REVIEW

CrossMark

Decompression with or without concomitant fusion in lumbar stenosis due to degenerative spondylolisthesis: a systematic review

M. L. Dijkerman[1] · G. M. Overdevest[1,2] · W. A. Moojen[1,2] · C. L. A. Vleggeert-Lankamp[1]

Abstract

Purpose The primary objective of this systematic review is to compare the outcome after decompression with and without concomitant instrumented fusion in patients with lumbar stenosis and degenerative spondylolisthesis. Does adding fusion to simple decompression lead to better results?

Methods PubMed, Embase, CENTRAL, Cochrane, Web of Science, CINAHL and Academic Search Premier were searched. All studies comparing outcome of decompression alone to decompression with concomitant-instrumented fusion in patients suffering from symptomatic lumbar stenosis with degenerative spondylolisthesis were included. Risk of bias was assessed using an adapted version of the Cowley checklist.

Results Eleven studies were included in the analysis involving 3119 patients in total. In the majority of studies, including two RCTs, clinical outcome of both patient groups was comparable regarding most clinical outcome measures.

Conclusion Currently there is not enough evidence that adding instrumented fusion to a decompression leads to superior outcomes compared to decompression only in patients with lumbar stenosis and degenerative spondylolisthesis. The most important clinical outcome measures, including the ODI, show comparable results. Therefore, the least invasive and least costly procedure, being decompression alone, is preferred in patients with low-grade spondylolisthesis with predominant leg pain.

Graphical abstract These slides can be retrieved under Electronic Supplementary Material.

✉ M. L. Dijkerman
m.l.dijkerman@lumc.nl

Extended author information available on the last page of the article

Introduction

Degenerative spondylolisthesis is a common indication for lumbar surgery in adults over 65 years. It is a condition where, due to facet arthropathy and intervertebral disc degeneration, one vertebra has slipped over the other [1]. This can eventually result in narrowing of the spinal canal

and neuroforamina. Compression of the cauda equina and nerve roots can lead to neurogenic claudication and radiculopathy. Back pain may coincide these complaints and can be attributed to the degenerative lumbar spine. Clinical symptoms vary amongst patients, but can have incapacitating effects on the patient's quality of life [2].

Degenerative spondylolisthesis is a common observation in patients suffering from the clinical symptoms of lumbar spinal stenosis; a 24% incidence of degenerative spondylolisthesis in patients with surgically treated lumbar stenosis has been reported [1, 3]. The severity of spondylolisthesis is commonly graded according to the Meyerding classification in case of degenerative spondylolisthesis; usually the slippage of the cephalic segment does not exceed 25% of the endplate of the adjacent caudal segment, and is therefore, mostly graded as Meyerding grade 1 or 'low-grade spondylolisthesis'.

Surgical treatment has proven to be more successful than conservative treatment [4]. However, the optimal surgical treatment in patients suffering from symptomatic lumbar stenosis, who also have degenerative spondylolisthesis, has long been a subject of debate. Decompression without fusion for lumbar stenosis in the presence of degenerative spondylolisthesis has been associated with postoperative instability of the spine [5–10]. This postoperative instability is presumed to be the result of the removal of the posterior elements, such as the spinous process, inter- and supraspinous ligaments and medial aspect of the facet joints. Therefore, concomitant instrumented fusion may result in better postoperative clinical and functional outcomes [5, 9, 11]. Furthermore, concomitant instrumented fusion may be beneficial in cases of lumbar stenosis presenting with predominant symptoms of low back pain [12]. Though, it remains debatable whether adding an instrumented spondylodesis to decompression is indeed required [13–16]. Since the population consists particularly of elderly patients who frequently have frailty and comorbidities, it is important to consider the necessity for more extensive surgery associated with concomitant instrumented fusion [17]. To contribute to the discussion on this controversy, the objective of this systematic review is to evaluate the quality of evidence with respect to the clinical outcome of decompression with and without concomitant instrumented fusion for patients with lumbar stenosis accompanied by degenerative spondylolisthesis.

Methods

Search strategy

A search was performed in PubMed, Embase, CENTRAL, Cochrane, Web of Science, CINAHL and Academic Search Premier. Pre-specified search terms were used; the

complete search strategy can be found in the Supplementary Appendix.

Inclusion criteria

Studies comparing decompression alone with decompression and concomitant fusion in patients with degenerative lumbar stenosis accompanied by degenerative spondylolisthesis were included. Randomized controlled trials, controlled cohort studies, both prospective and retrospective were included. From a methodological perspective, preferably only RCTs are to be included. Since RCTs or prospective studies on this subject are scarce, retrospective studies were also included. A prerequisite for inclusion was the evaluation of functional disability (e.g. Oswestry Disability Index), the sample size had to be more than 20 patients, follow-up had to be at least 2 months.

Exclusion criteria

Studies that were published before 1990 were excluded because decompression and fusion techniques were considered outdated. Furthermore, studies describing cases with isthmic spondylolisthesis were excluded. Studies concerning results of reoperations for degenerative spondylolisthesis were also excluded.

Selection

The selection process was carried out independently by two researchers (MD and GO) based on the aforementioned inlusion and exclusion criteria. Whenever disagreement arose about inclusion, a third reviewer (CVL) was consulted. A first selection of articles was made based on relevance of the article title and abstract. The full version of articles included in the first selection round was retrieved, after which a subsequent selection was made.

Data extraction

Data from the included studies regarding the following items were extracted: study characteristics (study design, sample size, description of intervention and control treatment, follow-up moments), demographic characteristics (mean age and gender), clinical characteristics (baseline characteristics, definition of degenerative spondylolisthesis, clinical and functional outcome using a validated outcome measure, complications, reoperation incidence).

Risk of bias assessment

The Cowley checklist [18], which was adapted for the purpose of this review (see Table 1), was used to assess the risk

of bias of the studies included in this review. The risk of bias assessment was carried out by two independent researchers (MD and CVL). A third reviewer (GO) was consulted in case of inconsistency. The studies were assessed on selection bias, for which four points per item could be obtained, and outcome bias, for which a maximum of six points could be attributed. Furthermore, studies were assessed on randomization bias and confounding bias, as well as whether a clear study objective and the independence of the investigators was stated. One point could be attributed to each item. In total, studies could be awarded a maximum of 13 points. Studies were then divided into low (11–13 points), medium (8–10 points) and high (7 or less) risk of bias.

The quality of evidence for all primary outcome measures was evaluated using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach [19].

Measures of treatment effect

The treatment effect is defined as the difference in outcome measures between the patient groups. Validated outcome measures that were included in the analysis were the Oswestry Disability Index (ODI), Short Form-36 (SF-36), Visual Analogue Scale (VAS), Numeric Rating Scale (NRS), European Quality of Life-5 Dimensions (EQ-5D), Core Outcome Measures Index (COMI) and Japanese Orthopaedic Association (JOA) score. Comparisons of continuous data are presented as mean differences with corresponding confidence intervals or standard deviations.

Included RCTs were assessed as to whether they were clinically and statistically homogeneous so that a quantitative analysis would generate meaningful data.

Due to the high potential risk for bias of retrospective studies, differences in outcome measures between patient groups were analyzed qualitatively rather than in quantitative manner.

To establish the minimal clinically important difference (MCID), the study by Ostelo et al. [20] was used as a reference. The minimal clinically important difference of outcome measures was defined as a 30% improvement compared to baseline. This corresponds to a difference of 15 on the Oswestry Disability Questionnaire (0–100), 15 on the Short Form-36 (PCS and MCS combined) (0–100), 2 on the Visual Analogue Scale (0–10), 2 on the Numeric Rating Scale (0–10), 3.5 on the complete Japanese Orthopaedic Association Scores (0–17), 0.2 on the European Quality of Life-5 Dimensions (0–1) and 2 on the Core Outcome Measures Index (0–10). These differences represent the minimum important change from baseline.

Table 1 Cowley's checklist for risk of bias assessment

Attrition bias	
Criteria for inclusion explicit	One point
Listhesis definition appropriately described	One point
Selection bias	
Number of men and women given	One point only if gender and age are reported properly
Age range and mean age reported	
Can selection-bias be ruled out	One point
Outcome bias	
Clearly defined criteria for measuring outcomes	One point
Patient-reported clinically relevant outcomes	One point
Preoperative status stated	One point
Follow-up period, range and mean given	One point
Valid statistical analysis undertaken	One point
Clinical evaluation independent of treating physician	One point only if clinical and radiological evaluation
Radiological evaluation independent and blinded to clinical results	were scored independently
Randomization bias	
Allocation bias	One point
Clear study objective	
	One point
Independence of investigators stated	
	One point

In total, studies could be awarded a maximum of 13 points. Studies were then divided into low (11–13 points), medium (8–10 points) and high (7 or less) risk of bias

Results

Search results

The literature search up to 26 May 2016 yielded 697 studies (Fig. 1). 683 studies were excluded based on the title and/or abstract. After evaluating the remaining 14 studies, 3 more studies were excluded (20–212). Bridwell et al. [21] used an outdated fusion technique (transverse process fusion with bone graft) and performed four different fusion techniques in 43 patients, making the results noncomparable to the other studies. Konno and Kikuchi [22] used a 'Graf' system, which was also not comparable to current fusion techniques. Kim et al. [23] did not report clinical or functional outcome measures and was, therefore, excluded. Finally, eleven studies remained eligible for inclusion. These studies were published between 1991 and 2016. A total of 3119 patients were included, with a mean age varying from 64 to 69 between studies.

Included studies

Two studies that were included were RCTs [24, 25]. Herkowitz and Kurz [5] was considered a quasi-RCT, since randomization methods were inadequate. Two studies were prospective controlled cohort studies [5, 26]. Seven of the studies were retrospective controlled cohort studies [2, 3, 7, 12, 26–29].

All studies compared decompression with or without concomitant fusion. In six studies, a laminectomy was performed [2, 7, 12, 25–28]. Two studies [2, 26] used a decompression technique with preservation of midline structures and Forsth et al. [24] used both techniques (laminectomy in 80% of the patients and a midline preserving technique in 20% of patients). The studies by Herkowitz and Kurz [5] and Ghogawala et al. [25] described the use of a laminectomy and concomitant partial facetectomy. In two studies [3, 7] the decompression technique was not specifically described.

Several, mostly instrumented, fusion techniques were compared with decompression alone. Ghogawala et al. [7], Kleinstueck et al. [12], Park et al. [28], Rampersaud [2] and Sigmundsson et al. [29] performed a posterior lumbar interbody fusion (PLIF) and Kleinstueck et al. [12] additionally used a transforaminal lumbar interbody fusion (TLIF) in a subgroup of patients. Four studies compared posterolateral fusion (PLF) with decompression alone [2, 12, 25, 26]. Forsth et al. [3] performed an instrumented fusion in 95% of the cases (5% was uninstrumented). Plotz and Benini [27] did a translaminar screw fixation and pedicular screw fixation with an AO internal fixator. The study of Herkowitz and Kurz [5] was the only study that did not make use of an instrumented fusion technique; instead it used a single-level bilateral intertransverse-process arthrodesis. Study characteristics of the included studies are described in Table 2.
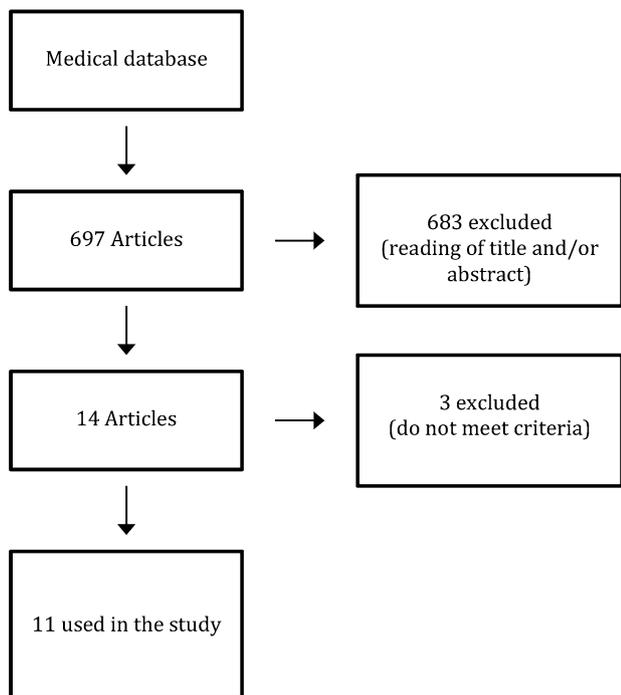
Risk of bias assessment

All studies were assessed on risk of bias using Cowley's checklist adapted for the subject of this review. An overview of the risk of bias assessment can be found in Table 3. Matsudaira [26] and Plotz and Benini [27] obtained seven points and are, therefore, considered having a high risk of bias. Rampersaud et al. [2], Forsth et al. [3], Herkowitz and Kurz [5] and Ghogawala et al. [7] scored a total of eight points; Park et al. [28] and Sigmundsson et al. [29] scored a total of nine points and Kleinstueck et al. [12] a total of ten points. These studies thus have a medium risk of bias. Forsth et al. [24] scored 11 points and Ghogawala et al. [25] scored the maximum amount of 13 points, and therefore, both studies are considered as having a low risk of bias. A critical note must be placed by these RCTs. Both studies were powered on a certain number of patients, however, due to loss to follow-up of patients, the conclusions that were drawn after 2–4 years of follow-up were not based on the number of patients that the studies were powered on.

Fig. 1 Selection of eligible articles

Table 2 Study characteristics

Study	Study type	Patient characteristics	Treatment
Prospective studies			
Forsth et al. [24]	Type: RCT Selection bias: no Randomization: yes Follow-up: 2 years	247 patients 155 female/92 male Mean age: 66.7	The method used for decompression and fusion surgery was determined solely by the surgeon
Ghogawala et al. [25]	Type: RCT Selection bias: no Randomization: yes Follow-up: 4 years	66 patients 53 female/13 male Mean age: 66.6	Group 1: complete laminectomy with partial removal of the medial facet joint Group 2: laminectomy with instrumented fusion
Herkowitz and Kurz [5]	Type: quasi-RCT Selection bias: no Randomization: yes Follow-up: 4 years	50 patients 36 female/14 male Mean age: 65 in group 1 and 63.5 in group 2	Group 1: bilateral laminectomy and facetectomy Group 2: single-level bilateral intertransverse-process arthrodesis
Matsudaira et al. [26]	Type: prospective cohort study Selection bias: yes Randomization: no Follow-up: 2 years	37 patients 22 female/15 male Mean age: 68 in group 1 and 67 in group 2	Group 1: decompression with preservation of the posterior elements Group 2: laminectomy with PLF
Retrospective studies			
Forsth et al. [3]	Type: retrospective cohort study Selection bias: no Randomization: no Follow-up: 2 years	1306 patients 956 female/350 male Mean age: 69	Group 1: decompression technique not specified Group 2: decompression and (un)instrumented fusion
Ghogawala et al. [7]	Type: retrospective cohort study Selection bias: no Randomization: no Follow-up: 1 year	34 patients 23 female/11 male Mean age: 68.8	Group 1: decompression technique not specified Group 2: decompression with PLIF
Kleinstueck et al. [12]	Type: retrospective cohort study Selection bias: no Randomization: no Follow-up: 1 year	213 patients 155 female/58 male Mean age: 69	Group 1: laminectomy, flavectomy, lateral recess decompression, partial medial facet resection and very limited laminectomies Group 2: decompression with TLIF, PLIF or PLF
Park et al. [28]	Type: retrospective cohort study Selection bias: yes Randomization: no Follow-up: 3 years	45 patients 37 female/8 male Mean age: 64	Group 1: unilateral laminotomy for bilateral compression (ULBD) Group 2: laminectomy with PLIF
Plotz and Benini [27]	Type: retrospective cohort study Selection bias: no Randomization: no Follow-up: 3 years	106 patients 72 female/34 male Mean age: 67	Group 1: undercutting decompression, laminectomy and decompression Group 2: undercutting decompression or laminectomy combined with translaminar screw fixation or internal pedicle fixator
Rampersaud et al. [2]	Type: retrospective cohort study Selection bias: yes Randomization: no Follow-up: 2 years	179 patients 125 female/54 male Mean age: 67.8 in group 1 and 62.47 in group 2	Group 1: midline-sparing, bilateral decompression from a unilateral approach using a tubular retractor system Group 2: decompression with PLF or PLIF
Sigmundsson et al. [29]	Type: retrospective cohort study Selection bias: yes Randomization: no Follow-up: 2 years	836 patients 646 female/190 male Mean age: 71.2 years	Group 1: decompression Group 2: decompression with PLIF

RCT randomized controlled trial, *PLF* posterolateral fusion, *PLIF* posterolateral instrumented fusion, *TLIF* transforaminal lumbar interbody fusion

Table 3 Risk of bias assessment

	Attrition bias (2)	Selection bias (2)	Outcome bias (6)	Randomization (1)	Clear study objective (1)	Independence of investigators (1)	Risk of bias
Forsth et al. [3]	**	**	*****	–	*	*	Medium
Forsth et al. [24]	*	**	*****	*	*	*	Low
Ghogawala et al. [7]	*	**	*****	–	*	–	Medium
Ghogawala et al. [25]	**	**	******	*	*	*	Low
Herkowitz and Kurz [5]	–	**	*****	–	*	*	Medium
Kleinstueck et al. [12]	*	**	******	–	*	*	Medium
Matsudaira et al. [26]	*	*	******	–	*	–	High
Park et al. [28]	**	*	*****	–	*	*	Medium
Plotz and Benini [27]	**	**	***	–	*	–	High
Rampersaud et al. [2]	*	*	******	–	*	–	Medium
Sigmundsson et al. [29]	*	*	******	–	*	–	Medium

Outcome

ODI

Four studies used the ODI as a primary outcome measure [3, 7, 24, 28]. The outcome measure was the difference between the decompression and fusion group regarding the ODI score (0–100). None of the studies, except for Ghogawala et al. [7], found a statistically significant difference in the ODI between the decompression group (D) and the decompression with fusion group (DF). Forsth et al. [3] found significant improvements in ODI scores in both the decompression and fusion group. ODI scores improved significantly from 45 (± 15) to 27 (± 1) in both groups, also meeting the MCID criteria. Park et al. [28] did not find a clinically meaningful improvement of the ODI [D: 29.8 (± 4.4) to 15.45 (± 7.06); DF: 24.6 (± 5.38) to 11 (± 7.09)], and the decrease was not significantly different between both groups (p = 0.96). Ghogawala et al. [7] demonstrated that the ODI score improved from 41.5 to 27.4 in the decompression group and from 41.5 to 14.0 in the fusion group (p = 0.02). Only in the fusion group the MCID of 15 was reached.

A quantitative analysis was performed for the RCTs from Forsth et al. [24] and Ghogawala et al. [25] (see Fig. 2). The outcome measure was the difference in the ODI scores between the two groups. In the study of Ghogawala et al. [25], patients in the decompression group went from an ODI of 36.3–18.4 in the first 2 years and to 21.6 after 4 years. In the fusion group, the ODI score went from 38.8 to 12.5 in 2 years and eventually to 15.1 after 4 years. No significant difference was found between both groups (p = 0.05). Forsth et al. [24] could not find a difference in improvements in ODI scores between the decompression and fusion group either. The ODI score went from 41 to 21 in the decompression group and from 41 to 25 in the fusion group (p = 0.11). When the mean differences of both studies are combined in a forest plot, the mean difference between both groups is − 1.45 (CI − 6.70; 3.81), with a moderate heterogeneity (I^2 = 52%). This illustrates that the pooled quantitative analysis of these two studies does not show a clinically nor statistically significant difference in ODI score improvement between patients treated with decompression and decompression with fusion.

Level of evidence The level of evidence is lowered by two levels, since only two of the included studies have a low risk of bias. Moreover, findings are inconsistent; four studies report no difference between decompression and instrumented fusion and one retrospective study claims that fusion has a better result based on ODI outcome. Therefore, the level of evidence is low.
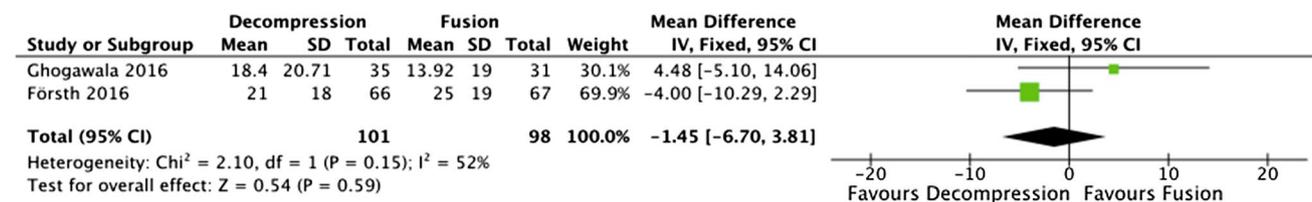

Fig. 2 Forest plot comparing differences in ODI scores between patients who received decompression with and without concomitant fusion

SF-36

In five studies, the SF-36 was used as a primary outcome [2, 7, 25, 28, 29]. The SF-36 score can be considered as a whole, or considering the physical component in the score separately (SF-36 Physical Component Summary), and/or the mental component in the score separately (SF-36 Mental Component Summary) [30].

Ghogawala used the SF-36 Physical Component Summary (SF-36 PCS) in both studies, while all other studies used the SF-36 as a whole, as is generally done. Both studies by Ghogawala found a significant difference in mean SF-36 PCS scores between the decompression and fusion group, while the other studies did not find a significant difference. In the study of Ghogawala et al. [7], the SF-36 PCS score in the decompression group went from 30.9 to 37.4 and in the fusion group from 29.8 to 45.7 after 1 year follow-up, $p = 0.003$. In the other study, there was no significant difference between both groups after 1 year of follow-up (D: 34.7–46; DF 31.5–46.8, $p = 0.16$). After 2 years, however, a significant difference between the decompression and fusion group was found (D: 34.7–44.2; DF: 31.5–46.7, $p = 0.046$). In both studies by Ghogawala, only the improvements of the SF-36 PCS scores in the fusion groups met the MCID criteria. Park et al. [28] found clinical meaningful improvements of both the SF-36 PCS and the SF-36 MCS scores, but there was no significant difference between groups ($p = 0.26$; $p = 0.25$). Rampersaud et al. [2] did not show a statistically significant increase in SF-36 PCS and MCS scores in both the decompression and fusion group ($p = 0.39$, $p = 0.06$). Sigmundsson [29] measured SF-36 scores in a predominant leg pain and predominant back pain group. In the leg pain group, there was no significant difference between the SF-36 PCS and the SF-36 MCS ($p = 0.16$; $p = 0.42$). In the back pain group, there was also no significant difference in the SF-36 and the SF-36 MCS ($p = 0.54$; $p = 0.09$). The exact improvements of the SF-36 PCS and MCS scores can be found in Table 4. The MCID was not met in any of the groups in this study. Differences in the reporting of SF-36 scores between studies did not allow for a quantitative comparison.

Level of evidence The level of evidence is lowered by two levels, since most studies have a medium to high risk of bias and findings were not consistent amongst studies. Therefore, the quality of evidence is low.

Leg pain

Leg pain was measured with the use of VAS, NRS and JOA back pain scores in six studies [3, 5, 12, 26, 28, 29]. None of the studies found a significant difference between the decompression and decompression and fusion group,

except Herkowitz and Kurz [5]. He states that patients who received decompression had significantly more leg pain compared to the group of patients who received arthrodesis. The decompression group showed an improvement in leg pain (scale 0–5) from 4.0 to 1.7, and an improvement from 4.3 to 1.0 in the fusion group. These improvements are both clinically significant, but the article states that the difference between the D and the DF group is significant, but a p value is not mentioned. The remaining five studies also showed significant improvements in leg pain, however, none of the studies demonstrated a statistical significant difference between the decompression and decompression and fusion group. Forsth et al. [3] showed a clinically relevant improvement in VAS leg pain scores after 2 years (D: 63 ($\pm$ 25) to 35 (CI 32–37); DF 62 ($\pm$ 26) to 32 (CI 30–35), $p = 0.17$). Kleinstueck et al. [12] measured a similar reduction in leg pain intensity (scale 0–10) in the decompression and fusion group from 3.1 ($\pm$ 3.0) (D) and 3.9 ($\pm$ 3.4) (DF), with baseline scores of 6.5 ($\pm$ 2.3) (D) and 6.2 ($\pm$ 2.7) (DF), $p = 0.13$. In the study of Matsudaira et al. [26], both groups showed a significant improvement in JOA leg pain scores (D: 1.0 ($\pm$ 0.4) to 1.8 ($\pm$ 0.9); DF: 1.1 ($\pm$ 0.6) to 2.2 ($\pm$ 0.7), $p = 0.5208$). Park et al. [28] showed comparable improvements in NRS leg pain scores in both groups [D: 7.8 ($\pm$ 0.91) to 2.4 ($\pm$ 2.53); DF 8.0 ($\pm$ 0.87) to 2.5 ($\pm$ 1.80), $p = 0.99$]. In the patient group with predominant leg pain, Sigmundsson et al. [29] also found significant improvements in VAS scores (scale 0–100) of leg pain after 2 years, without a difference between the decompression and decompression and fusion group. The decompression group had a reduction in leg pain of 36.5 ($\pm$ 36.2) [baseline 71.3 ($\pm$ 20.1)] and the fusion group a reduction of 43 ($\pm$ 34.1) [baseline 73.8 ($\pm$ 18.9)], $p = 0.24$. The improvements in leg pain scores met the MCID criteria in all six studies.

Level of evidence The level of evidence is lowered by one level, since there were no studies included with a low risk of bias. Although Herkowitz and Kurz [5] claim that there is a significant difference in leg pain between both groups, no p value was given and the difference in leg pain between both groups does not seem clinically relevant (with a difference of 0.7 on a 5 point scale). It can, therefore, be concluded that none of the studies found a difference in leg pain between both groups. The level of evidence is, therefore, moderate.

Back pain

All six studies that measured improvements in leg pain scores, scored improvement of back pain as well. Four studies found a significant better outcome in the fusion group regarding back pain [5, 12, 28, 29]. Herkowitz and Kurz [5] rated pain on a scale from 0 to 5 (0 = no pain, 5 = severe pain). He reports back pain at baseline

Table 4 Study outcomes

Primary outcome	Level of evidence	Study	Results (D = decompression alone, DF = decompression + fusion)	Conclusion
ODI[a]	Low	Forsth et al. [24]	D: change in ODI (41–21) DF: change in ODI (41–25)	No significant difference ($p = 0.11$)
		Ghogawala et al. [25]	D (after 2 years): change in ODI (36.3–18.4) D (after 4 years): change in ODI (36.3–21.6) DF (after 2 years): change in ODI (38.8–12.5) DF (after 4 years): change in ODI (38.8–15.1)	No significant difference after 2 years ($p = 0.06$) and 4 years ($p = 0.05$)
		Forsth et al. [3]	D: change in ODI (45–27) DF: change in ODI (45–27)	No significant difference ($p = 0.93$)
		Ghogawala et al. [7]	D: change in ODI (41.0–27.4; $p = 0.003$) DF: change in ODI (41.5–14.0; $p < 0.001$)	Significant difference ($p = 0.02$)
		Park et al. [28]	D: change in ODI (29.8–15.45) DF: change in ODI (24.6–11.0)	No significant difference ($p = 0.96$)
		Sigmundsson et al. [29]	D: predominant leg pain: change in ODI (43.3–25) Predominant back pain: change in ODI (44.9–30.4) DF: predominant leg pain: change in ODI (44.3–24.4) Predominant back pain: change in ODI (46.4–28.8)	Predominant leg pain: no significant difference ($p = 0.71$) Predominant back pain: no significant difference ($p = 0.20$)
SF-36[b]	Very low	Ghogawala et al. [25]	D (after 2 years): change in SF-36 PCS (34.7–44.2) D (after 4 years): change in SF-36 PCS (34.7–42.1) DF (after 2 years): change in SF-36 PCS (31.5–46.7) DF (after 4 years): change in SF-36 PCS (31.5–45.6)	Significant difference after 2 years ($p = -0.046$) and 4 years ($p = 0.02$)
		Ghogawala et al. [7]	D: change in SF-36 PCS (30.9–37.4; $p = 0.005$) DF: change in SF-36 PCS (29.8–45.7; $p < 0.001$)	Significant difference ($p = 0.003$)
		Park et al. [28]	D: change in SF-36 PCS (29.2–47.2) and SF-36 MCS (28.0–46.7) DF: change in SF-36 PCS (26.1–46.3) and SF-36 MCS (29.3–44.5)	No significant difference ($p = 0.26$, $p = 0.25$)
		Rampersaud et al. [2]	D: change in SF-36 PCS (28.89–39.02). and SF-36 MCS (42.91–50.23) DF: change in SF-36 PCS (29.97–41.39 (10.59) and SF-36 MCS (46.78–50.94)	No significant difference ($p = 0.39$, $p = 0.06$)
		Sigmundsson et al. [29]	D: predominant leg pain: change in SF-36 PCS (37.9–48) and SF-36 MCS (29.2–36.5) Predominant back pain: change in SF-36 PCS (38.2–44.1) and SF-36 MCS (29.0–33.8) DF: predominant leg pain: change in SF-36 PCS (40.7–47.4) and SF-36 MCS (28.2–36.8) Predominant back pain: change in SF-36 PCS (38.6–44.8) and SF-36 MCS (27.5–35.3)	Predominant leg pain: no significant difference ($p = 0.16$, $p = 0.42$) Predominant back pain: no significant difference ($p = 0.54$, $p = 0.09$)

Table 4 (continued)

Primary outcome	Level of evidence	Study	Results (D = decompression alone, DF = decompression + fusion)	Conclusion
Back and leg pain	Low/moderate	Forsth et al. [3]	D: VAS[c] back (54–35), VAS leg (63–35) DF: VAS back (61–32) VAS leg (62–32)	No significant difference (back pain: $p = 0.12$, leg pain: $p = 0.17$)
		Herkowitz and Kurz [5]	D: change in back pain (2.9–2.5) and leg pain (4.0–1.7) DF: change in back pain (3.3–1.3) and leg pain (4.3–1.0)	Significant difference ($p < 0.01$).
		Kleinstueck et al. [12]	D: reduction back pain: 1.7 (baseline 4.1) and leg pain: 3.1 (baseline 6.5) DF: reduction back pain: 2.9 (baseline 5.3); reduction leg pain: 3.9 (baseline 6.2)	Significant difference reduction in back pain ($p = 0.01$) No significant difference reduction in leg pain ($p = 0.13$)
		Matsudaira et al. [26]	D: change in back pain (1.4–2.5, $p < 0.0001$) and leg pain (1.1–2.2, $p < 0.0001$) DF: change in back pain (1.4–2.4, $p = 0.0001$) and leg pain (1.0–1.8, $p = 0.0001$)	No significant difference ($p = 0.69$, $p = 0.52$)
		Park et al. [28]	D: change in NRS[d] back (2.8–3.1) and NRS leg (7.8–2.4) DF: change in NRS back (6.6–2.4) and NRS leg (8.0–2.5)	Significant difference NRS back ($p = 0.001$). No significant difference NRS leg ($p = 0.99$)
		Sigmundsson et al. [29]	D: predominant leg pain: Change in back pain (38.4–30.3) and leg pain (71.3–34.8) Predominant back pain: change in back pain (65.6–38.3) and LP (55.8–35.5) DF: predominant leg pain: Change in back pain (50.8–28.5) and leg pain (73.8–30.8) Predominant back pain: change in back pain (69.4–35.8) and leg pain (56.6–30.8)	Predominant leg pain: significant difference change in back pain ($p = 0.0008$). No significant difference change in leg pain ($p = 0.24$) Predominant back pain: no significant differences ($p = 0.17$, $p = 0.38$)
Good outcome	Low	Herkowitz and Kurz [5]	D: excellent (8%) good (36%), fair (48%), poor (8%) DF: excellent (44%), good (52%), fair (4%), poor (0%)	Significant difference ($p = 0.0001$)
		Kleinstueck et al. [12]	D: 70.4% good outcome DF: 86.2% good outcome	Significant difference ($p = 0.01$)
EQ-5D[e]	No evidence	Forsth et al. [3]	D: EQ-5D (0.36–0.63) DF: EQ-5D (0.33–0.62)	No significant difference ($p = 0.34$)
		Sigmundsson et al. [29]	D: predominant leg pain: change in EQ-5D (0.34–0.65) Predominant back pain: change in EQ-5D (0.33–0.56) DF: predominant leg pain: change in EQ-5D (0.34–0.66) Predominant back pain: change in EQ-5D (0.30–0.59)	Predominant LP: no significant difference ($p = 0.63$) Predominant LBP: no significant difference ($p = 0.41$)

Table 4 (continued)

Primary outcome	Level of evidence	Study	Results (D = decompression alone, DF = decompression + fusion)	Conclusion
COMI[f]	No evidence	Kleinstueck et al. [12]	D: reduction in COMI: 3.1 (baseline 7.0) DF: reduction in COMI: 4.2 (baseline 7.6)	Significant difference ($p = 0.009$)

[a]Oswestry Disability Index (ranging from 0 to 100 with a higher score being an indication of more severe disability)

[b]Short Form-36 (with a range from 0 to 100, with higher scores demonstrating less disability). PCS: physical component summary, MCS: mental component summary

[c]Visual Analogue Scale (with a range from 0 to 10, with higher scores demonstrating more pain)

[d]Numeric Rating Scale (with a range from 0 to 10, with higher scores demonstrating more pain)

[e]European Quality of Life-5 Dimensions (with a range from 0 to 1, with a higher score illustrating a higher quality of life)

[f]Core Outcome Measures Index (with a range from 0 to 10, with higher scores demonstrating worst outcome)

of 2.9 (D) and 3.3 (DF), that decreases to 2.5 (D) and 1.3 (DF), $p < 0.01$. Only the reduction of back pain in the fusion group met the MCID criteria. Kleinstueck et al. [12] measured low back pain intensity at baseline in both groups (scale 0–10). The decompression group had a reduction in low back pain of 1.7 ($\pm$ 3.4), which was not clinically relevant [baseline score 4.1 ($\pm$ 3.0)]. The fusion group did show a clinical relevant reduction of 2.9 ($\pm$ 2.9), with a baseline score of 5.3 ($\pm$ 2.9). There was a significant statistical difference found between both groups ($p = 0.01$). However, patients in the fusion group had a significant higher preoperative back pain score compared to the decompression group ($p = 0.04$). Park used the NRS pain score and reports only minor back pain in the decompression group, which hardly changed after surgery. However, in the fusion group, preoperative back pain is clinically relevant (NRS = 6.6) and, significantly decreases to 2.4, meeting the MCID criteria. Sigmundsson et al. [29] showed a significant reduction in VAS scores (scale 0–100) of back pain after 1 year, with a reduction of 29.5 ($\pm$ 30.1) in the decompression group [baseline score 65.6 ($\pm$ 22.5)] and a reduction of 39.1 ($\pm$ 32.6) in the fusion group [baseline score 69.4 ($\pm$ 21.2)], $p = 0.005$. However, after 2 years the difference in reduction between both groups was not significant anymore [D: 27.3 ($\pm$ 31.2); DF: 33.6 ($\pm$ 27.4), $p = 0.17$].

Two studies showed a comparable reduction of back pain in both the decompression and fusion group. Forsth et al. [3] showed a clinical relevant decrease in VAS back pain scores [D: 54 ($\pm$ 27) to 35 (CI 32–37); DF 61 ($\pm$ 25) to 42 (CI 30–34), $p = 0.12$]. Matsudaira et al. [26] also found a clinical relevant improvement of the JOA back pain scores in both groups [D: 1.4 ($\pm$ 0.7) to 2.5 ($\pm$ 0.5); DF: 1.4 ($\pm$ 0.6) $\rightarrow$ 2.4 ($\pm$ 0.6), $p = 0.69$].

Level of evidence Since all studies have a medium to high risk of bias and outcomes are not consistent amongst studies, with four studies showing a statistical difference

between groups and two studies that did not, the level of evidence is lowered by two levels. The level of evidence is, therefore, low.

Perceived recovery

Both Herkowitz and Kurz [5] and Kleinstueck et al. [12] asked patients about their self-perceived recovery. Patients included in the study by Herkowitz and Kurz [5] had to rate the operative results as excellent, good, fair or poor. Kleinstueck et al. [12] used a five-point Likert scale (1 = surgery made things worse, 5 = surgery helped a lot). Herkowitz and Kurz [5] reported that 96% of patients had a good or excellent outcome in the fusion group, compared to 42% in the decompression group, $p = 0.0001$. Kleinstueck also demonstrated a significant higher percentage of good outcomes in the fusion group (86.2%, CI 80–92) than in the decompression group [70.4% (CI 58–83)], $p = 0.01$.

Level of evidence Because both studies have a medium risk of bias and direct comparison is not possible due to the fact that Herkowitz used a non-instrumented fusion technique, the level of evidence is lowered by two levels and is, therefore, considered to be low.

EQ-5D

Two studies measured clinical outcome with the EQ-5D [3, 29]. Sigmundsson et al. [29] only found a significant difference in improvement of EQ-5D in the predominant back pain group after 1 year, with an improvement of 0.24 ($\pm$ 0.42) in the decompression group [baseline 0.33 ($\pm$ 0.32)] and 0.34 ($\pm$ 0.35) in the fusion group [baseline 0.30 ($\pm$ 0.32)], $p = 0.04$. However, after 2 years there was no significant difference in EQ-5D scores in this group anymore ($p = 0.41$). In the group with predominant leg pain, there was also no significant difference between the decompression [improvement

of 0.31 (± 0.31)] and the fusion group [improvement of 0.32 (± 0.28)], $p = 0.56$. Försth found a comparable improvement of the EQ-5D in both groups, with the EQ-5D improving from 0.36 (± 0.32) to 0.63 (CI 0.61–0.66) in the decompression group and from 0.33 (± 0.31) to 0.62 (CI 0.59–0.64) in the fusion group, $p = 0.34$. The improvement of the EQ-5D scores met the MCID criteria in both studies.

Level of evidence Since there were only retrospective studies with a high risk of bias included and no RCTs, there is no evidence for this outcome.

COMI

Only one study used the COMI summary score (scale 0–10) as a primary outcome score [12]. A significant difference in reduction of the COMI score was found, with COMI scores reducing with 3.1 (± 2.9) points in the decompression group [baseline 7.0 (± 2.1)] and 4.2 (± 2.7) in the fusion group [baseline 7.6 (± 1.7)], $p = 0.009$. These scores met the MCID criteria.

Level of evidence Since only one study was included, which is a retrospective study with a high risk of bias; there is no evidence for this outcome.

Complications

Six studies reported on postoperative surgical and medical complications. Ghogawala et al. [7], Forsth et al. [24] and Matsudaira et al. [26] found a larger number of complications in the fusion group. Forsth et al. [24] showed a complication rate of 24% in the fusion group compared to 19% in the decompression group. In the study of Ghogawala et al. [7], 14% of the patients in the fusion group and 5% of the patients in the decompression group had postoperative complications. Matsudaira et al. [26] reported only one complication in the fusion group (5.3%). Kleinstueck et al. [12], Ghogawala et al. [25] and Park et al. [28] showed almost comparable complication rates in both treatment groups. In the study of Ghogawala et al. [25], the complication rate in the decompression group was 6%, compared to 3% in the fusion group. Kleinstueck et al. [12] showed a 17.9% complication rate in the decompression group and 17.2% in the fusion group. Park et al. [28] had 5% complications in the decompression group and 4% in the fusion group. Complications that were reported on included dura lesion, wound infection, postoperative hemorrhage, recurrent pain, adjacent segment stenosis and instrumentation related complications.

Level of evidence The level of evidence is lowered by three levels, since studies showed conflicting results, have different definitions of complications (some including persistent pain after lumbar decompression alone) and non-standardized assessment of complication incidence. The level of evidence is, therefore, very low.

Reoperations

Five studies reported reoperation rates. In three studies, a higher number of reoperations were performed in the decompression group. Ghogawala et al. [25] found a reoperation rate of 34% in the decompression group compared to 14% in the fusion group. An earlier study by Ghogawala et al. [7] reported a reoperation rate of 15% in the decompression group, while no reoperations were performed in the fusion group. In the study of Plotz and Benini [27], 65% of the patients in the decompression group needed reoperation. In the fusion group, 6% needed reoperation. Two studies found a higher number of reoperations in the fusion group. Matsudaira et al. [26] reported one case of reoperation (5%) in the fusion group and Rampersaud et al. [2] performed reoperation in 36% of patients who received fusion and in 11% of patients who received decompression. The study of Forsth et al. [24] showed similar reoperation rates in both groups, with 21% in the decompression group and 22% in the fusion group. No statistical significant differences were found, because of the small number of reoperations.

Level of evidence The level of evidence is lowered by three levels, since indications for reoperations are poorly reported on, results are conflicting amongst studies and the majority of included studies has a high risk of bias. Thus, the level of evidence is very low.

An overview of study outcomes can be found in Table 4.

Discussion

This review concludes that decompression with or without concomitant fusion seems to be leading to comparable results regarding the most important clinical outcome measures. The quality of evidence, however, is minimal, with the quality of evidence for comparable postoperative leg pain being moderate, while the quality of evidence for the other outcome parameters is low to very low. Although this review cannot conclude decompression has equal results to decompression with fusion in all clinical cases, the findings illustrate that adding fusion to a decompression as usual care in a degenerative spondylolisthesis stenosis does not inevitably lead to better clinical outcomes.

Multiple clinical outcome measures that were used in 11 studies were evaluated in this review. The ODI being the most specific clinical outcome for neurogenic claudication is used internationally as a disease-specific measure of functional disability. The absence of a difference in ODI outcome

between the two treatment strategies in the pooled quantitative analysis of two recently published RCTs is a strong indicator that there is no difference in functional outcome. The added qualitative analysis of three retrospective studies showing comparable improvements in both the decompression and decompression and fusion group, adds to this conclusion. One relatively small study by Ghogawala et al. [7] showed a significant better outcome regarding the ODI in favour of fusion, however, this conclusion was drawn after a relatively short follow-up period of 1 year. Ghogawala fails to show a significant difference between the two groups in the ODI score in the later RCT.

Ghogawala et al. found a significant difference in SF-36 PCS scores between the decompression and fusion group in two of their studies, with the most recent one being a RCT. This difference is presented as evidence for preference to add instrumented spondylodesis to decompression in degenerative spondylolisthesis stenosis. The SF-36 was not the primary outcome measure of this study and after 4 years only two-third of the original population was included in the analysis. The study was not powered on this number of people. Besides, The SF-36 is a very generic scale and the SF-36 physical component represents only a part of the functionality scale SF-36. It is by no means a standard parameter to evaluate outcome in neurogenic claudication. Moreover, three other studies that used the SF-36 PCS did not find any difference between fusion and decompression alone.

Regarding leg pain, all studies found that leg pain improved after surgery, whether that would be decompression or fusion, except for Herkowitz and Kurz [5] who found significantly better results in the fusion group. This outcome is disputable, since an extensive decompression with removal of the facet joints was performed. This may have contributed to the worse outcome in the decompression alone group, since extensive bony decompression may induce spinal instability. Furthermore, the generalizability of the results of this study is limited because non-instrumented fusion procedures are rarely performed in current clinical practice.

All studies that evaluated leg pain also looked at the improvement of back pain in their patients groups. A number of studies showed that back pain improved more in the fusion group than in the decompression alone group, which might lead to the conclusion that patients with degenerative spondylolisthesis with predominant back might benefit from an additional fusion surgery. Yet most studies were not able to show clinical meaningful differences between the patient groups. In the retrospective study by Park et al. [28], patients with low back pain tended to be preferentially included in the fusion group, which was based on the surgeon's decision. In the study by Kleinstueck et al. [12] patients in the fusion group had significantly more back pain compared to patients who received decompressive surgery, which adds

to the assumption that outcomes are biased due to the allocation of patients. Since the six studies that evaluated back pain were not randomized and results were conflicting, the evidence that back pain would decrease more in patients who received fusion is low. Prospective studies are necessary to obtain higher levels of evidence showing whether or not adding a fusion leads to a significant decrease in lower back pain in patients with degenerative spondylolisthesis.

Herkowitz and Kurz [5] and Kleinstueck et al. [12] were the only two studies reporting on perceived recovery. Both studies showed that patients had a better recovery after receiving fusion. Results must be interpreted with cause as baseline symptom severity, different patient characteristics among treatment groups, and non-blinding of the patients can influence self-perceived recovery. This outcome measure is reliable only in randomized studies or high-quality prospective controlled cohort studies that are preferably blinded to patients for treatment allocation.

Reoperation rates varied considerably among studies. Five of the included studies reported reoperation rates. Rampersaud et al. [2] and Matsudaira et al. [26] found a higher reoperation rate after fusion, which is presumed to be the result of a higher complication incidence that requires reoperation and instrumentation related reoperations. However, Ghogawala et al. [7], Plotz and Benini [27] and Park et al. [28] found a higher reoperation rate after decompression alone. This might be due to a higher risk of surgery-induced instability, but also includes patients with unsatisfactory relief of symptoms after decompression alone. Unfortunately, indications for reoperations were poorly described. Furthermore, if reoperation with concomitant fusion for unsatisfactory recovery after decompression alone is considered an outcome measure it is of particular importance to assess the outcome after these reoperations. Unfortunately, these are not reported.

Six studies included in this review reported on complications after decompression with or without concomitant fusion. The studies by Ghogawala et al. [7], Forsth et al. [24] and Matsudaira et al. [26] found higher complication rates in the fusion group, whereas Kleinstueck et al. [12], Ghogawala et al. [25] and Park et al. [28] found similar rates in both groups. Complications were assessed in a non-standardized way, in which some studies reported recurrent pain as a complication. None of the studies were able to find a significant difference, probably since the incidence of complications was too low to detect a statistically significant difference. Older studies that were not included in this review showed higher complication rates in the fusion group [9, 31], with a more recent study by Deyo et al. [17] clearly showing significant higher complications rates in the decompression and fusion group compared to the decompression alone group. After adjusting for age, comorbidities and other factors, there was a higher risk of life-threatening

life-events for patients receiving decompression with concomitant fusion.

Since there is only low evidence that adding fusion does not necessarily lead to better clinical outcomes, other factors that can lead to deciding whether or not to perform instrumented fusion should be taken into account as well. Cost-effectiveness studies conclude that decompression alone is associated with lower costs than decompression with (non-)instrumented fusion. A non-randomized study from the US calculated mean hospital costs of $12,615 for decompression alone, $18,495 for non-instrumented fusion and $25,914 for instrumented fusion [32]. Another study calculated similar costs, with decompression costing $14,700, non-instrumented fusion $21,500 and instrumented fusion $30,200 [33]. More recently, Deyo et al. [17] calculated that the adjusted mean hospital charges for fusion procedures were $80,888 for a complex fusion and $58,511 for a simple fusion compared to $23,724 for decompression alone. The RCT of Forsth et al. [24] showed that direct costs are $6800 higher for fusion. Decompression alone is, therefore, clearly the most cost-effective technique.

Although it seems that results of decompression and fusion are equal, possibly there are subgroups of patients that may benefit from decompression in combination with fusion. Most studies in this review included patients with spinal canal stenosis in combination with degenerative spondylolisthesis. Besides spinal canal stenosis, there are also patients who suffer from foraminal stenosis, induced by the spondylolisthesis. To adequately treat foraminal stenosis, the facet joint has to be severely reduced, which increases the risk of surgically induced instability. Therefore, in those cases, additional instrumented fusion is strongly recommended [6, 29, 34]. Another reason to recommend additional instrumented fusion might be preoperative hypermobility at the level of the spondylolisthesis as indicated by a dynamic radiograph. A review by Leone et al. [35] concludes that patients with symptomatic lumbar stenosis and a slip of more than 3 mm should receive concomitant fusion. However, in a study they performed in 2009 [36], they conclude that it is a challenge to determine a relationship between imaging instability and its symptoms. This is partly due to the fact that measurement errors may exist in plain X-rays. In a study by Herron and Mangelsdorf [14] concomitant fusion was routinely performed in patients with a greater than 3 mm slip. Most of the included studies in this review included patients with a minimum slip of 3 mm. Forsth et al. [24] reports a 7.4 mm slip in both the decompression and decompression with concomitant fusion group without showing differences in clinical outcome after surgery. In both studies performed by Ghogawala no significant difference in slippage was found in both groups,

with a slippage up to 8.5 mm and a maximum translation of 1.6 mm. They found significant differences in both the ODI and the SF-36 scores. Further research is necessary to identify these anatomical characteristics that justify additional instrumented fusion.

Although lumbar stenosis with degenerative spondylolisthesis is frequently an indication for lumbar surgery, no consensus for the optimal surgical management exists in the current literature. American guidelines consider both decompression alone and decompression with concomitant fusion as effective for the treatment of lumbar stenosis with concomitant stenosis. However, decompression with concomitant instrumented fusion is still common practice [37, 38]. With this review, demonstrating that decompression and decompression and fusion are equally effective in treating lumbar degenerative spondylolisthesis, especially regarding the ODI, being the most important clinical outcome measure, and leg pain, with higher costs and presumably higher complications rates associated with decompression with fusion, we hope to add to the discussion.

Conclusion

Currently, there is not enough evidence that concomitant instrumented fusion in patients with symptomatic lumbar spinal canal stenosis and degenerative spondylolisthesis leads to better clinical outcome than decompression alone. Decompression alone is a more cost-effective technique and is presumably more associated with fewer complications compared to decompression with concomitant fusion. This might be a fair reason to choose for decompression alone in patients with low-grade spondylolisthesis with predominant leg pain. These results must be interpreted with caution as not all cases of degenerative spondylolisthesis may be managed with decompression alone. Patients with high-grade spondylolisthesis or low-grade spondylolisthesis in combination with foraminal stenosis or vertebral instability may still benefit from concomitant fusion.

Compliance with ethical standards

References

1. Postacchini F et al (1991) Degenerative lumbar spondylolisthesis. II. Surgical treatment. Ital J Orthop Traumatol 17(4):467–477
2. Rampersaud YR et al (2014) Health-related quality of life following decompression compared to decompression and fusion for degenerative lumbar spondylolisthesis: a Canadian multicentre study. Can J Surg 57(4):E126–E133
3. Forsth P et al (2013) Does fusion improve the outcome after decompressive surgery for lumbar spinal stenosis? A two-year follow-up study involving 5390 patients. Bone Jt J 95-B(7):960–965
4. Weinstein JN et al (2007) Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 356(22):2257–2270
5. Herkowitz HN, Kurz LT (1991) Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Jt Surg Am 73(6):802–808
6. Fischgrund JS et al (1997) 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. Spine (Phila Pa 1976) 22(24):2807–2812
7. Ghogawala Z et al (2004) Prospective outcomes evaluation after decompression with or without instrumented fusion for lumbar stenosis and degenerative Grade I spondylolisthesis. J Neurosurg Spine 1(3):267–272
8. Kornblum MB et al (2004) Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. Spine (Phila Pa 1976) 29(7):726–733
9. Martin CR et al (2007) The surgical management of degenerative lumbar spondylolisthesis: a systematic review. Spine (Phila Pa 1976) 32(16):1791–1798
10. Okuda S et al (2007) Surgical outcomes of posterior lumbar interbody fusion in elderly patients. Surgical technique. J Bone Jt Surg Am 89(Suppl 2 Pt.2):310–320
11. Mardjetko SM et al (1994) Degenerative lumbar spondylolisthesis. A meta-analysis of literature 1970–1993. Spine (Phila Pa 1976) 19(20 Suppl):2256s–2265s
12. Kleinstueck FS et al (2012) To fuse or not to fuse in lumbar degenerative spondylolisthesis: do baseline symptoms help provide the answer? Eur Spine J 21(2):268–275
13. Dall BE, Rowe DE (1985) Degenerative spondylolisthesis. Its surgical management. Spine (Phila Pa 1976) 10(7):668–672
14. Herron LD, Mangelsdorf C (1991) Lumbar spinal stenosis: results of surgical treatment. J Spinal Disord 4(1):26–33
15. Kristof RA et al (2002) Degenerative lumbar spondylolisthesis-induced radicular compression: nonfusion-related decompression in selected patients without hypermobility on flexion-extension radiographs. J Neurosurg 97(3):281–286
16. Kelleher MO et al (2010) Success and failure of minimally invasive decompression for focal lumbar spinal stenosis in patients with and without deformity. Spine (Phila Pa 1976) 35(19):E981–E987
17. Deyo RA et al (2010) Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. J Am Med Assoc 303(13):1259–1265
18. Cowley DE (1995) Prostheses for primary total hip replacement: a critical appraisal of the literature. Int J Technol Assess Health Care 11(04):770–778
19. Atkins D et al (2004) Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. BMC Health Serv Res 4(1):38
20. Ostelo RW et al (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
21. Bridwell KH et al (1993) The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. J Spinal Disord 6(6):461–472
22. Konno S, Kikuchi S (2000) Prospective study of surgical treatment of degenerative spondylolisthesis—comparison between decompression alone and decompression with Graf system stabilization. Spine 25(12):1533–1537
23. Kim S et al (2012) Cost-utility of lumbar decompression with or without fusion for patients with symptomatic degenerative lumbar spondylolisthesis. Spine J 12(1):44–54
24. Forsth P et al (2016) A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. N Engl J Med 374(15):1413–1423
25. Ghogawala Z et al (2016) Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. N Engl J Med 374(15):1424–1434
26. Matsudaira K et al (2005) Spinal stenosis in grade I degenerative lumbar spondylolisthesis: a comparative study of outcomes following laminoplasty and laminectomy with instrumented spinal fusion. J Orthop Sci 10(3):270–276
27. Plotz GMJ, Benini A (1998) Lumbar degenerative spondylolisthesis: review of 106 operated cases with degenerative anterior vertebral translation as the predominant aspect of the spondylosis. Neurosurg Q 8(4):271–287
28. Park JH et al (2012) A comparison of unilateral laminectomy with bilateral decompression and fusion surgery in the treatment of grade I lumbar degenerative spondylolisthesis. Acta Neurochir (Wien) 154(7):1205–1212
29. Sigmundsson FG et al (2015) Outcome of decompression with and without fusion in spinal stenosis with degenerative spondylolisthesis in relation to preoperative pain pattern: a register study of 1,624 patients. Spine J 15(4):638–646
30. Carreon LY et al (2010) Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. Spine J 10(6):469–474
31. Weir B et al (1981) Lumbar stenosis: analysis of factors affecting outcome in 81 surgical cases. Can J Neurol Sci 8(4):295–298
32. Katz JN et al (1997) Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. Spine (Phila Pa 1976) 22(10):1123–1131
33. Kuntz KM et al (2000) Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. Spine (Phila Pa 1976) 1(25):1132–1139
34. Sengupta DK, Herkowitz HN (2003) Lumbar spinal stenosis. Treatment strategies and indications for surgery. Orthop Clin N Am 34(2):281–295
35. Leone A et al (2007) Lumbar intervertebral instability: a review. Radiology 245(1):62–77
36. Leone A et al (2009) Degenerative lumber intervertebral instability: what is it and how does imaging contribute? Skelet Radiol 38(6):529–533
37. Kepler CK et al (2014) National trends in the use of fusion techniques to treat degenerative spondylolisthesis. Spine 39(19):1584–1589
38. Fox MW et al (1996) Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. J Neurosurg 85(5):793–802

Affiliations

M. L. Dijkerman[1] · **G. M. Overdevest[1,2]** · **W. A. Moojen[1,2]** · **C. L. A. Vleggeert-Lankamp[1]**

1 Department of Neurosurgery, Leiden University Medical Centre, Albinusdreef 2, 2300 RC Leiden, The Netherlands

2 Department of Neurosurgery, Haga Teaching Hospital and Haaglanden Medical Center, The Hague, The Netherlands