The Spine Journal 23 (2023) 1613–1622

Clinical Study

Surgery in degenerative spondylolisthesis: does fusion improve outcome in subgroups? A secondary analysis from a randomized trial (NORDSTEN trial)

Christian Hellum, PhD[a,*], Frode Rekeland, MD[b,c], Milada Cvancarova Småstuen, PhD[x], Tore Solberg, PhD[i,j], Erland Hermansen, PhD[d], Kjersti Storheim, PhD[e,f], Jens Ivar Brox, PhD[g,h], Håvard Furunes, PhD[k,l], Eric Franssen, MD[m], Clemens Weber, PhD[n,o], Helena Brisby, PhD[p,q], Oliver Grundnes, PhD[r], Knut Robert Hector Algaard, MD[s], Tordis Böker, MD[t,u], Hasan Banitalebi, MD[v], Kari Indrekvam, PhD[b,w], Ivar Magne Austevoll, PhD[b]

[a] *Division of Orthopedic Surgery, Oslo University Hospital Ullevål, Fallanveien 6c, Oslo, Norway*
[b] *Kysthospitalet in Hagevik, Orthopedic Department, Haukeland University Hospital, Bergen, Norway*
[c] *Department of Clinical Medicine, University of Bergen, Bergen, Norway*
[d] *Orthopedic Department, Møre and Romsdal Hospital Trust, Ålesund Hospital, Norway*
[e] *Research and Communication Unit for Musculoskeletal Health, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway*
[f] *Department of Physiotherapy, Oslo Metropolitan University, Oslo, Norway*
[g] *Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway*
[h] *Medical Faculty, University of Oslo, Oslo, Norway*
[i] *Institute of Clinical Medicine, The Arctic University of Norway (UiT), Tromsø, Norway*
[j] *The Norwegian Registry for Spine Surgery (NORspine), University Hospital of North Norway, Tromsø, Norway*
[k] *Gjøvik Hospital, Innlandet Hospital Trust, Brumunddal, Norway*
[l] *Institute of Health and Society Studies, University of Oslo, Oslo, Norway*
[m] *Department of Orthopedic, Stavanger University Hospital, Stavanger, Norway*
[n] *Department of Neurosurgery, Stavanger University Hospital, Stavanger, Norway*
[o] *Department of Quality and Health Technology, Faculty of Health Sciences, University of Stavanger, Stavanger, Norway*
[p] *Spine Surgery Team, Department of Orthopedics, Sahlgrenska University Hospital, Gothenburg, Sweden*
[q] *Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Sweden*
[r] *Department of Orthopedic, Akershus University Hospital, Oslo, Norway*
[s] *Radiology, Unilabs Radiology, Oslo, Norway*
[t] *Division of Radiology and Nuclear Medicine, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway*
[u] *Department of Radiology and Nuclear Medicine, Oslo University Hospital Ullevål, Oslo, Norway*
[v] *Department of Diagnostic Imaging, Akershus University Hospital, Lorenskog, Norway*
[w] *Institute of Clinical Medicine, University of Oslo, Oslo, Norway*
[x] *Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway*

Received 14 February 2023; revised 16 May 2023; accepted 17 June 2023

FDA device/drug status: Not applicable.

Author disclosures: ***CH***: Nothing to disclose. ***FR***: Nothing to disclose. ***MCS:*** Nothing to disclose. ***TS***: Nothing to disclose. ***EH***: Nothing to disclose. ***KS***: Nothing to disclose. ***JIB***: Nothing to disclose. ***HF***: Nothing to disclose. ***EF***: Nothing to disclose. ***CW***: Nothing to disclose. ***HB:*** Nothing to disclose. ***OG:*** Nothing to disclose. ***KRHA:*** Nothing to disclose. ***TB***: Nothing to disclose. ***HB:*** Nothing to disclose. ***KI***: Nothing to disclose. ***IMA:*** Nothing to disclose.

*Corresponding author. Christian Hellum, PhD, Division of Orthopedic Surgery, Oslo University Hospital Ullevål, Fallanveien 6c, Oslo, 0495 Norway. Tel.: +(479)-008-5615.

E-mail address: Christianhellum1@gmail.com (C. Hellum).

https://doi.org/10.1016/j.spinee.2023.06.386

Abstract

BACKGROUND CONTEXT: Patients with spinal stenosis and degenerative spondylolisthesis are treated surgically with decompression alone or decompression with fusion. However, there is debate regarding which subgroups of patients may benefit from additional fusion.

PURPOSE: To investigate possible treatment effect modifiers and prognostic variables among patients operated for spinal stenosis and degenerative spondylolisthesis.

DESIGN: A secondary exploratory study using data from the Norwegian Degenerative Spondylolisthesis and Spinal Stenosis (NORDSTEN-DS) trial. Patients were randomized to decompression alone or decompression with instrumented fusion.

PATIENT SAMPLE: The sample in this study consists of 267 patients from a randomized multicenter trial involving 16 hospitals in Norway. Patients were enrolled from February 12, 2014, to December 18, 2017. The study did not include patients with degenerative scoliosis, severe foraminal stenosis, multilevel spondylolisthesis, or previous surgery.

OUTCOME MEASURES: The primary outcome was an improvement of ≥ 30% on the Oswestry Disability Index score (ODI) from baseline to 2-year follow-up.

METHODS: When investigating possible variables that could modify the treatment effect, we analyzed the treatment arms separately. When testing for prognostic factors we analyzed the whole cohort (both treatment groups). We used univariate and multiple regression analyses. The selection of variables was done a priori, according to the published trial protocol.

RESULTS: Of the 267 patients included in the trial (183 female [67%]; mean [SD] age, 66 [7.6] years), complete baseline data for the variables required for the present analysis were available for 205 of the 267 individuals. We did not find any clinical or radiological variables at baseline that modified the treatment effect. Thus, none of the commonly used criteria for selecting patients for fusion surgery influenced the chosen primary outcome in the two treatment arms. For the whole cohort, less comorbidity (American Society of Anesthesiologists Classification [ASA], OR = 4.35; 95% confidence interval (CI [1.16−16.67]) and more preoperative leg pain (OR = 1.23; CI [1.02 −1.50]) were significantly associated with an improved primary outcome.

CONCLUSIONS: In this study on patients with degenerative spondylolisthesis, neither previously defined instability criteria nor other pre-specified baseline variables were associated with better clinical outcome if fusion surgery was performed. None of the analyzed variables can be applied to guide the decision for fusion surgery in patients with degenerative spondylolisthesis. For both treatment groups, less comorbidity and more leg pain were associated with improved outcome 2 years after surgery.

TRIAL REGISTRATION: NORDSTEN-DS ClinicalTrials.gov, NCT02051374.

Keywords: Decompression; Degenerative; Fusion; Predictor; Spinal stenosis; Spondylolisthesis; Subgroup; Surgery

Introduction

There is an ongoing debate on whether additional fusion surgery is superior to decompression alone in patients with spinal stenosis and degenerative spondylolisthesis. Previous studies and one recently published randomized controlled trial (RCT) support additional fusion in these patients [1−3]. However, the results of two RCTs have contradicted these findings [4,5]. As a consequence, there is great diversity in the management of degenerative spondylolisthesis on both national and individual levels [6]. In spite of increasing documentation that decompression alone is as good as decompression with fusion, surgery has increased in complexity, resulting in higher costs for patients and for society [7−9]. Still, there may be subgroups of patients that could benefit from additional fusion. Some authors have suggested that clinical characteristics and specific radiological changes indicating "instability" could be used to identify these patients [10,11]. However, no evidence for any specific radiological criteria, or a valid definition of instability, has been documented in well-designed studies [10,12]. The aim of this study was to assess possible treatment effect modifiers and prognostic variables that can support decisions about additional fusion surgery in selected patients with degenerative spondylolisthesis.

Methods

This is a secondary explorative study based on data from a national multicenter randomized trial, the Norwegian Degenerative Spondylolisthesis and Spinal Stenosis (NORDSTEN-DS) trial, in which 267 patients were included at 16 Norwegian hospitals between February 2014 and December 2017 [5]. The trial has been approved by the Regional Committee for Medical and Health Research Ethics of Central Norway (project identifier 2013/366) and

performed in accordance with the Helsinki declaration. The protocol and statistical analysis plan has been published and the trial was registered at ClinicalTrials.gov in January 2014 (NORDSTEN-DS ClinicalTrials.gov, NCT02051374) [13,14]. The trial was reported in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines, and the present study is reported in accordance with recommendations for studies of subgroups and the STAndards for the Reporting of Diagnostic accuracy studies (STARD) reporting guidelines [15–17]. The NORDSTEN-DS trial was supported by the Western Norway Regional Health Authority and the Møre and Romsdal Hospital Trust. The funding authorities had no part in trial design, data analysis, or reporting of the trial.

Study participants

Included patients were between 18 and 80 years of age with neurogenic claudication or radiating pain into the lower limbs not responding to at least 3 months of nonoperative treatment and spinal stenosis with degenerative spondylolisthesis of at least 3 mm verified on standing radiographs. Patients were not included if they had a grade 3 foraminal stenosis (deformed nerve root in the intervertebral foramen) on MRI according to the classification of Lee et al., had previous surgery at the level of the spondylolisthesis, or scoliosis of more than 20 degrees (Cobb angle) [5,18]. Further exclusion criteria were ASA- grade>3, cauda equina syndrome or complete motor deficit, Isthmic defect in pars interarticularis, fracture or former fusion of the thoracolumbar region, symptoms in one or both legs due to other diseases such as polyneuropathy, vascular claudication, or osteoarthritis. Patients with multilevel degenerative spondylolisthesis were not included. Patients were included regardless of the grade of slippage above 3 mm or motion on flexion-extension radiographs. Baseline characteristics are presented in Table 1 and a flow chart is available in the published protocol and the recently published clinical results of the RCT [5,13].

Interventions

Patients were randomized to decompression with or without fusion. In the decompression alone group, the approach could be bilateral, ipsilateral, or ipsilateral with crossover with preservation of the midline structures. In the fusion group, an optional technique was used regarding the preservation of midline structures and the use of intervertebral cages. In the fusion group implantation of pedicle screws, rods, and bone grafting was mandatory. Furthermore, the use of a microscope or magnifying glasses during decompression was mandatory. All participating surgeons were experienced with the treatments used in the trials. Before the trials began, investigators from the Scientific Board visited the hospitals to ensure a common understanding and performance of the surgical methods described in study protocols.

Table 1
Distribution of baseline variables

Variable	Mean (SD) or %	No.
Age	66 (7.6)	260
Sex (male)	32%	84/267
Comorbidity (ASA)		256
ASA 1	11%	27
ASA 2	73%	187
ASA 3	16%	42
BMI	28 (4.4)	258
Smoking, yes	14%	37/258
ODI score baseline	39 (13.2)	262
NRS back pain	6.7 (2.0)	261
NRS leg pain	6.7 (2.0)	261
HSCL-25	1.6 (0.4)	261
Spondylolisthesis, ≥20% slippage*	35%	83/235
Instability, ≥10 degrees and or ≥3 mm slippage †	21%	51/244
Foraminal stenosis, grade 0 and 1‡	85%	217/255
Facet joint fluid, ≥2 mm §	19%	49/252
Disc height, mm	7.8 (2.0)	251
Lumbar lordosis, angle	54 (11.1)	237
Pelvic incidence, angle	57 (9.4)	205
Facet joint angulation, angle	56 (8.7)	251
ODI 30% improvement at 2 y	73%	174/239

BMI, body mass index; N, number of variables; SD, standard deviation.

ODI, Oswestry Disability Index. Score from 0 to 100. Higher scores indicate more severe pain and disability [19].

NRS, Numeric Rating Scale range from 0 to 10, with lower scores indicating less pain during the past week.

HSCL-25, Hopkin's symptom check list - emotional distress score ranges from 1 to 4 with lower scores indicating less severe symptoms [13].

ASA, American Society of Anesthesiologists Classification; a physical status classification system that ranges from 1 to 6 (1 indicates a completely healthy patient) [20].

* Categorized into more or less than 20% slippage of the upper vertebrae on standing radiographs.

† Categorized as instable if angulation between flexion and extension radiographs were ≥10 degrees and or ≥3 mm translation, measured according to Dupuis et al. [21].

‡ Categorized into 0/1 and 2/3 according to classification by Lee et al. [18].

§ Categorized into more or less than 2 mm facet joint fluid [22]. Angulation of facet joints measured according to Berlemann et al. [23]. The disc height is measured on MRI as the mid-sagittal distance between the mid-superior and mid-inferior disc borders, Masharawi et al. [24]. Lumbar lordosis is measured as the angle between the superior end plate of sacrum and superior end plate of L1. Pelvic Incidence measured as an angle between the perpendicular and the sacral plate at its midpoint and the line connecting this point to the center of the femoral head. [25].

Outcome measures

The primary outcome was a change from baseline to 2-year follow-up measured on the Oswestry Disability Index (ODI) (ODI version 2.0) [19]. ODI comprises 10 questions with a sum score from 0 to 100, and higher scores indicate more severe pain and disability. Change in ODI was dichotomized to at least 30% ODI improvement (success) or not (nonsuccess) [13]. A sensitivity analysis was performed for the Zurich Claudication Questionnaire as the dependent variable. Success was defined as meeting two of three

predefined criteria at 2-year follow-up: a decrease from baseline in the score on the symptom-severity scale of 0.46 or more, a decrease from baseline on the physical-function scale of 0.42 or more, and 2.42 or less on the patient-satisfaction scale [5,26].

Selected possible treatment effect modifiers and prognostic factors

Treatment effect modifiers are patient characteristics associated with superior outcomes in one treatment group compared to another [27]. Possible treatment effect modifiers and prognostic variables were selected and predefined before we had access to the locked database (on February 5, 2020). All variables were carefully selected and categorized based on previous literature and a biological or clinical rationale (Table 1) [11,22,28−31]. The baseline characteristics of the cohort have been published at Clinicaltrials.gov and in the study protocol [13]. The radiographs and standard MRI images were acquired less than 6 months prior to surgery and evaluated independently by two experienced musculoskeletal radiologists and one experienced surgeon based on means or majority ratings. The MRI protocols included sagittal T1 and axial and sagittal T2-weighted images with slice thickness: 3 to 5 mm. An integrated measurement tool in a Picture Archiving and Communication System (IDS7 PACS, Sectra, Sweden) was used to evaluate all MRIs. The following classification systems and categories were used for the predefined baseline characteristics: Foraminal stenosis by Lee et al dichotomized into category 0 and 1 vs category 2 and 3 [18]. Disc height, lumbar lordosis, pelvic incidence, and facet joint angle were categorized into tertiles [23−25]. The facet joint fluid was measured with a cut-off at 2 mm of facet joint fluid in the joint with the most fluid (right/left) [22]. A slippage of the upper vertebrae in a standing position was categorized into more or less than 20% according to the treatment guidelines by NASS [32]. Radiographs were categorized as unstable if angulation between flexion and extension radiographs in standing position were ≥ 10 degrees and/or ≥ 3 mm translation [11,21,33]. Comorbidity was graded according to the American Society of Anesthesiologists (ASA) classification [20]. Age was analyzed as a continuous variable and additionally graded into tertiles. For more details on baseline characteristics, see Table 1. All data were collected by a local trial coordinator and entered into a database by the NORDSTEN study coordinating center at Oslo University Hospital and were inaccessible to the research group until database unlock.

Statistical analyses

Statistical analyses were performed with Stata (version 16) or SPSS (version 27, SPSS Inc., Chicago, IL, USA). Missing values of treatment effect modifiers or prognostic factors were not imputed. The effect of possible treatment effect modifiers in each treatment group was analyzed according to intention-to-treat principles. To check for multicollinearity, we computed Spearman's correlation coefficients for all pairwise comparisons of possible predictive factors. A multiple logistic regression analysis was utilized in accordance with the purposeful selection model by Hosmer Lemeshow [34]. In the first step, univariate analyses of all selected baseline variables were performed. Variables with p<.25 from univariate analyses were included in the multivariate model. Age and gender were included in all multivariate models and in the final model according to our prespecified protocol. The lowest category was chosen as the reference category for all categorical variables. In the next step, covariates were removed in descending order according to the p-value of .1. In the third step, covariates not selected in the univariate analysis were added one at a time and evaluated as possible confounders. Confounders were defined as variables associated with a change in remaining covariates of more than 15% and were consequently included in the final model. Finally, the correct parametric forms for continuous variables and plausible interactions were investigated. The final model was tested for goodness-of-fit and consists of the variables considered to be of predictive value adjusted for age and gender [34]. The described analyses were performed for the whole cohort of patients and stratified by treatment groups. The results from the stratified analyses are depicted graphically using Forest plots of odds ratios with 95% confidence intervals (CI).

Additionally, we performed two sensitivity analyses: One analysis for the Zurich Claudication Questionnaire as the dependent variable, the other for ODI as a dependent variable, considering patients who underwent fusion surgery after decompression as nonsuccesses.

All analyses were considered exploratory, so no correction for multiple testing was performed and p-values <.05 were considered statistically significant.

Results

Out of the 267 patients included in the original randomized trial (183 female [67%]; mean [SD] age, 66 [7.6] years), baseline variables were available for 205 to 267 individuals (Table 1). The flow chart and further baseline data have recently been published [5]. The final model had acceptable goodness-of-fit (p=.94). We did not detect any multicollinearity among the selected possible predictive factors or any significant correlations when performing pairwise comparisons. The Spearman's correlation coefficient was <0.7.

Stratified analyses of potential treatment effect modifiers

Assessed by the primary outcome, ODI, no clinical or radiological baseline variables modified the treatment effect in the two treatment groups analyzed separately. We did not fit multivariate regression models as none of the possible predictive factors reached a sufficient level of statistical significance and, consequently, only the univariate analyses

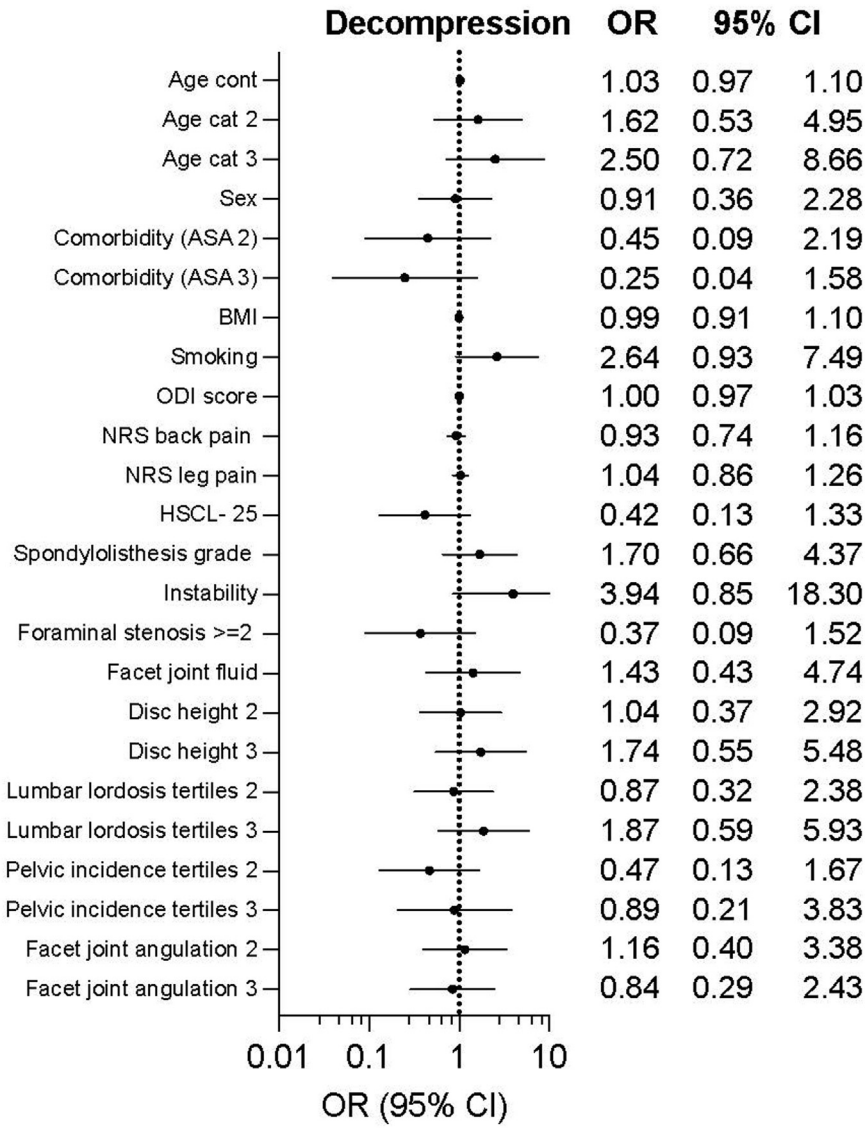

Fig. 1. Univariate logistic regression of possible treatment modifiers for 30% ODI points improvement as dependent variable for the Decompression group. The results are expressed as odds ratios (OR) with 95% confidence intervals (CI). The vertical line through 1 indicates no statistically significant association. When confidence intervals (horizontal lines around the point estimate of OR) cross the line, there is no significant association. Values to the right of the vertical lines indicate increased odds for 30% ODI improvement, values to the left decreased odds of ODI improvement. ODI, Oswestry Disability Index. Score from 0 to 100. Higher scores indicate more severe pain and disability [19]. NRS, Numeric Rating Scale range 0 to 10, with lower scores indicating less pain during the past week; HSCL-25, Hopkin's symptom check list - emotional distress score ranges from 1 to 4 with lower scores indicating less severe symptoms [13]. ASA, American Society of Anesthesiologists Classification, a physical status classification system that ranges from 1 to 6 (1 indicates a completely healthy patient) [20]. BMI, Body Mass Index. Spondylolisthesis grade; categorized into more or less than 20% slippage of the upper vertebrae on standing radiographs. Instability; categorized as instable if angulation between flexion and extension radiographs were ≥10 degrees and or ≥3 mm translation, measured according to Dupuis et al. [21]. Foraminal stenosis; categorized into 0/1 and 2/3 according to classification by Lee et al. [18]. Facet joint fluid; categorized into more or less than 2 mm facet joint fluid [22]. Angulation of facet joints measured according to Berlemann et al. [23]. The disc height is measured on MRI as the mid-sagittal distance between mid-superior and mid-inferior disc borders Masharawi et al. [24]. Lumbar lordosis is measured as the angle between the superior end plate of sacrum and superior end plate of L1. Pelvic incidence is measured as an angle between the perpendicular and the sacral plate at its midpoint and the line connecting this point to the center of the femoral head [25]. ASA 2 and 3 are compared to ASA 1 and tertiles 2 and 3 are compared to tertile 1.

for the two groups are presented using forest plots (Figs. 1 and 2).

Additionally, the sensitivity analyses of the Zurich Claudication Questionnaire as a dependent variable and for ODI as a dependent variable, considering patients who underwent fusion surgery after decompression as nonsuccesses, were in line with the main prespecified analysis (e Figs. 1–4)

Analysis of potential prognostic factors for the whole cohort

The univariate analyses are presented in Table 2. The following prognostic factors for the whole cohort were identified in the final multivariate model: less comorbidity and more leg pain were associated with an improvement of

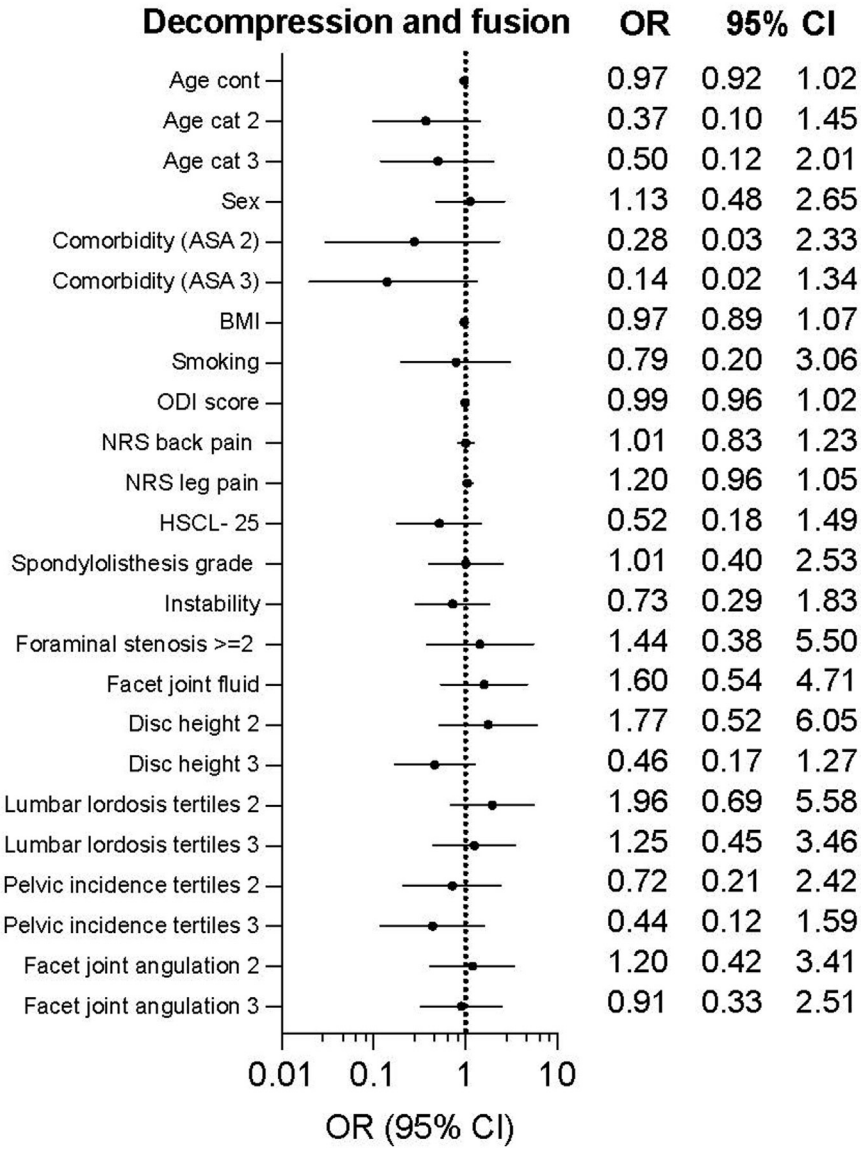

Fig. 2. Univariate logistic regression of possible treatment modifiers for 30% ODI points improvement as dependent variable for the Decompression and Fusion group. The results are expressed as odds ratios (OR) with 95% confidence intervals (CI). The vertical line through 1 indicates no statistically significant association. When confidence intervals (horizontal lines around the point estimate of OR) cross the line, there is no significant association. Values to the right of the vertical lines indicate increased odds for 30% ODI improvement, values to the left decreased odds of ODI improvement. ODI, Oswestry Disability Index. Score from 0 to 100. Higher scores indicate more severe pain and disability [19]. NRS, Numeric Rating Scale ranges from 0 to 10, with lower scores indicating less pain during the past week. HSCL-25; Hopkin's symptom checklist - emotional distress score ranges from 1 to 4 with lower scores indicating less severe symptoms [13]. ASA, American Society of Anesthesiologists Classification, a physical status classification system that ranges from 1 to 6 (1 indicates a completely healthy patient) [20]. BMI, Body Mass Index. Spondylolisthesis grade; categorized into more or less than 20% slippage of the upper vertebrae on standing radiographs. Instability; categorized as instable if angulation between flexion and extension radiographs were ≥10 degrees and or ≥3mm translation, measured according to Dupuis et al. [21]. Foraminal stenosis; categorized into 0/1 and 2/3 according to classification by Lee et al. [18]. Facet joint fluid; categorized into more or less than 2 mm facet joint fluid [22]. Angulation of facet joints measured according to Berlemann et al. [23]. The disc height is measured on MRI as the mid-sagittal distance between mid-superior and mid-inferior disc borders Masharawi et al. [24]. Lumbar lordosis is measured as the angle between the superior end plate of sacrum and superior end plate of L1. Pelvic incidence measured as an angle between the perpendicular and the sacral plate at its midpoint and the line connecting this point to the center of the femoral head [25]. ASA 2 and 3 are compared to ASA 1 and tertiles 2 and 3 are compared to tertile 1.

at least 30% ODI points from baseline to 2-year follow-up (OR = 4.35; 95% CI [1.16−16.67] and (OR = 1.23; 95% CI [1.02−1.50]).

The association between leg pain and a positive outcome after surgery was reduced with increasing preoperative back pain intensity. For details, see Table 3.

Discussion

Neither the degree of slippage, instability, or any other radiological or clinical variables had a significant impact on outcome when the two treatment groups were analyzed separately. Hence, we cannot recommend using any of the

Table 2
Univariate model. Associations of baseline characteristics and 30% ODI points improvement in 267 patients treated surgically for spinal stenosis and degenerative spondylolisthesis (analysis of the whole cohort)

	B	OR (CI)	p-value
Age cont.	0.00	1.00 (0.97–1.04)	.91
Age tertiles			.77
Age tertile 2	-0.01	0.99 (0.45–2.17)	.98
Age tertlel 3	0.21	1.24 (0.54–2.84)	.61
Sex (male)	-0.4	0.96 (0.52–1.78)	.90
Comorbidity (ASA)			.09
ASA 2	-0.73	0.48 (0.16–1.47)	.20
ASA 3	-1.34	0.26 (0.07–0.92)	.04
BMI	0.01	1.01 (0.94–1.07)	.85
Smoking	0.51	1.67 (0.77–3.63)	.20
ODI score	-0.01	0.99 (0.97–1.02)	.59
NRS back pain	-0.04	0.96 (0.84–1.11)	.62
NRS leg pain	0.10	1.11 (0.96–1.28)	.16
HSCL-25	-0.71	0.49 (0.23–1.05)	.07
Spondylolisthesis grade*	0.04	1.04 (0.56–1.94)	.90
Instability†	0.19	1.21 (0.58–2.51)	.61
Foraminal stenosis ≥2‡	-0.25	0.78 (0.35–1.76)	.55
Facet joint fluid§	0.34	1.40 (0.65–3.03)	.39
Disc height tertiles			.44
Disc height tertile 2	0.28	1.32 (0.62–2.81)	.47
Disc height tertile 3	-0.20	0.82 (0.40–1.65)	.58
Lumbar lordosis tertiles			.44
Lumbar lordosis tertile 2	0.28	1.32 (0.62–2.81)	.47
Lumbar lordosis tertile 3	-0.20	0.82 (0.40–1.65)	.58
Pelvic incidence tertiles			.66
Pelvic incidence tertile 2	-0.36	0.70 (0.31–1.58)	.38
Pelvic incidence tertile 3	-0.34	0.71 (0.29–1.75)	.46
Facet joint angulation tertiles			.76
Facet joint angulation tertile 2	0.14	1.15 (0.56–2.38)	.71
Facet joint angulation tertile 3	-0.13	0.88 (0.43–1.79)	.72

BMI, body mass index; OR, odds ratio; B, regression coefficient, CI, 95% confidence interval.

ODI, Oswestry Disability Index. Score from 0 to 100. Higher scores indicate more severe pain and disability [19].

NRS, Numeric Rating Scale range from 0 to 10, with lower scores indicating less pain during the past week.

HSCL-25, Hopkin's symptom check list - emotional distress score ranges from 1 to 4 with lower scores indicating less severe symptoms [13].

ASA, American Society of Anesthesiologists Classification; a physical status classification system that ranges from 1 to 6 (1 indicates a completely healthy patient) [20].

p values; Indicate if the baseline variable is associated with 30% ODI points improvement or not at 2-year follow-up

* Categorized into more or less than 20% slippage of the upper vertebrae on standing radiographs.

† Categorized as instable if angulation between flexion and extension radiographs were ≥10 degrees and or ≥3mm translation, measured according to Dupuis et al. [21].

‡ Categorized into 0/1 and 2/3 according to classification by Lee et al. [18].

§ Categorized into more or less than 2 mm facet joint fluid [22].

Angulation of facet joints measured according to Berlemann et al. [23]. The disc height is measured on MRI as the mid-sagittal distance between mid-superior and mid-inferior disc borders Masharawi et al. [24]. Lumbar lordosis is measured as the angle between the superior end plate of the sacrum and superior end plate of L1. Pelvic incidence is measured as an angle between the perpendicular and the sacral plate at its midpoint and the line connecting this point to the center of the femoral head [25].

ASA 2 and 3 are compared to ASA 1 and tertiles 2 and 3 are compared to tertile 1.

Table 3
Final model. Associations of baseline characteristics and 30% ODI points improvement in 267 patients treated surgically for spinal stenosis and degenerative spondylolisthesis (analysis of the whole cohort)

Variable	No.	B	OR (CI)	p-value
Comorbidity (ASA)	230			.07
ASA 2		-0.79	0.45 (0.14–1.43)	.18
ASA 3		-1.49	0.23 (0.06–0.86)	.030
NRS leg pain	230	0.21	1.23 (1.02–1.50)	.035
NRS back pain	230	-0.14	0.87 (0.71–1.07)	.18
Sex	230	-0.16	0.85 (0.44–1.63)	.63
Age continuous	230	0.01	1.01 (0.97–1.05)	.56

ASA (ref = no comorbidity), Sex (ref = males).

N represents availability of baseline variables and outcome.

OR, odds ratio; B, regression coefficient; CI, 95% Confidence Interval.

p-values; Indicate if the baseline variable is associated with 30% ODI points improvement or not at 2-year follow-up.

ASA, American Society of Anesthesiologists Classification; a physical status classification system that ranges from 1 to 6 (1 indicates a completely healthy patient) [20].

ASA 2 and 3 are compared to ASA 1.

NRS, Numeric Rating Scale range from 0 to 10, with lower scores indicating less pain during the past week.

analyzed criteria to guide the decision for fusion surgery in patients with degenerative spondylolisthesis. Analyzing the whole cohort regardless of treatment, we found that healthier patients with more preoperative leg pain were more likely to report an improvement of at least 30% ODI points. The association between leg pain and a 30% improvement decreased with increasing preoperative back pain.

Patients with instability were eligible for this study, and 21% had a slippage of at least 3 mm, or at least 10 degrees of angulation assessed by dynamic standing radiographs. In line with recommendations from NASS by Matz et al. [32], we also investigated a slippage of more or less than 20% of the upper vertebrae on standing radiographs, but there is no firm evidence that patients with slippage of more than 20% benefit from additional fusion. In our study, patients with a high-grade spondylolisthesis (>20%) or dynamic slippage of the vertebral bodies on flexion-extension radiographs reported a similar treatment effect on outcome in the decompression alone and decompression with fusion groups. Aihara et al. [35] reported results similar to those of our study for patients with a slippage over 20% in neutral position radiographs. Pearson et al. used the cut-off of more than 10 degrees according to recommendations by Hanley et al., based on the biomechanical studies by Posner et al. [11,30,33]. In line with our findings, they concluded that instability on dynamic standing radiographs was not associated with outcome. Contrary, Yone et al. [31] suggested that the instability criteria published by Posner et al. may aid in selecting patients that might benefit from fusion.

The remaining radiological variables assessed in this study have been discussed in several papers without any consensus or consistent prediction of outcome [10,36–38]. The value of radiological parameters in determining segmental instability

or indicating fusion surgery, therefore remains unknown. Several studies reported that increased facet joint fluid may suggest instability [22,29,39]. Similar to the study by Lattig et al., we chose a cut-off of 2 mm facet joint fluid (19% had at least 2 mm facet joint fluid) [22]. We found that the amount of facet joint fluid did not modify the treatment effect. In recent years, researchers have attempted to evaluate the role of sagittal balance and facet joint orientation among patients with degenerative spondylolisthesis. Increased pelvic incidence and lumbar lordosis can lead to larger shear forces, enhance the forces on the facet joints, and lead to remodeling of the joints [40]. These variables did not modify the treatment effect in patients with degenerative spondylolisthesis in our study. According to our protocol, patients with grade 3 foraminal stenosis were not included, but foraminal stenosis grade 2 did not modify the treatment effect. We considered age to be a possible treatment effect modifier and analyzed age as both a continuous and a categorical variable [33,41–44]. In this study, age was neither a treatment effect modifier nor a prognostic factor at 2-year follow-up.

In line with previously published studies, preoperative level of leg pain predicted the outcome. In the SPORT study, patients with predominant leg pain improved more than patients with predominant back pain [45]. Among patients treated surgically for spinal stenosis with and without degenerative spondylolisthesis, several studies have reported that predominant back pain decreases the odds of patient satisfaction [46,47].

Strength and limitations

This study has several limitations. The generalizability of the findings is limited by our eligibility criteria and our patient population, thus our conclusions may not apply to patients with degenerative scoliosis, extensive foraminal stenosis (Lee grade 3), previous surgery, or multilevel spondylolisthesis [18]. For many of these patients, fusion is an established practice. The conclusions may also be limited by the choice of short-term follow-up, the variables analyzed, and cut-off criteria used. We did not investigate the influence of local kyphosis at the treated lumbar segment or difference in slippage between radiographs in standing and supine position. Further, the choice of angulation and translation on dynamic radiographs could impact our conclusions. However, the selected instability criteria are frequently mentioned in the literature [10]. Furthermore, there is a chance of overfitting due to the size of the dataset. The ODI as primary outcome was used in the original article and prespecified in the protocol article [5,13]. Consequently, ODI was chosen as the main outcome of this study. ODI is the most commonly utilized outcome measure for patients with spinal stenosis, but other more specific outcomes, such as Zurich Claudication Score, have also been recommended-[32,48,49]. Strengths of this study are the randomized design that allows us to investigate suggested instability criteria among similar groups of patients treated both with or without fusion; the prospective collection of predefined baseline variables; and the evaluation of the radiological data by three independent raters.

About 10% of the patients in the randomized trial (NORDSTEN-DS trial) were fused after decompression alone which might justify a search for baseline characteristics to identify small subgroups that could benefit from additional fusion [5]. Based on existing evidence, we question the validity of existing instability criteria used to select patients for fusion and suggest that the existing criteria should be re-evaluated. To improve surgical treatment for individual patients with degenerative spondylolisthesis, high-quality multicenter studies or register studies with sufficient power and high-quality radiographs at baseline are required.

Conclusions

Neither previously defined instability criteria nor other prespecified baseline variables modified the treatment effect of decompression alone or decompression with an additional fusion for patients with degenerative spondylolisthesis. Based on the variables investigated in this study, we cannot recommend fusion surgery as the first choice of treatment. For the whole cohort, patients with less comorbidity, or more leg pain had better outcomes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Role of the Funder/Sponsor

The funding authorities had no part in trial design, data analysis, or reporting the trial.

Data Sharing Statement

Deidentified participant data (including data dictionary) will be available to medical researchers by request in accordance with local registration and ethical approval. Data requests can be submitted at any time and the data will be accessible for 12 months after the article has been published, with possible extensions considered. All proposals requesting data access will need to specify a statistical analysis plan and will need approval of the scientific board before any data can be released. Requests to access data should be addressed to christianhellum1@gmail.com.

Acknowledgments

The NORDSTEN study is funded by governmental organizations (the Central Norway Regional Health Authority [grant 10174] and the Western Norway Regional Health Authority [grant 912046]) and by the nonprofit foundation Sophies Minde Ortopedi (www.sophiesminde.no, grant31/

2021). We thank the patients who volunteered for this trial; Inger Ljostad, our patient representative, for her involvement in implementing the trial; Katarina Mølsæter (Data and Safety Monitoring Board, Møre and Romsdal Health Trust) for her dedication and her diligent review of the data; Anniken Remseth and Bodil Røyset for managing the trial finances; Inger Ljostad, Jørn Aaen, and Masoud Anwar for their involvement in the NORDSTEN study group; the members of the NORDSTEN working group, including Turid Rognsvåg, Janne Haugland, Anita K. Vassbakk, Vidar Punsvik, Kirstine Eikenes, Therese Gundersen, Trine Myrvold, Yngve Sporstøl, Ted P. Lundgren, Terje Fallås, Paal Arnesen, Anne-Charlotte Fosse Haug, Elisabeth Lilleholdt Muller, Kjartan Krogedal, Bettina Timenes, Vidar Opland, Merete Finjarn, Hege Bergum, Samer Habiba, Kristine Helland, Maria Rieber-Mohn, Andreas Seip, Greger Lønne, Gisle Szacinski, Espen Sjåberg, Ellen Aksnes, Hege Andresen, Andreas Sørlie, Kristine Evanger, Knut Jørgen Haug, Ingrid Edborg, and all the local trial coordinators and surgeons not named here who contributed to the success of this trial; the staff at the Research and Communication Unit for Musculoskeletal Health (FORMI) at Oslo University Hospital (Ingrid Fjellheim Bånerud, Maren Hjelle Guddal, Eira Ebbs, Marie Skovli Pettersen, and Ørjan Nesse Vigdal) for meticulous work with data collection; Eira Ebbs (FORMI) for providing linguistic assistance in writing the manuscript, and Filip C Dolatowski for critical review of the manuscript.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.spinee.2023.06.386.

References

[1] Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. 1991;73(6):802−8.

[2] Bridwell KH, Sedgewick TA, O'Brien MF, Lenke LG, Baldus C. The role of fusion and instrumentation in the treatment of degenerative spondylolisthesis with spinal stenosis. J Spinal Disord 1993;6(6):461–72.

[3] Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. 2016;374(15):1424−34.

[4] Försth P, Ólafsson G, Carlsson T, et al. A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. 2016;374(15):1413−23.

[5] Austevoll IM, Hermansen E, Fagerland MW, et al. Decompression with or without fusion in degenerative lumbar spondylolisthesis. N Engl J Med 2021;385(6):526–38.

[6] Lønne G, Schoenfeld AJ, Cha TD, Nygaard Ø P, Zwart JAH, Solberg T. Variation in selection criteria and approaches to surgery for Lumbar Spinal Stenosis among patients treated in Boston and Norway. Clin Neurol Neurosurg 2017;156:77–82.

[7] Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. 2010;303(13):1259−65.

[8] Kepler CK, Vaccaro AR, Hilibrand AS, et al. National trends in the use of fusion techniques to treat degenerative spondylolisthesis. Spine (Phila Pa 1976) 2014;39(19):1584–9.

[9] Kuntz KM, Snider RK, Weinstein JN, Pope MH, Katz JN. Cost-effectiveness of fusion with and without instrumentation for patients with degenerative spondylolisthesis and spinal stenosis. 2000;25(9):1132−9.

[10] Simmonds AM, Rampersaud YR, Dvorak MF, Dea N, Melnyk AD, Fisher CG. Defining the inherent stability of degenerative spondylolisthesis: a systematic review. J Neurosurg Spine 2015;23(2):178–89.

[11] Posner I, White 3rd AA, Edwards WT, Hayes WC. A biomechanical analysis of the clinical stability of the lumbar and lumbosacral spine. Spine (Phila Pa 1976) 1982;7(4):374–89.

[12] Steiger F, Becker HJ, Standaert CJ, et al. Surgery in lumbar degenerative spondylolisthesis: indications, outcomes and complications. A systematic review. Eur Spine J 2014;23(5):945–73.

[13] Austevoll IM, Hermansen E, Fagerland M, et al. Decompression alone versus decompression with instrumental fusion the NORDSTEN degenerative spondylolisthesis trial (NORDSTEN-DS); study protocol for a randomized controlled trial. BMC Musculoskelet Disord 2019;20(1):7.

[14] Austevoll I. The NORDSTEN Study - Degenerative Spondylolisthesis (NORDSTEN/DS). Clinicaltrials.gov 2014 Available at: https://clinicaltrials.gov/ct2/show/NCT02051374?term=nordsten&cond=-Spondylolisthesis&draw=2&rank=1. [accessed January 6, 2022]

[15] Kent P, Keating JL, Leboeuf-Yde C. Research methods for subgrouping low back pain. BMC Med Res Methodol 2010;10:62.

[16] Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016;6(11). e012799-e.

[17] Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg (London, England) 2012;10(1):28–55.

[18] Lee S, Lee JW, Yeom JS, et al. A practical MRI grading system for lumbar foraminal stenosis. AJR Am J Roentgenol 2010;194(4):1095–8.

[19] Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66(8):271–3.

[20] Ament R. Origin of the ASA classification. Anesthesiology 1979;51(2):179.

[21] Dupuis PR, Yong-Hing K, Cassidy JD, Kirkaldy-Willis WH. Radiologic diagnosis of degenerative lumbar spinal instability. Spine (Phila Pa 1976) 1985;10(3):262–76.

[22] Lattig F, Fekete TF, Grob D, Kleinstück FS, Jeszenszky D, Mannion AF. Lumbar facet joint effusion in MRI: a sign of instability in degenerative spondylolisthesis? Eur Spine J 2012;21(2):276–81.

[23] Berlemann U, Jeszenszky DJ, Bühler DW, Harms J. Facet joint remodeling in degenerative spondylolisthesis: an investigation of joint orientation and tropism. Eur Spine J 1998;7(5):376–80.

[24] Masharawi Y, Kjaer P, Bendix T, et al. The reproducibility of quantitative measurements in lumbar magnetic resonance imaging of children from the general population. Spine (Phila Pa 1976) 2008;33(19):2094–100.

[25] Duval-Beaupère G, Schmidt C, Cosson P. A barycentremetric study of the sagittal shape of spine and pelvis: the conditions required for an economic standing position. Ann Biomed Eng 1992;20(4):451–62.

[26] Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. Spine (Phila Pa 1976) 2006;31(11):1276–80.

[27] Christensen R, Bours MJL, Nielsen SM. Effect modifiers and statistical tests for interaction in randomized trials. J Clin Epidemiol 2021;134:174–7.

[28] Kleinstueck F, Fekete T, Mannion A, et al. To fuse or not to fuse in lumbar degenerative spondylolisthesis: do baseline symptoms help provide the answer? 2012;21(2):268−75.

[29] Lattig F, Fekete TF, Kleinstück FS, Porchet F, Jeszenszky D, Mannion AF. Lumbar facet joint effusion on MRI as a sign of unstable degenerative spondylolisthesis: should it influence the treatment decision? 2015;28(3):95−100.
[30] Pearson AM, Lurie JD, Blood EA, et al. Spine patient outcomes research trial: radiographic predictors of clinical outcomes after operative or nonoperative treatment of degenerative spondylolisthesis. Spine (Phila Pa 1976) 2008;33(25):2759–66.
[31] Yone K, Sakou T. Usefulness of Posner's definition of spinal instability for selection of surgical treatment for lumbar spinal stenosis. J Spinal Disord 1999;12(1):40–4.
[32] Matz PG, Meagher R, Lamer T, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. 2016;16(3):439−48.
[33] Hanley Jr. EN. The indications for lumbar spinal fusion with and without instrumentation. Spine (Phila Pa 1976) 1995;20(Suppl. 24):143s–53s.
[34] Hosmer DW, Lemeshow S, May S. Applied survival analysis. Regression modeling of time-to-event data. Hoboken: John Wiley and Sons; 2008.
[35] Aihara T, Toyone T, Murata Y, Inage K, Urushibara M, Ouchi J. Degenerative lumbar spondylolisthesis with spinal stenosis: a comparative study of 5-year outcomes following decompression with fusion and microendoscopic decompression. Asian Spine J 2018;12(1):132–9.
[36] Aalto TJ, Malmivaara A, Kovacs F, et al. Preoperative predictors for postoperative clinical outcome in lumbar spinal stenosis: systematic review. Spine (Phila Pa 1976) 2006;31(18):E648–63.
[37] Blumenthal C, Curran J, Benzel EC, et al. Radiographic predictors of delayed instability following decompression without fusion for degenerative grade I lumbar spondylolisthesis. 2013;18(4):340−6.
[38] Joaquim AF, Milano JB, Ghizoni E, Patel A. Is there a role for decompression alone for treating symptomatic degenerative lumbar spondylolisthesis?: A systematic review. 2016;29(5):191−202.
[39] Snoddy MC, Sielatycki JA, Sivaganesan A, Engstrom SM, McGirt MJ, Devin CJ. Can facet joint fluid on MRI and dynamic instability be a predictor of improvement in back pain following lumbar fusion for degenerative spondylolisthesis? 2016;25(8):2408−15.
[40] Guo M, Kong C, Sun S, Sun X, Li X, Lu SJ. Predictors of L4− L5 degenerative lumbar spondylolisthesis: L4 inclination angle and facet joint angle. 2019;130:e680−e6.
[41] Mummaneni PV, Bydon M, Alvi MA, et al. Predictive model for long-term patient satisfaction after surgery for grade I degenerative lumbar spondylolisthesis: insights from the Quality Outcomes Database. Neurosurg Focus 2019;46(5):E12.
[42] Sigmundsson FG, Jönsson B, Strömqvist B. Determinants of patient satisfaction after surgery for central spinal stenosis without concomitant spondylolisthesis: a register study of 5100 patients. Eur Spine J 2017;26(2):473–80.
[43] Crawford 3rd CH, Glassman SD, Djurasovic M, Owens 2nd RK, Gum JL, Carreon LY. Prognostic factors associated with best outcomes (minimal symptom state) following fusion for lumbar degenerative conditions. Spine J 2019;19(2):187–90.
[44] Matsunaga S, Ijiri K, Hayashi K. Nonsurgically managed patients with degenerative spondylolisthesis: a 10- to 18-year follow-up study. J Neurosurg 2000;93(2):194–8.
[45] Pearson A, Blood E, Lurie J, et al. Predominant leg pain is associated with better surgical outcomes in degenerative spondylolisthesis and spinal stenosis: results from the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976) 2011;36(3):219–29.
[46] Kleinstück FS, Grob D, Lattig F, et al. The influence of preoperative back pain on the outcome of lumbar decompression surgery. Spine (Phila Pa 1976) 2009;34(11):1198–203.
[47] Atlas SJ, Keller RB, Robson D, Deyo RA, Singer DE. Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the maine lumbar spine study. Spine (Phila Pa 1976) 2000;25(5):556–62.
[48] Winebrake JP, Lovecchio F, Steinhaus M, Farmer J, Sama A. Wide variability in patient-reported outcomes measures after fusion for lumbar spinal stenosis: a systematic review. Global Spine J 2020;10(2):209–15.
[49] Fujimori T, Ikegami D, Sugiura T, Sakaura H. Responsiveness of the Zurich Claudication Questionnaire, the Oswestry Disability Index, the Japanese Orthopaedic Association Back Pain Evaluation Questionnaire, the 8-Item Short Form Health Survey, and the Euroqol 5 dimensions 5 level in the assessment of patients with lumbar spinal stenosis. Eur Spine J 2022;31(6):1399–412.