF (2004) Hierarchy of evidence: differences in results between non-randomized studies and randomized trials in patients with femoral neck fractures. *Arch Orthop Trauma Surg* 124(1):10–16
- Carragee EJ, Hurwitz EL, Weiner BK (2011) A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *Spine J* 11(6):471–491
- Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD (2009) Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine (Phila Pa 1976)* 34(10):1094–1109
- Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, Castel LD, Kalsbeek WD, Carey TS (2009) The rising prevalence of chronic low back pain. *Arch Intern Med* 169(3):251–258
- Freeman BJ (2006) IDET: a critical appraisal of the evidence. *Eur Spine J* 15(Suppl 3):S448–S457
- Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 34(18):1929–1941
- Furlan AD, Tomlinson G, Jadad AA, Bombardier C (2008) Examining heterogeneity in meta-analysis: comparing results of randomized trials and nonrandomized studies of interventions for low back pain. *Spine (Phila Pa 1976)* 33(3):339–348
- Gibson J, Waddell G (2005) Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev* (4):CD001352
- Gibson JNA, Waddell G (2007) Surgical interventions for lumbar disc prolapse. *Cochrane Database Syst Rev* (2):CD001350
- Hestbaek L, Leboeuf-Yde C, Manniche C (2003) Low back pain: what is the long-term course? A review of studies of general patient populations. *Eur Spine J* 12(2):149–165
- Hirsch JA, Singh V, Falco FJ, Benyamin RM, Manchikanti L (2009) Automated percutaneous lumbar discectomy for the contained herniated lumbar disc: a systematic assessment of evidence. *Pain Physician* 12(3):601–620
- Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, Contopoulos-Ioannidis DG, Lau J (2001) Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 286(7):821–830

17. Jacobs WC, van Tulder M, Arts M, Rubinstein SM, van Middelkoop M, Ostelo R, Verhagen A, Koes B, Peul WC (2011) Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. *Eur Spine J* 20(4):513–522
18. Jacobs WC, Vreeling A, de Kleuver M (2006) Fusion for low-grade adult isthmic spondylolisthesis: a systematic review of the literature. *Eur Spine J* 15(4):391–402
19. Kabir SM, Gupta SR, Casey AT (2010) Lumbar interspinous spacers: a systematic review of clinical and biomechanical evidence. *Spine (Phila Pa 1976)* 35(25):E1499–E1506
20. Kovacs FM, Urrutia G, Alarcon JD (2011) Surgery versus conservative treatment for symptomatic lumbar spinal stenosis: a systematic review of randomized controlled trials. *Spine (Phila Pa 1976)* 36(20):E1335–E1351
21. Kuijpers T, van Middelkoop M, Rubinstein SM, Ostelo R, Verhagen A, Koes BW, van Tulder MW (2011) A systematic review on the effectiveness of pharmacological interventions for chronic non-specific low-back pain. *Eur Spine J* 20(1):40–50
22. Lambeek LC, van Tulder MW, Swinkels IC, Koppes LL, Anema JR, van Mechelen W (2011) The trend in total cost of back pain in The Netherlands in the period 2002 to 2007. *Spine (Phila Pa 1976)* 36(13):1050–1058
23. Lewis R, Williams N, Matar H, Din N, Fitzsimmons D, Phillips C, Jones M, Sutton A, Burton K, Nafees S, Hendry M, Rickard I, Chakraverty R, Wilkinson C (2011) The clinical effectiveness and cost-effectiveness of management strategies for sciatica: systematic review and economic model. *Health Technol Assess* 15(39):1–578
24. Martin CR, Gruszczynski AT, Braunsfurth HA, Fallatah SM, O’Neil J, Wai EK (2007) The surgical management of degenerative lumbar spondylolisthesis: a systematic review. *Spine (Phila Pa 1976)* 32(16):1791–1798
25. McGirt MJ, Ambrossi GL, Dato G, Sciubba DM, Witham TF, Wolinsky JP, Gokaslan ZL, Bydon A (2009) Recurrent disc herniation and long-term back pain after primary lumbar discectomy: review of outcomes reported for limited versus aggressive disc removal. *Neurosurgery* 64(2):338–344
26. Mirza SK, Deyo RA (2007) Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine (Phila Pa 1976)* 32(7):816–823
27. Moller H, Hedlund R (2000) Surgery versus conservative management in adult isthmic spondylolisthesis—a prospective randomized study: part I. *Spine (Phila Pa 1976)* 25(13):1711–1715
28. Moojen WA, Arts MP, Bartels RH, Jacobs WC, Peul WC (2011) Effectiveness of interspinous implant surgery in patients with intermittent neurogenic claudication: a systematic review and meta-analysis. *Eur Spine J* 20(10):1596–1606
29. Nellensteijn J, Ostelo R, Bartels R, Peul W, van Royen B, van Tulder M (2010) Transforaminal endoscopic surgery for symptomatic lumbar disc herniations: a systematic review of the literature. *Eur Spine J* 19(2):181–204
30. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korf M, Bouter LM, de Vet HC (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)* 33(1):90–94
31. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, Thomeer RT, Koes BW (2007) Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 356(22):2245–2256
32. Postacchini F, Cinotti G, Perugia D, Gumina S (1993) The surgical treatment of central lumbar stenosis. Multiple laminotomy compared with total laminectomy. *J Bone Joint Surg Br* 75(3):386–392
33. Pratt RK, Fairbank JC, Virr A (2002) The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 27(1):84–91
34. Revel M, Payan C, Vallee C, Laredo JD, Lassale B, Roux C, Carter H, Salomon C, Delmas E, Roucoules J (1993) Automated percutaneous lumbar discectomy versus chemonucleolysis in the treatment of sciatica. A randomized multicenter trial. *Spine (Phila Pa 1976)* 18(1):1–7
35. Savigny P, Kuntze S, Watson P, Underwood M, Ritchie G, Cotterell M, Hill D, Buchanan E, Coffey P, Dixon P, Drummond C, Flanagan M, Greenough C, Griffiths M, Halliday-Bell J, Hettinga D, Vogel S, Walsh D (2009) NICE clinical guideline 88 low back pain: early management of persistent non-specific low back pain. National Collaborating Centre for Primary Care and Royal College of General Practitioners, London
36. Shah RV, Albert TJ, Bruegel-Sanchez V, Vaccaro AR, Hilibrand AS, Grauer JN (2005) Industry support and correlation to study outcome for papers published in *Spine*. *Spine (Phila Pa 1976)* 30(9):1099–1104
37. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 7:10
38. Shojania KG, Bero LA (2001) Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. *Eff Clin Pract* 4(4):157–162
39. Transfeldt EE, Mehbood AA (2007) Evidence-based medicine analysis of isthmic spondylolisthesis treatment including reduction versus fusion in situ for high-grade slips. *Spine (Phila Pa 1976)* 32(19 Suppl):S126–S129
40. US National library of medicine Pubmed clinical queries filter (2011) 1 June 2011. <http://www.ncbi.nlm.nih.gov/pubmed/clinical>
41. van den Eerenbeemt KD, Ostelo RW, van Royen BJ, Peul WC, van Tulder MW (2010) Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature. *Eur Spine J* 19(8):1262–1280 [1432-0932 (Electronic), 0940-6719 (Linking)]
42. van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, van Tulder MW (2011) A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J* 20(1):19–39
43. van Tulder MW, Koes B, Seitsalo S, Malmivaara A (2006) Outcome of invasive treatment modalities on back pain and sciatica: an evidence-based review. *Eur Spine J* 15(Suppl 1):S82–S92
44. Walker BF (2000) The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *J Spinal Disord* 13(3):205–217
45. Watters WC III, Bono CM, Gilbert TJ, Kreiner DS, Mazanec DJ, Shaffer WO, Baisden J, Easa JE, Fernand R, Ghiselli G, Heggeness MH, Mendel RC, O’Neill C, Reitman CA, Resnick DK, Summers JT, Timmons RB, Toton JF (2009) An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. *Spine J* 9(7):609–614
46. Watters WC III, McGirt MJ (2009) An evidence-based review of the literature on the consequences of conservative versus aggressive discectomy for the treatment of primary disc herniation with radiculopathy. *Spine J* 9(3):240–257
47. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA (2006) Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 296(20):2441–2450

48. Yajun W, Yue Z, Xiuxin H, Cui C (2010) A meta-analysis of artificial total disc replacement versus fusion for lumbar degenerative disc disease. *Eur Spine J* 19(8):1250–1261 [1432-0932 (Electronic), 0940-6719 (Linking)]
49. Zucherman JF, Hsu KY, Hartjen CA, Mehalic TF, Implicito DA, Martin MJ, Johnson DR, Skidmore GA, Vessa PP, Dwyer JW, Puccio S, Cauthen JC, Ozuna RM (2004) A prospective randomized multi-center study for the treatment of lumbar spinal stenosis with the X STOP interspinous implant: 1-year results. *Eur Spine J* 13(1):22–31