Decompression alone versus decompression and instrumented fusion for the treatment of isthmic spondylolisthesis: a randomized controlled trial

Azizpour, K.; Schutte, P.; Arts, M.P.; Pondaag, W.; Bouma, G.J.; Coppes, M.; ... ; Vleggeert-Lankamp, C.L.A.

Decompression alone versus decompression and instrumented fusion for the treatment of isthmic spondylolisthesis: a randomized controlled trial

Kayoumars Azizpour, MSc,[1] Pieter Schutte, MD,[1,6] Mark P. Arts, MD, PhD,[3]
Willem Pondaag, MD, PhD,[1,6] Gerrit J. Bouma, MD, PhD,[4] Maarten Coppes, MD, PhD,[5]
Erik van Zwet, PhD,[2] Wilco C. Peul, MD, PhD,[1,3,6] and
Carmen L. A. Vleggeert-Lankamp, MD, MSc, PhD[1,3,7]

Departments of [1]Neurosurgery and [2]Biostatistics, Leiden University Medical Center, Leiden; [3]Department of Neurosurgery, Haaglanden Medical Center, The Hague; [4]Department of Neurosurgery, OLVG, Amsterdam; [5]University Medical Center Groningen; [6]Alrijne Hospital, Leiden and Leiderdorp, and [7]Spaarne Gasthuis, Haarlem/Hoofddorp, The Netherlands

OBJECTIVE The most advocated surgical technique to treat symptoms of isthmic spondylolisthesis is decompression with instrumented fusion. A less-invasive classical approach has also been reported, which consists of decompression only. In this study the authors compared the clinical outcomes of decompression only with those of decompression with instrumented fusion in patients with isthmic spondylolisthesis.

METHODS Eighty-four patients with lumbar radiculopathy or neurogenic claudication secondary to low-grade isthmic spondylolisthesis were randomly assigned to decompression only (n = 43) or decompression with instrumented fusion (n = 41). Primary outcome parameters were scores on the Roland Disability Questionnaire (RDQ), separate visual analog scales (VASs) for back pain and leg pain, and patient report of perceived recovery at 12-week and 2-year follow-ups. The proportion of reoperations was scored as a secondary outcome measure. Repeated measures ANOVA according to the intention-to-treat principle was performed.

RESULTS Decompression alone did not show superiority in terms of disability scores at 12-week follow-up (p = 0.32, 95% CI −4.02 to 1.34), nor in any other outcome measure. At 2-year follow-up, RDQ disability scores improved more in the fusion group (10.3, 95% CI 3.9–8.2, vs 6.0, 95% CI 8.2–12.4; p = 0.006, 95% CI −7.3 to −1.3). Likewise, back pain decreased more in the fusion group (difference: −18.3 mm, CI −32.1 to −4.4, p = 0.01) on a 100-mm VAS scale, and a higher proportion of patients perceived recovery as showing "good results" (44% vs 74%, p = 0.01). Cumulative probabilities for reoperation were 47% in the decompression and 13% in the fusion group (p < 0.001) at the 2-year follow-up.

CONCLUSIONS In patients with isthmic spondylolisthesis, decompression with instrumented fusion resulted in comparable short-term results, significantly better long-term outcomes, and fewer reoperations than decompression alone. Decompression with instrumented fusion is a superior surgical technique that should in general be offered as a first treatment option for isthmic spondylolisthesis, but not for degenerative spondylolisthesis, which has a different etiology.

Clinical trial registration number: NTR1300 (Netherlands Trial Register)

https://thejns.org/doi/abs/10.3171/2021.1.SPINE201958

KEYWORDS decompression alone; fusion; isthmic; sciatica; spondylodesis; spondylolisthesis; degenerative; lumbar

ISTHMIC spondylolisthesis is the slip of a vertebral body relative to the adjacent segment below and is caused by a discontinuity (lysis) of the pars interarticularis.[1] This anatomical malalignment should be differentiated from degenerative spondylolisthesis, which also leads to a slippage of vertebrae but has a different etiology. Surgical treatment of symptomatic isthmic spondylolisthesis consists of decompression of the nerve root,[2] as described by Gill et al.[3] Nowadays, most authors perform an additional instrumented fusion, with transpedicular screw fixation and interbody fusion being the most advocated surgical techniques;[4–9] however, Gill and White advocated that additional fusion is not mandatory.[10]

Previous studies evaluating noninstrumented decompression have demonstrated satisfactory results.[11–13] Although most spine surgeons perform decompression with

instrumented fusion in isthmic spondylolisthesis, scientific justification for instrumented fusion over decompression alone is lacking. The main argument in favor of instrumented fusion is that upon reduction of the spondylolisthesis, the neuroforamen enlarges, which adds to nerve root decompression. Furthermore, the procedure has been reported to prevent progressive slippage.[14] Drawbacks of instrumented fusion include longer surgical time, higher risk of nerve root damage, more tissue damage, and additional costs.[15,16] The arguments to omit fusion are shorter surgical time and less paraspinal muscle injury, which may lead to less surgery-related low-back pain, faster mobilization, and earlier resumption of work.

The comparison of decompression alone with decompression and fusion in spondylolisthesis is not new: two randomized trials comparing these methods in patients with degenerative spondylolisthesis were recently published.[17,18] Both studies demonstrated similar functional outcomes in the study groups at 2-year follow-up. These studies weakened the position of fusion in the standard surgical care of degenerative spondylolisthesis.

This controversy is also present in the treatment of low-grade isthmic spondylolisthesis, which justified a randomized controlled trial (RCT) comparing decompression with instrumented fusion to decompression alone. The purpose of this study was to assess clinical outcomes of both surgical strategies in the short (12 weeks) and long term (2 years).

We hypothesized that decompression alone gives superior short-term results compared to decompression and fusion, due to the less-invasive nature of the surgical procedure, but that this less-invasive procedure results in comparable outcomes in the long term.

Methods

A multicenter, prospective randomized trial was conducted to compare surgical treatment of symptomatic isthmic spondylolisthesis with decompression of the nerve root as described by Gill et al.[3] with instrumented decompression and fusion (Sciatica-Gill trial) among patients with lumbosacral radiculopathy or neurogenic claudication. The protocol was approved in all participating centers by the Medical Ethical Committee of Leiden University Medical Center. Written informed consent was obtained from all patients. Details of the design and study protocol have been published previously.[19] This trial was registered with the Netherlands Trial Register (identifier: NTR1300).

Eligibility and Randomization

Eligible patients were 18 to 70 years of age with radiologically proven low-grade isthmic spondylolisthesis (Meyerding grade I or II) and had received a diagnosis from a neurosurgeon of lumbar radicular syndrome or neurogenic claudication that had lasted for more than 3 months, in agreement with radiological evaluation. At the time of enrollment, an independent research nurse verified the persistence of the symptoms and evaluated exclusion criteria. Patients were excluded from eligibility if one of the following criteria were applicable: patients with high-grade isthmic spondylolisthesis (Meyerding

grade III or IV), disc herniation on the affected segment which required discectomy, low-back pain only, abnormal instability on dynamic radiograph (> 3 mm), and progressive spondylolisthesis. Other exclusion criteria were severe obesity (BMI > 35), previous spine surgery, severe osteoporosis, chronic use of steroids, severe comorbidity, contraindication for surgery, short-term planned migration, no or limited understanding of the Dutch language, and pregnancy.

Between June 2008 and January 2015, patients were enrolled from each of the following study centers: Spaarne Hospital and Leiden University Medical Center (n = 16), Alkmaar Medical Center (n = 3), Groene Hart Hospital and Vlietland Hospital and Haaglanden Medical Center (n = 34), Alrijne Hospital Leiden and Leiderdorp (n = 18), Sint Lucas Andreas Hospital (n = 2), and University Medical Center Groningen (n = 11).

A computer-generated permuted-block scheme was used for randomization (1:1), with patients stratified according to center. Allocations were stored in opaque, coded, and sealed envelopes prepared by a data manager not involved in the selection and allocation of patients. The key was only accessible to the ProMISe (Project Manager Internet Server) data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. After randomization, the prepared envelope was opened and the patient was informed of the allocated intervention, and subsequently the logistics for the allocated intervention were arranged. The patients could not be blinded to treatment group.

Data from 10 participating medical centers in the Netherlands were collected and entered in the ProMISe database through case report forms.

Interventions

Patients were randomized to undergo nerve root decompression by removal of the floating lamina and the fibrocartilaginous mass of the pseudo-joint according to Gill (D group) or comparable nerve root decompression combined with instrumented spondylodesis (DF group).

Patients randomized to decompression only who had persisting leg pain on the first follow-up visit at 12 weeks were offered secondary spondylodesis. These patients were analyzed in the Gill (D) group and their care conformed to the intention-to-treat principle.

In the decompression-alone group (D group), a lumbosacral midline incision was made and the paravertebral muscles were dissected unilaterally or bilaterally, depending on the patient's symptoms. The floating lamina (due to the fracture at the isthmus) and inferior articular process were removed, together with the fibrocartilaginous mass of the pseudo-joint. In case of a unilateral procedure (hemi-Gill), a vertical hemilaminectomy was performed. The affected nerve root(s) were decompressed adequately, implicating an additional reduction of the superior articular process if necessary. The wound was closed in layers with a suction drain if necessary. Decompression with instrumented spondylodesis (DF group) was performed bilaterally as described above. Subsequently, pedicle screws were placed in the affected segment under fluoroscopic control. O-arm navigation was not routinely used during the inclu-

sion period. If the intervertebral space was admissible, the disc tissue was removed, and the intervertebral space was filled with two polyetheretherketone (PEEK) or titanium cages. Sagittal realignment was achieved if possible and screws were fixed to the rods under slight compression. The wound was closed in layers with a suction drain if necessary.

Postoperatively, all patients were encouraged to mobilize as soon as possible. No orthoses were prescribed.

Outcomes

The primary outcomes were the 23-item Roland Disability Questionnaire for Sciatica (RDQ),[20] the 100-mm visual analog scales (VASs) for leg pain and back pain,[21,22] and a 7-point Likert self-rating scale of global perceived recovery.[23,24] Secondary outcomes were scores on the SF-36 scale[25] and the reoperation frequency.

All outcome scores were assessed at baseline and at 3, 6, 12, 26, 52, and 104 weeks. The Likert perceived recovery score was not evaluated at baseline. Questionnaires were sent to patients and in-person interviews were performed by research nurses and surgeons in an outpatient setting. Patients were not informed of the results of earlier assessments.

Complications

Both perioperative and postoperative complications were recorded during the study period. Complications were measured and registered in a database using a standardized classification involving type of complication and implications for (adjustment of) care according to usual care in Dutch hospitals. Perioperative complications were defined as complications occurring during surgery and were registered immediately following surgery. Postoperative complications were defined as complications occurring during inpatient care and the postoperative follow-up period, and were registered daily during the clinical period and at 2 and 6 months postoperatively in an outpatient care setting.

Statistical Analysis

The sample size was calculated based on the RDQ, considering superiority of decompression alone at short-term and noninferiority of decompression alone at long-term follow-up. At the 12-week follow-up treatment was considered successful in cases of improvement of at least 7 points on the RDQ score; furthermore, a clinically relevant difference between the groups was deemed present if the RDQ score values at the 12-week follow-up differed by at least 20%. The noninferiority margin at the 2-year follow-up was defined as a maximum 4-point difference in RDQ score between decompression alone and instrumented fusion, and superior clinically relevant outcome is deemed to be present if the difference in RDQ score was more than 4 points. With a power of 90% and a two-tailed significance level of 0.05, 220 patients with symptomatic isthmic spondylolisthesis grades I and II were calculated to be needed (110 patients in both treatment groups, including 10% loss to follow-up) for both the 12-week and the 2-year outcome data. The numbers used for this calcu-

lation were retrieved from the 1- and 5-year results of the Maine Lumbar Spine Study.[26,27]

We did not include the intended number of 220 patients. The inclusion rate was slow due to absence of willingness of the majority of patients to be randomized for surgery. Moreover, in some of the included hospitals the trial was not discussed with every patient routinely. Extrapolation of the inclusion rate led to another 10 years of inclusion time for a total of 220 patients. We decided that outcome data were more relevant if they would be available within due time. More importantly, surgeons noticed that patients randomized to the D group frequently underwent reoperation for additional fusion. Although this frequent reoperation in the D group was not foreseen in the protocol, an interim analysis for reoperation was performed and the very high reoperation rate in the D group, as opposed to the DF group, convinced the authors to terminate the study.

Groups were compared based on an intention-to-treat analysis. We tested the short-term (12 weeks) and long-term (2 years) effects of treatment. The primary and secondary outcome measures were calculated using a linear mixed model containing the interaction of randomization group and follow-up time. Heterogeneous compound symmetry was used as the covariance structure. Improvement of the outcome compared to baseline was used in the linear mixed model. Multiple imputation for missing outcome measurements was not performed, because the robust nature of mixed modeling for missing data provided valid inferences.[28] Times to perceived recovery and risk of reoperation over time were estimated for both groups using Kaplan-Meier curves and compared with log-rank tests; $p < 0.05$ was considered significant. All statistical analyses were performed using IBM SPSS version 22 (IBM Corp.).

Results

A total of 226 consecutive patients with symptomatic isthmic spondylolisthesis were eligible for inclusion. Thirty-nine patients were excluded during the second stage and 84 patients were enrolled in the Sciatica-Gill trial (Fig. 1). Of the 103 patients who declined to enroll in the RCT, 101 patients agreed to be followed in a separate prospective nonrandomized cohort study.

The included 84 patients were randomly assigned to decompression (D: 43 patients) or decompression and instrumented fusion (DF: 41 patients). Baseline variables did not show significant differences between the two groups (Table 1). A vast majority of patients had Meyerding grade I spondylolisthesis in the D and DF groups (37 and 33 patients, respectively). Likewise, the numbers of patients with Meyerding grade II did not significantly differ between the D and DF groups (6 patients and 8 patients, respectively; p = 0.57). Dynamic flexion-extension radiographs were used to evaluate movement at the level of spondylolisthesis. Thirty-five patients in the D group had movement of 0–1 mm, compared to 32 patients in the DF group. Movement of 2–3 mm was present in 6 patients in the D group compared to 7 patients in the DF group (p = 0.77). Due to spontaneous clinical improvement after randomization, 3 patients did not undergo surgery. One patient in the D group underwent an additional spondy-
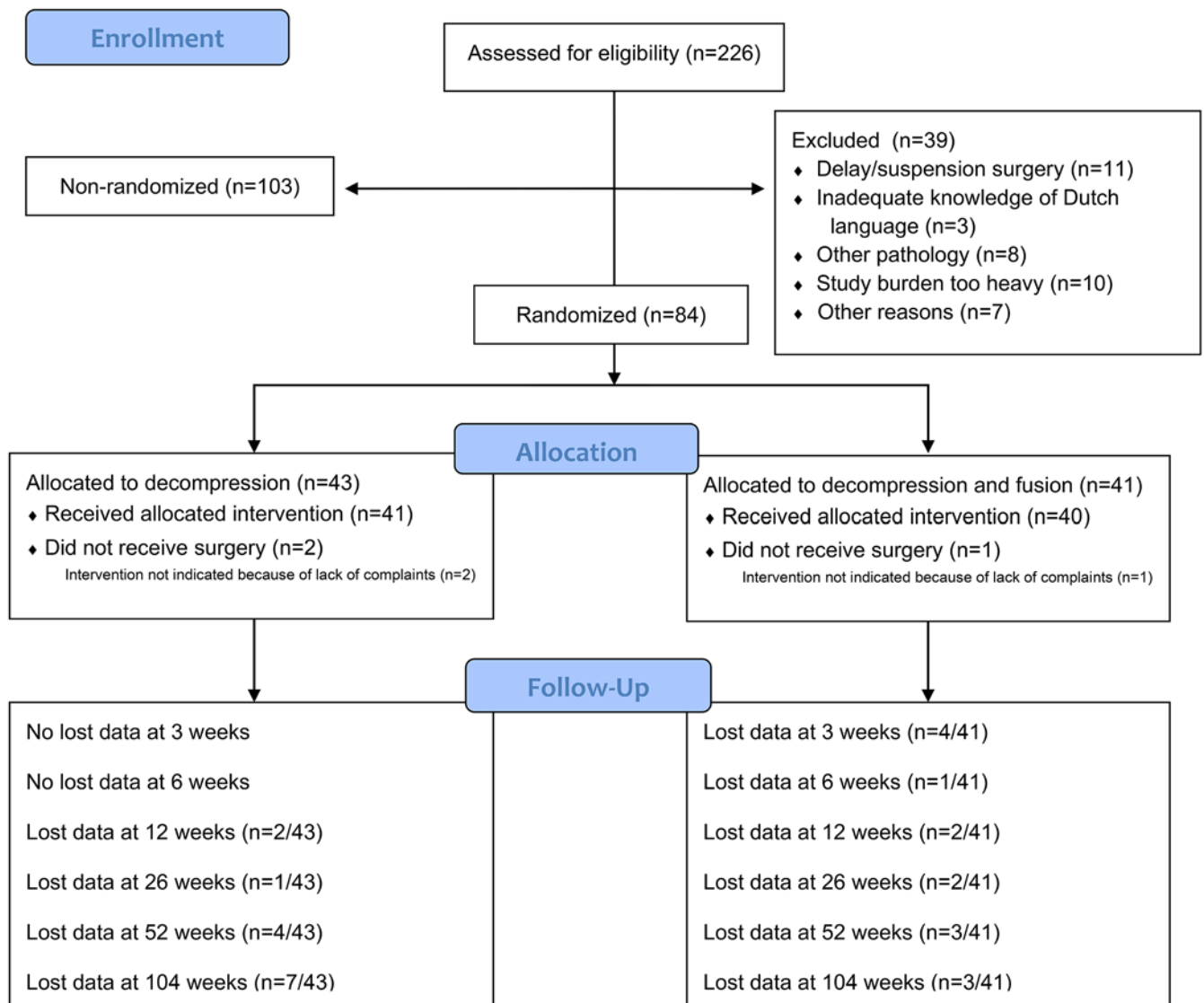
FIG. 1. Overview of patient enrollment, randomization, and follow-up. All patients were analyzed according to the group to which the patient had been randomized. Nonrandomized patients (n = 103) refused randomization and were followed in an additional cohort study. "Other reasons" (n = 7) for exclusion were as follows: patient could not be reached (n = 1), patient was not enrolled in the study protocol in time (n = 5), and patient was older than 70 years (n = 1). Figure is available in color online only.

lodesis because the patient insisted on additional instrumentation prior to the intervention. One patient in the DF group underwent decompression only. The reason could not be retrieved. Follow-up data at 12 weeks and at 2 years were available for 80 patients (95% compliance rate) and 74 patients (88% compliance rate), respectively (Fig. 1). Data for surgical time, blood loss, hospital stay, and reoperations are summarized in Tables 2 and 3.

During the first 12 weeks after surgery, RDQ scores for disability declined comparably in the two groups (Fig. 2A). The decrease of 7 points in RDQ was nearly achieved in the D group, but successfully realized in the DF group. Although the difference in decline was more than 20%, there was no significant difference in decline over the course of the first 12 weeks after surgery. Likewise, leg and back pain in the D and DF groups demonstrated a

comparable decline during this period (Fig. 2B and C). The recovery of overall health was perceived by a comparable number of patients in both groups, but the perceived recovery of leg pain was present in a higher percentage of patients in the DF group (p = 0.01; Table 4). The SF-36 subscale and summary scores were comparable between both groups at the 12-week follow-up (Table 4).

Two years after randomization, however, observed disability in the D group remained at the same level as at the 12-week follow-up, while the disability in the DF group continued to decline (p = 0.04). The difference exceeded the previously indicated clinically relevant 4-point border (Fig. 2A). Similar to the 12-week follow-up, the SF-36 subscale and summary scores did not show a statistically significant difference between both groups at the 2-year follow-up (Table 5).

TABLE 1. Baseline characteristics for all patients

Characteristic	Study Group		p Value
	D (n = 43)	DF (n = 41)	
Age, yrs	50.5 ± 10.4	53.5 ± 9.9	0.19
Age ≥60 yrs	10	12	0.62
Males	24	27	0.38
BMI, kg/m^2	27.8 ± 4.6	27.0 ± 3.6	0.36
BMI ≥30	12	8	0.45
Smoking status, often	17	17	
Daily cigarettes*	15.0 (6.0–20.0)	20.0 (10.0–22.5)	0.07
Duration of symptoms, mos			
Leg pain	12.0 (5.0–24.0)	12.0 (10.5–24.0)	0.15
Low-back pain	16.0 (6.0–60.0)	24.0 (12.0–150.0)	0.15
Nature of symptoms†			0.25
Neurogenic claudication	18	14	
Radicular pain	19	13	
Both	6	11	
Motor deficit‡	19	13	0.27
SF-36 score			
Physical functioning	44.9 ± 20.1	39.6 ± 17.7	0.21
Role: physical	26.8 ± 37.2	20.1 ± 36.3	0.41
Role: emotional	63.5 ± 44.1	46.3 ± 46.5	0.09
Vitality	47.4 ± 23.2	51.7 ± 17.9	0.35
Mental health	66.6 ± 21.5	70.0 ± 20.3	0.46
Social functioning	59.3 ± 26.5	55.8 ± 26.1	0.54
Bodily pain	36.5 ± 21.6	33.3 ± 19.6	0.49
General health	56.2 ± 22.7	56.6 ± 21.0	0.93
Physical Component Summary score	33.3 ± 8.6	32.0 ± 7.8	0.47
Mental Component Summary score	46.4 ± 13.3	46.0 ± 13.0	0.89
Preop RDQ	14.3 ± 4.2	14.6 ± 4.3	0.69
Preop VAS leg	66.4 ± 24.5	68.5 ± 21.9	0.68
Preop VAS back	56.3 ± 25.6	60.3 ± 27.4	0.49
Operated level			0.44§
L3–4	1	1	
L4–5	9	5	
L4–S1	0	1	
L5–S1	29	32	
L5–6	2	1	
Spondylolisthesis grade			0.57
Meyerding I	37	33	
Meyerding II	6	8	
Flexion-extension movement at listhesis¶			0.77
0–1 mm	35	32	
2–3 mm	6	7	

D = decompression-alone; DF = decompression and instrumented fusion.

Values are presented as number of patients, mean ± SD, or median (IQR) unless otherwise indicated.

* Data were available for 17 patients in the D group and 17 patients in the DF group.

† Data were available for 43 patients in the D group and 38 patients in the DF group.

‡ Motor deficit defined as measured weakness in motor function detected on neurological exam. Data were available for 43 patients in the D group and 41 patients in the DF group.

§ Fisher's exact test based on dichotomized values (L5–S1 vs rest). Data were available for 41 patients in the D group and 40 patients in the DF group.

¶ Data measured from preoperative dynamic flexion-extension radiographs. Data were available for 41 patients in the D group and 39 patients in the DF group.

TABLE 2. Surgical data and preoperative and postoperative complications

Characteristic	Study Group		p Value
	D (n = 43)	DF (n = 41)	
Surgical time, mins	45.0 (30.0–74.5)	180.0 (120.0–210.0)	<0.001
Blood loss, ml	100.0 (50.0–150.0)	420.0 (300.0–800.0)	<0.001
Length of stay, days	2 (1–2)	4 (3–5)	<0.001
Periop complications, no.	3	6	
Dural tear	3	3	0.31
Pedicle breach	0	3	
Postop complications, no.	4	8	
Postop hemorrhage†	0	1	
CSF leak‡	1	3	
Micturition disorder§	0	2	0.19*
Sensory dysfunction¶	2	2	
Motor dysfunction	1	0	

Values are presented as number of patients or median (IQR) unless otherwise indicated. Postoperative complications were defined during hospitalization by a research nurse. There was a statistical difference in blood loss, surgical time, and length of stay between the two groups. There were no differences in terms of perioperative and postoperative complications. For variables with missing data, frequencies are based on available cases.
* Fisher's exact test based on 3 patients in the D group versus 7 patients in the DF group.
† Postoperative hematoma needing reintervention.
‡ CSF leakage requiring prolonged bed rest.
§ Temporary bladder dysfunction requiring urinary catheter.
¶ Numbness in the dermatome of the decompressed nerve root.

Leg pain decreased comparably in both groups, but back pain improved significantly more in the DF group (p = 0.04; Fig. 2B and C). Perceived recovery was present in a higher percentage of patients in the DF group, both global recovery (p = 0.01) and perceived recovery (p = 0.04) of leg pain. The survival analysis showed a difference in speed to perceived recovery of global health in favor of the DF group (p = 0.012; Fig. 2D).

Complications and Reoperations

In total, dural tears occurred in 3 patients in both groups. In addition, pedicle breaches occurred in 3 patients in the DF group. There was no difference between the two groups in perioperative complications (p = 0.55; Table 2). Postoperative complications in the D group consisted of 1 case of cerebrospinal fluid (CSF) leak, 2 cases of sensory dysfunction, and 1 case of motor dysfunction.

TABLE 3. Reoperations during the study period

Cause of Reop	D (n = 19)	DF (n = 5)
Persisting radicular leg pain	17	1
Adjacent segment disease	0	2
Hardware failure*	1†	1
Invalidating back pain	0	1
Herniated nucleus pulposus	1	0

Data are presented as number of patients.
* Hardware failure comprised pedicle screw loosening.
† Patient was randomized to decompression alone, but received decompression and instrumented fusion.

Postoperative complications in the DF group consisted of 1 case of postoperative hemorrhage, defined as a postoperative hematoma requiring reintervention; 3 cases of CSF leaks; 2 cases of bladder dysfunction, defined as dysfunction requiring a short-term urinary catheter; and 2 cases of sensory dysfunction. There was no difference between the two groups in postoperative complications (p = 0.19).

In the decompression group, 19 patients underwent a reoperation after a median time of 7.3 months (range 2.3–13.0 months; Fig. 2E). One patient, who was randomized to the D group but underwent decompression and fusion instead, received revision surgery due to hardware failure. The remaining 18 patients failed to recover from radicular leg pain. Seventeen of these patients underwent a subsequent instrumented fusion and the remaining patient underwent a secondary decompression at a different level (Table 3).

In the decompression and instrumented fusion group, 5 patients were reoperated after a median time of 5.0 months (range 3.3–20.4 months). One patient inadvertently underwent primary decompression. Due to disabling back pain, the patient subsequently received instrumented fusion. Three patients underwent extension of fusion to the adjacent level. In 1 case the fusion material failed and a subsequent revision was performed (Table 3).

Discussion

In this study we demonstrated that instrumented fusion in addition to decompression of the nerve roots is superior to decompression alone in patients with symptomatic isthmic spondylolisthesis. Early outcomes at the 12-week
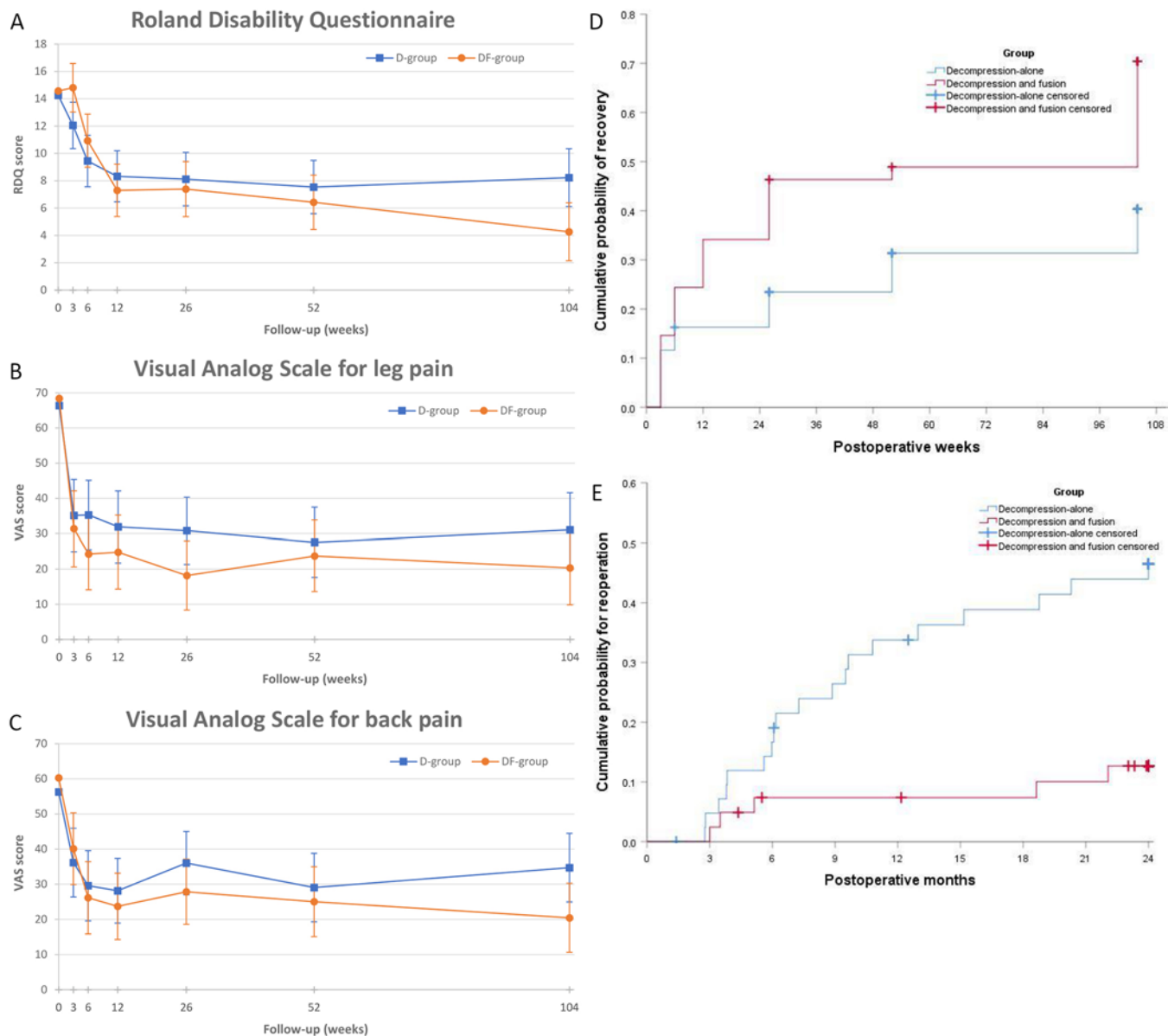
FIG. 2. Courses of primary outcomes, Kaplan-Meier curves for risk of reoperation, and probability of perceived recovery in the decompression (D) and decompression and fusion (DF) groups. Shown are the courses of the RDQ score (**A**), VAS leg pain score (**B**), and VAS back pain score (**C**) over time, in which outcome at 0 weeks was set as the mean value at baseline. The cumulative probability of perceived recovery of global health over time is shown (**D**) for the two study groups. Patients lost to follow-up or without perceived recovery were censored. Mean time to recovery was 78.0 weeks (95% CI 65.47–90.49) in the D group and 59.8 weeks (95% CI 45.49–74.19) in the DF group. There was a significant difference in survival distributions between the two groups (p = 0.012 by the log-rank test). The cumulative risk of reoperation over time for the two study groups is also shown (**E**). Patients lost to follow-up or patients not reoperated within 2 years after the primary operation were censored. The cumulative probability for reoperation at 2-year follow-up was 47% in the D group and 13% in the DF group (p < 0.001). Figure is available in color online only.

follow-up demonstrated that fusion tended to be preferable to decompression alone, in contrast to the hypothesis that early recovery would be hampered by the more extensive surgery of fusion. At long-term follow-up, outcomes in the fusion group were superior to those in the decompression group with respect to functional disability, back pain, and perceived recovery.

Although almost half of the patients in the decompression alone (D) group underwent reoperation and received

additional instrumented spondylodesis, the decompression and fusion (DF) group still showed superior results. Despite D-group patients eventually receiving the same treatment as patients in the DF group, the decompression-alone group as a whole showed inferior results. An explanation for this finding may be that the median time to reoperation was 7.3 months, and therefore reoperated patients in the D group were still not fully recovered at the 2-year follow-up, resulting in an unequal comparison.

TABLE 4. Summary of study outcomes at 12-week follow-up

Variable	D	Difference vs Baseline (95% CI)	DF	Difference vs Baseline (95% CI)	Improvement in D vs DF (95% CI)	p Value*
Primary outcome						
RDQ score	8.1 ± 6.7	5.9 (4.1 to 7.8)	7.7 ± 6.6	7.3 (5.4 to 9.2)	−1.3 (−4.0 to 1.3)	0.32
VAS leg pain	30.5 ± 30.1	34.5 (24.3 to 44.7)	25.2 ± 30.8	43.6 (33.2 to 54.1)	−9.1 (−23.7 to 54.1)	0.22
VAS back pain	29.2 ± 28.3	28.1 (18.9 to 37.4)	24.7 (22.2)	36.6 (27.2 to 46.0)	−8.4 (−22.6 to 4.7)	0.21
Perceived recovery: global health†‡	16 (39)	NA	23 (59)	NA	NA	0.07
Perceived recovery: leg pain†‡	17 (42)	NA	27 (69)	NA	NA	0.01
Secondary outcomes						
SF-36						
Physical functioning	67.7 (25.2)	22.6 (15.4 to 29.7) (p < 0.001)	63.7 (22.0)	24.4 (16.8 to 32.0) (p < 0.001)	4.0 (−6.6 to 14.5)	0.46
Role: physical	32.3 (44.1)	2.5 (−10.3 to 15.3) (p = 0.69)	27.6 (35.3)	7.1 (−6.6 to 20.7) (p = 0.30)	4.8 (−13.0 to 22.5)	0.60
Role: emotional	61.0 (48.2)	−2.5 (−15.8 to 10.8) (p = 0.71)	70.9 (39.9)	25.6 (10.1 to 41.2) (p = 0.002)	−10.0 (−29.6 to 9.7)	0.32
Vitality	59.6 (22.7)	11.5 (5.9 to 17.1) (p < 0.001)	66.4 (18.1)	15.6 (10.1 to 21.1) (p < 0.001)	−6.8 (−15.9 to 2.4)	0.15
Mental health	75.1 (20.1)	8.1 (3.0 to 13.2) (p = 0.003)	81.0 (16.6)	12.3 (7.2 to 17.4) (p < 0.001)	−5.9 (−13.9 to 2.1)	0.15
Social functioning	77.7 (25.5)	16.2 (9.3 to 23.1) (p < 0.001)	75.6 (19.2)	20.5 (10.9 to 30.2) (p < 0.001)	2.1 (−7.9 to 12.1)	0.68
Bodily pain	58.2 (22.9)	19.7 (11.8 to 27.6) (p < 0.001)	57.4 (17.0)	24.6 (18.2 to 31.1) (p < 0.001)	0.9 (−8.1 to 9.9)	0.85
General health	62.7 (23.4)	7.7 (3.1 to 12.3) (p = 0.002)	70.0 (17.7)	13.5 (7.5 to 19.4) (p < 0.001)	−7.3 (−16.5 to 1.9)	0.12
Physical Component Summary score	41.0 ± 10.2	7.3 (4.4 to 10.2) (p < 0.001)	39.3 (8.1)	7.3 (4.4 to 10.2) (p < 0.001)	1.7 (−2.5 to 5.8)	0.43
Mental Component Summary score	48.7 ± 13.0	1.9 (−1.1 to 5.0) (p = 0.21)	52.7 (9.6)	7.4 (3.7 to 11.0) (p < 0.001)	−4.1 (−9.2 to 1.0)	0.11

NA = not applicable.

Data are expressed as number (%) of patients or mean ± SD unless otherwise indicated.

* p value of difference in improvement between D and DF at 12-week time point.

† Likert scale: "complete recovery" and "almost complete recovery" were defined as a good result, which was used to dichotomize data. A Likert perceived recovery evaluation was performed for recovery of overall health and recovery of leg pain separately.

‡ Summary data are based on available cases.

In the D group, half of the patients did not require an additional instrumented spondylodesis. It may be argued that a subgroup of patients could achieve satisfactory results with decompression alone, and that decompression alone is suitable for a select group of patients who are unfit to undergo spondylodesis surgery, such as fragile patients, patients with osteoporosis, and patients with a high chance of implant failure. However, the current study did not have enough power to detect this finding, and future studies would be necessary to address this issue.

Before the start of the trial, a decrease of 7 or more points on the RDQ was defined as a successful result. Only the DF group reached this absolute decrease in RDQ at the 12-week follow-up. Although there was no significant difference in leg and back pain at the 12-week follow-up, significantly more patients in the fusion group perceived their recovery of leg pain as good. The comparable short-term outcomes in the fusion group may be explained by the fact that instrumented spondylodesis is a surgical procedure that is commonly performed by surgeons in our study group, increasing the overall technical proficiency of the treatment.

At the 2-year follow-up, a significant and clinically rele-vant difference in effectiveness was demonstrated, since a more than 4-point difference on the RDQ score was measured, favoring decompression and fusion. Theoretically, patients who did not respond well to surgery, who were potentially present in both randomization groups, could leave the D group and enter the DF group, but not vice versa. This situation could have theoretically negatively influenced the results in the DF group. However, the results in the DF group are even better, emphasizing the superiority of decompression with fusion. This difference is most likely attributable to stabilization of the involved segment, preventing progressive slippage and further nerve root compression. Restoration of the sagittal lumbosacral alignment also may have played an important role.[11,29]

Even though the inclusion of 220 patients was not reached and the study was underpowered, a statistically significant difference in the decrease of functional disability was found at the 2-year follow-up in favor of decompression and fusion. This conclusion is even strengthened by the observation that half of the patients in the D group were reoperated to undergo additional fusion. However, the conclusion that the results did not differ between the two groups at the 12-week follow-up is not solid, since

TABLE 5. Summary of study outcomes at 2-year follow-up

Variable	D	Difference vs Baseline (95% CI)	DF	Difference vs Baseline (95% CI)	Difference in Improvement D vs DF (95% CI)	p Value*
Primary outcome						
RDQ score	7.4 (6.3)	6.0 (3.9 to 8.2)	4.8 (6.5)	10.3 (8.2 to 12.4)	−4.3 (−7.3 to −1.3)	0.006
VAS leg pain	26.6 (27.8)	35.3 (24.8 to 45.9)	21.4 (26.5)	48.1 (37.5 to 58.6)	−12.7 (−27.6 to 2.2)	0.09
VAS back pain	32.9 (27.4)	21.6 (11.8 to 31.3)	21.1 (22.3)	39.8 (30.0 to 49.7)	−18.3 (−32.1 to −4.4)	0.01
Perceived recovery in global health†‡	16 (44)	NA	28 (74)	NA	NA	0.01
Perceived recovery in leg pain†‡	18 (50)	NA	28 (74)	NA	NA	0.04
Secondary outcome						
SF-36						
Physical functioning	68.9 (25.2)	23.5 (14.5 to 32.4) (p < 0.001)	76.3 (25.3)	37.6 (30.3 to 45.0) (p < 0.001)	−7.4 (−19.1 to 4.3)	0.21
Role: physical	56.3 (43.7)	30.0 (15.6 to 44.4) (p < 0.001)	69.1 (43.7)	49.3 (33.8 to 64.9) (p < 0.001)	−12.8 (−33.1 to 7.4)	0.21
Role: emotional	77.8 (37.4)	13.3 (−3.1 to 29.8) (p = 0.11)	78.9 (39.8)	31.6 (13.9 to 49.2) (p = 0.001)	−1.2 (−19.1 to 16.8)	0.90
Vitality	62.6 (19.1)	13.4 (6.6 to 20.3) (p < 0.001)	69.3 (18.9)	18.6 (12.4 to 24.7) (p < 0.001)	−6.7 (−15.5 to 2.1)	0.13
Mental health	75.1 (15.4)	5.8 (0.4 to 11.3) (p = 0.04)	80.0 (17.1)	9.7 (5.4 to 14.0) (p < 0.001)	−4.9 (−12.5 to 2.6)	0.20
Social functioning	79.2 (27.2)	16.0 (6.9 to 25.0) (p = 0.001)	87.5 (19.3)	32.6 (23.0 to 42.1) (p < 0.001)	−8.3 (−19.2 to 2.6)	0.13
Bodily pain	61.6 (24.6)	22.7 (14.0 to 31.5) (p < 0.001)	72.0 (22.1)	39.6 (32.7 to 46.5) (p < 0.001)	−10.4 (−21.2 to 0.4)	0.06
General health	63.1 (21.8)	5.7 (0.7 to 10.7) (p = 0.03)	72.8 (21.6)	14.7 (9.0 to 20.4) (p < 0.001)	−9.7 (−19.8 to 0.4)	0.06
Physical Component Summary score	43.1 (11.3)	9.9 (6.4 to 13.4) (p < 0.001)	47.7 (10.3)	16.0 (13.2 to 18.8) (p < 0.001)	−4.6 (−9.6 to 0.4)	0.07
Mental Component Summary score	50.3 (9.9)	2.4 (−1.3 to 6.2) (p = 0.20)	52.0 (10.0)	5.8 (2.3 to 9.3) (p = 0.002)	−1.7 (−6.3 to 3.0)	0.48

Data are expressed as number (%) of patients unless otherwise indicated.

* p value of difference in improvement between D and DF at 2-year time point.

† Likert scale: "complete recovery" and "almost complete recovery" were defined as a good result, which was used to dichotomize data. A Likert perceived recovery evaluation was performed for recovery of overall health and recovery of leg pain separately.

‡ Summary data are based on available cases.

the sample size was not reached. For the same reason, the conclusion that the D-group patients had an unsatisfactory result after 12 weeks is also not valid.

With respect to back pain, the difference between the two groups is, although significant, so small that the clinical relevance can be debated. The power of the study was not calculated based on back pain, and with the current study population the better VAS back values in the fusion group may be underestimated. Reduction of the affected spinal segments may contribute to a decline in back pain, and therefore the results of this trial may lead to somewhat more optimism regarding recovery from back pain associated with decompression and fusion in spondylolisthesis surgery.

A drawback of instrumented spondylodesis could be an additional risk of nerve root and tissue damage.[15] In this study, we did not demonstrate any differences in perioperative and postoperative complications. Also, instrumented spondylodesis is a more extensive procedure, which could result in more blood loss and longer surgical time than de-

compression only. Even though our findings confirm this assumption, the fact that computer-assisted navigation is increasingly being used intraoperatively by present-day surgeons may result in shorter surgical times and thus the possibility that these potential operative disadvantages were overestimated. Another drawback for instrumented spondylodesis could be a longer required hospital stay. However, this difference can be explained by the availability of additional standard physical therapy programs, which required a similar hospital stay of two nights postoperatively for both groups.

Our study has several important limitations. Most importantly, the power of the study has not been met, thus reducing the chance of detecting a predefined clinical difference. This also applies to the complication rates found, given that the study was not powered to detect these differences. The results of the primary outcome, however, demonstrated a statistically and clinically relevant difference between the two groups at the 2-year follow-up. Not only is the difference in reoperation rates convincing, and the

main reason for premature termination of the study, but the RDQ score outcomes are convincing as is. However, this is a limitation that was not defined in the stopping rules in our protocol. Another limitation is a higher loss to follow-up in the decompression-alone group (n = 7) compared to the decompression and fusion group (n = 3) at the 2-year follow-up. This could yield attrition bias by potentially losing those patients that scored in a particular direction. However, mixed model analyses reduce this bias by using the correlation within and between cases to impute the missing values, provided that there are enough follow-up data.

Conclusions

Decompression with fusion was superior to decompression alone in low-grade isthmic spondylolisthesis during long-term follow-up and did not show the expected slow recovery and complications during the first 3 months. In contrast to patients with degenerative spondylolisthesis, patients with low-grade isthmic spondylolisthesis should be treated with decompression with instrumented fusion as the first choice.

Acknowledgments

References

1. Kreiner DS, Baisden J, Mazanec DJ, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of adult isthmic spondylolisthesis. *Spine J.* 2016;16(12):1478–1485.
2. Möller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis—a prospective randomized study: part 1. *Spine (Phila Pa 1976).* 2000;25(13):1711–1715.
3. Gill GG, Manning JG, White HL. Surgical treatment of spondylolisthesis without spine fusion; excision of the loose lamina with decompression of the nerve roots. *J Bone Joint Surg Am.* 1955;37-A(3):493–520.
4. La Rosa G, Conti A, Cacciola F, et al. Pedicle screw fixation for isthmic spondylolisthesis: does posterior lumbar interbody fusion improve outcome over posterolateral fusion? *J Neurosurg.* 2003;99(2)(suppl):143–150.
5. Knight M, Goswami A. Management of isthmic spondylolisthesis with posterolateral endoscopic foraminal decompression. *Spine (Phila Pa 1976).* 2003;28(6):573–581.
6. Carragee EJ. Single-level posterolateral arthrodesis, with or without posterior decompression, for the treatment of isthmic spondylolisthesis in adults. A prospective, randomized study. *J Bone Joint Surg Am.* 1997;79(8):1175–1180.
7. Deguchi M, Rapoff AJ, Zdeblick TA. Posterolateral fusion for isthmic spondylolisthesis in adults: analysis of fusion rate and clinical results. *J Spinal Disord.* 1998;11(6):459–464.
8. Müslüman AM, Yılmaz A, Cansever T, et al. Posterior lumbar interbody fusion versus posterolateral fusion with instrumentation in the treatment of low-grade isthmic spondylolisthesis: midterm clinical outcomes. *J Neurosurg Spine.* 2011;14(4):488–496.
9. Schnee CL, Freese A, Ansell LV. Outcome analysis for adults with spondylolisthesis treated with posterolateral fusion and transpedicular screw fixation. *J Neurosurg.* 1997;86(1):56–63.
10. Gill GG, White HL. Surgical treatment of spondylolisthesis without spine fusion: a long term follow-up of operated cases. *Acta Orthop Scand Suppl.* 1965;36(Suppl 85):5–99.
11. Arts M, Pondaag W, Peul W, Thomeer R. Nerve root decompression without fusion in spondylolytic spondylolisthesis: long-term results of Gill's procedure. *Eur Spine J.* 2006;15(10):1455–1463.
12. Osterman K, Lindholm TS, Laurent LE. Late results of removal of the loose posterior element (Gill's operation) in the treatment of lytic lumbar spondylolisthesis. *Clin Orthop Relat Res.* 1976;(117):121–128.
13. Weiner BK, McCulloch JA. Microdecompression without fusion for radiculopathy associated with lytic spondylolisthesis. *J Neurosurg.* 1996;85(4):582–585.
14. Sudo H, Oda I, Abumi K, et al. Biomechanical study on the effect of five different lumbar reconstruction techniques on adjacent-level intradiscal pressure and lamina strain. *J Neurosurg Spine.* 2006;5(2):150–155.
15. Fritzell P, Hägg O, Nordwall A. Complications in lumbar fusion surgery for chronic low back pain: comparison of three surgical techniques used in a prospective randomized study. A report from the Swedish Lumbar Spine Study Group. *Eur Spine J.* 2003;12(2):178–189.
16. Möller H, Hedlund R. Instrumented and noninstrumented posterolateral fusion in adult spondylolisthesis—a prospective randomized study: part 2. *Spine (Phila Pa 1976).* 2000;25(13):1716–1721.
17. Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. *N Engl J Med.* 2016;374(15):1424–1434.
18. Försth P, Ólafsson G, Carlsson T, et al. A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. *N Engl J Med.* 2016;374(15):1413–1423.
19. Arts MP, Verstegen MJT, Brand R, et al. Cost-effectiveness of decompression according to Gill versus instrumented spondylodesis in the treatment of sciatica due to low grade spondylolytic spondylolisthesis: a prospective randomised controlled trial [NTR1300]. *BMC Musculoskelet Disord.* 2008;9:128.
20. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976).* 1983;8(2):141–144.
21. Scott J, Huskisson EC. Graphic representation of pain. *Pain.* 1976;2(2):175–184.
22. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain.* 1983;16(1):87–101.
23. Bombardier C. *Spine* Focus Issue introduction. Outcome assessments in the evaluation of treatment of spinal disorders. *Spine (Phila Pa 1976).* 2000;25(24):3097–3099.
24. Main CJ, Wood PL, Hollis S, et al. The Distress and Risk Assessment Method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine (Phila Pa 1976).* 1992;17(1):42–52.
25. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ.* 1992;305(6846):160–164.
26. Atlas SJ, Keller RB, Wu YA, et al. Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the Maine Lumbar Spine Study. *Spine (Phila Pa 1976).* 2005;30(8):927–935.
27. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine (Phila Pa 1976).* 1996;21(15):1777–1786.
28. Biering K, Hjollund NH, Frydenberg M. Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clin Epidemiol.* 2015;7:91–106.
29. Cunningham JE, Elling EM, Milton AH, Robertson PA.

What is the optimum fusion technique for adult isthmic spondylolisthesis—PLIF or PLF? A long-term prospective cohort comparison study. *J Spinal Disord Tech*. 2013;26(5):260–267.

Disclosures

Author Contributions

Conception and design: Schutte, Arts, Peul, Vleggeert-Lankamp. Acquisition of data: Azizpour, Arts, Pondaag, Bouma, Coppes, Peul, Vleggeert-Lankamp. Analysis and interpretation of data: Azizpour, Schutte, Pondaag, van Zwet, Peul, Vleggeert-Lankamp. Drafting the article: Azizpour. Critically revising the article: Azizpour, Schutte, Arts, Peul, Vleggeert-Lankamp. Reviewed submitted version of manuscript: Schutte, Arts, Pondaag, Bouma, Coppes, van Zwet, Peul, Vleggeert-Lankamp. Approved the final version of the manuscript on behalf of all authors: Azizpour. Statistical analysis: Azizpour, Schutte, van Zwet, Peul, Vleggeert-Lankamp. Administrative/technical/material support: Azizpour, Bouma, Coppes, van Zwet. Study supervision: Pondaag, Peul, Vleggeert-Lankamp.

Supplemental Information

Previous Presentations

Presented at a plenary session of the EUROSPINE 2019 Annual Meeting, October 18, 2019, Helsinki, Finland.

Correspondence

Kayoumars Azizpour: Leiden University Medical Center, Leiden, The Netherlands. k.azizpour@lumc.nl.