ORIGINAL ARTICLE

Decompression with or without Fusion in Degenerative Lumbar Spondylolisthesis

I.M. Austevoll, E. Hermansen, M.W. Fagerland, K. Storheim, J.I. Brox, T. Solberg, F. Rekeland, E. Franssen, C. Weber, H. Brisby, O. Grundnes, K.R.H. Algaard, T. Böker, H. Banitalebi, K. Indrekvam, and C. Hellum, for the NORDSTEN-DS Investigators*

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Austevoll at Kysthospitalet in Hagevik, Orthopedic Department, Haukeland University Hospital, Hagaviksbakken 25, 5217 Hagevik, Bergen, Norway, or at imau@helse-bergen.no.

*A full list of the investigators in the NORDSTEN-DS trial is provided in the Supplementary Appendix, available at NEJM.org.

N Engl J Med 2021;385:526-38.
DOI: 10.1056/NEJMoa2100990

ABSTRACT

BACKGROUND

In patients with lumbar spinal stenosis and degenerative spondylolisthesis, it is uncertain whether decompression surgery alone is noninferior to decompression with instrumented fusion.

METHODS

We conducted an open-label, multicenter, noninferiority trial involving patients with symptomatic lumbar stenosis that had not responded to conservative management and who had single-level spondylolisthesis of 3 mm or more. Patients were randomly assigned in a 1:1 ratio to undergo decompression surgery (decompression-alone group) or decompression surgery with instrumented fusion (fusion group). The primary outcome was a reduction of at least 30% in the score on the Oswestry Disability Index (ODI; range, 0 to 100, with higher scores indicating more impairment) during the 2 years after surgery, with a noninferiority margin of –15 percentage points. Secondary outcomes included the mean change in the ODI score as well as scores on the Zurich Claudication Questionnaire, leg and back pain, the duration of surgery and length of hospital stay, and reoperation within 2 years.

RESULTS

The mean age of patients was approximately 66 years. Approximately 75% of the patients had leg pain for more than a year, and more than 80% had back pain for more than a year. The mean change from baseline to 2 years in the ODI score was –20.6 in the decompression-alone group and –21.3 in the fusion group (mean difference, 0.7; 95% confidence interval [CI], –2.8 to 4.3). In the modified intention-to-treat analysis, 95 of 133 patients (71.4%) in the decompression-alone group and 94 of 129 patients (72.9%) in the fusion group had a reduction of at least 30% in the ODI score (difference, –1.4 percentage points; 95% CI, –12.2 to 9.4), showing the noninferiority of decompression alone. In the per-protocol analysis, 80 of 106 patients (75.5%) and 83 of 110 patients (75.5%), respectively, had a reduction of at least 30% in the ODI score (difference, 0.0 percentage points; 95% CI, –11.4 to 11.4), showing noninferiority. The results for the secondary outcomes were generally in the same direction as those for the primary outcome. Successful fusion was achieved with certainty in 86 of 100 patients (86.0%) who had imaging available at 2 years. Reoperation was performed in 15 of 120 patients (12.5%) in the decompression-alone group and in 11 of 121 patients (9.1%) in the fusion group.

CONCLUSIONS

In this trial involving patients who underwent surgery for degenerative lumbar spondylolisthesis, most of whom had symptoms for more than a year, decompression alone was noninferior to decompression with instrumented fusion over a period of 2 years. Reoperation occurred somewhat more often in the decompression-alone group than in the fusion group. (NORDSTEN-DS ClinicalTrials.gov number, NCT02051374.)

Degenerative spondylolisthesis is a condition in which one vertebra has slipped forward in relation to the next lower vertebra. Patients, usually older than 60 years of age, typically have leg and back pain and restricted function due to lumbar spinal stenosis caused by spondylolisthesis, disc bulging, and ligamentous and facet hypertrophy.[1] Surgical treatment is often recommended in patients whose pain has not decreased with conservative management.[2,3] In the past few decades, instrumented fusion (the use of screws, rods, plates, or other apparatus to assist in achieving fusion between vertebral bodies by bone grafts) in addition to decompression of the lumbar canal has been widely used to treat lumbar spondylolisthesis and accounts for more than 90% of decompression surgeries in some countries.[4,5] In the United States, the hospital costs of lumbar instrumented fusion procedures were estimated at $13 billion in 2011, higher than for any other surgical procedure, such as knee and hip arthroplasty and percutaneous coronary angioplasty.[6] Degenerative spondylolisthesis accounts for the largest proportion of these procedures, and the number of surgeries for this condition more than doubled from 2004 to 2015.[6,7] Nevertheless, there is a large variation in practice, and in some countries approximately 50% of patients undergo decompression surgery without fusion.[8] Instead of laminectomy, which was the original method of decompression, less invasive procedures that preserve the integrity of stabilizing structures of the spine are now commonly used.[9,10]

Two randomized, controlled trials that were published in the *Journal* in 2016 investigated the benefit of adding fusion to decompression in patients with lumbar spondylolisthesis.[11,12] The conclusions of the trials diverged, but an editorial that summarized the results suggested that fusion does not add value for this patient group.[13] There has been debate about the interpretation of these two trials because of different entry criteria,[14,15] equivocal conclusions from subsequent meta-analyses and systematic reviews,[16-18] and requests for further high-quality studies.[19,20] In the Norwegian Degenerative Spondylolisthesis and Spinal Stenosis (NORDSTEN-DS) trial, we investigated whether decompression surgery was noninferior to decompression surgery with instrumented fusion in patients who have spinal stenosis with degenerative spondylolisthesis.

METHODS

TRIAL DESIGN AND OVERSIGHT

The NORDSTEN-DS trial is an investigator-initiated, multicenter, randomized, open-label, parallel-group, noninferiority[21] trial. The Regional Committee for Medical and Health Research Ethics of Central Norway approved the trial (project identifier 2013/366). The trial was independently monitored in accordance with a modified model of the International Council for Harmonisation guidelines for Good Clinical Practice.[22] Information about monitoring and patient involvement is provided in Section 3 in the Supplementary Appendix, available with the full text of this article at NEJM.org. The trial is reported according to the recommendations of Consolidated Standards of Reporting Trials.[23] The initial draft of the manuscript was written by the first, third, and last authors, who had full access to all data. The analyses were conducted by the third author, who was unaware of the treatment-group assignments. All the authors vouch for the completeness and accuracy of the data and for the fidelity of the trial to the protocol. The trial protocol[24] and the statistical analysis plan[25] have been published previously and are available at NEJM.org.

When 150 of the trial patients had completed their 1-year assessment, a third-party statistician, who was unaware of the treatment-group assignments, performed an interim analysis for safety and efficacy according to the protocol.[24] Permission to continue the trial, and no other results, was disclosed to the research group on February 28, 2017.

TRIAL POPULATION

Eligible patients were 18 to 80 years of age with neurogenic claudication or radicular radiating pain in the lower limbs that had not responded to at least 3 months of conservative care. Patients had to have radiographic evidence of spinal stenosis verified by magnetic resonance imaging (MRI) and have degenerative spondylolisthesis solely at the stenotic level of at least 3 mm verified by standing plain radiographs obtained in the lateral view. Patients were included regardless of the grade of slippage above 3 mm and regardless of the result of the flexion–extension radiographs to detect dynamic slippage of vertebral bodies. Patients were not to be included if they had foraminal stenosis of grade 3 (i.e., a deformed nerve

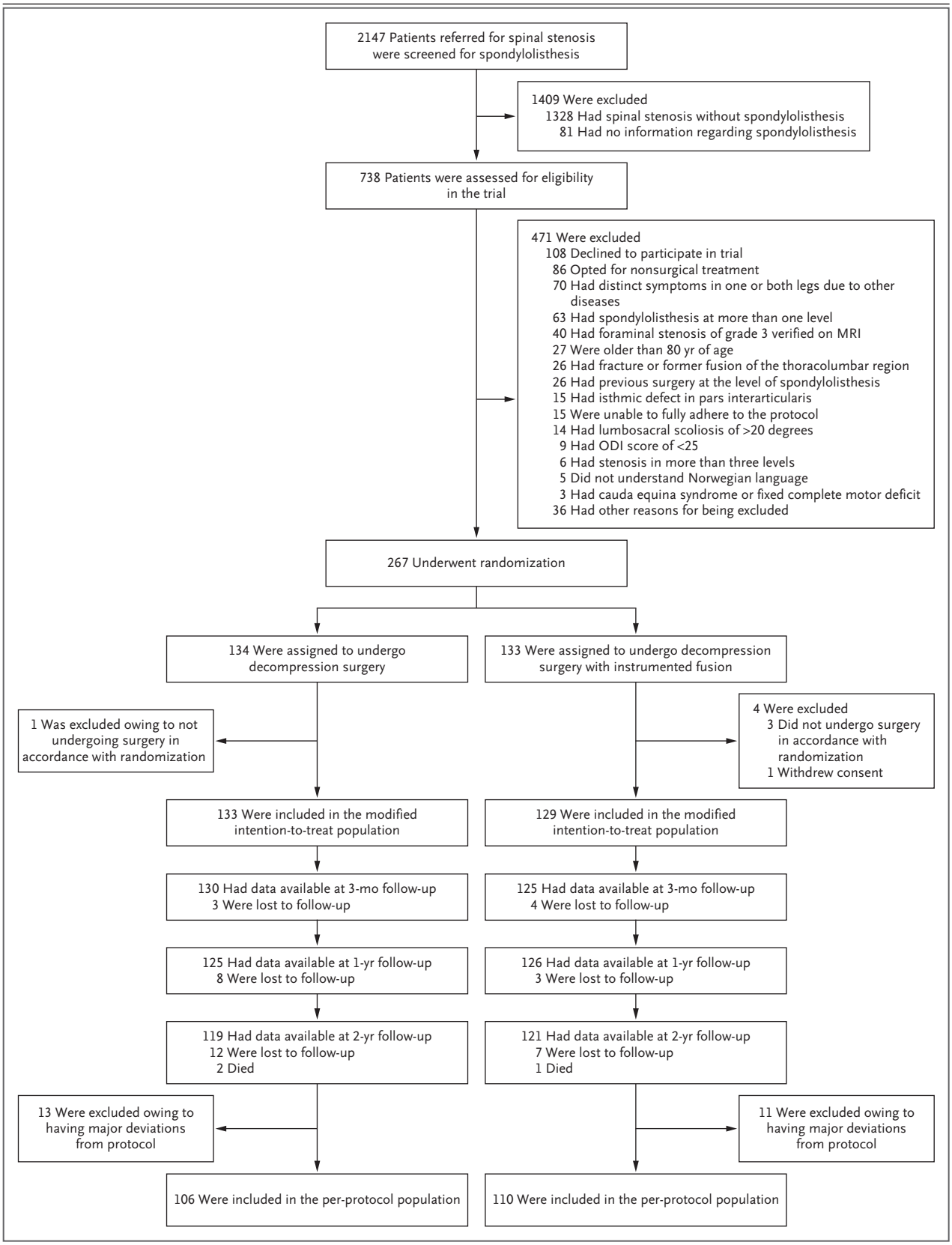

Figure 1 (facing page). Screening, Randomization, and Follow-up of the Trial Patients.

Patients who were referred with the diagnosis of lumbar spinal stenosis were screened for the presence of spondylolisthesis, which was an inclusion criterion. The 738 patients who were assessed for eligibility in the trial could be excluded for more than one reason. The diagnosis of foraminal stenosis of grade 3 was made according to the classification of Lee et al.[26] From April 15, 2014 (the start of inclusion), to August 29, 2015, a score of less than 25 on the Oswestry Disability Index (ODI) was an exclusion criterion (see amendments to the statistical analysis plan, available with the protocol at NEJM.org). The modified intention-to-treat population consisted of all the patients who received the trial treatment in accordance with the randomization assignment and had available data after randomization. The three deaths (two in the decompression-alone group and 1 in the fusion group) were not related to participation in the trial. The per-protocol population consisted of all the patients in the modified intention-to-treat population who did not undergo reoperation during the follow-up period and who had available data on the primary outcome. MRI denotes magnetic resonance imaging.

root in the intervertebral foramen) on MRI, according to the classification of Lee et al.[26]; had previous surgery at the level of spondylolisthesis; or had a former fracture or fusion surgery in the thoracolumbar region. Patients were referred to public orthopedic and neurosurgical clinics and screened for eligibility by surgeons involved in performing the trial operations. Patients who opted to undergo surgery and who fulfilled the eligibility criteria received written and spoken information about the trial. Those who were willing to take part in the trial provided written informed consent before randomization. Information about the surgical departments and their contribution to the enrollment of patients is provided in Section 2 in the Supplementary Appendix.

Patients were randomly assigned in a 1:1 ratio to undergo decompression surgery alone or decompression surgery with instrumented fusion. Randomization was performed within the MedInsight database hosted by the Clinical Trial Unit at Oslo University Hospital (OUS). The sequence was concealed from the investigators, computer-generated, and stratified according to center with the use of permuted blocks. The block sizes varied randomly between four patients and six patients. Information on block size, generation of the randomization sequence, and treatment assignments was unavailable to the investigators. The trial coordinating center at the Research and Communication Unit for Musculoskeletal Health at OUS emailed output information from MedInsight regarding treatment assignments to local trial coordinators who documented the assignments in the patients' records and informed the surgeons.

INTERVENTIONS

For patients who were assigned to undergo decompression without fusion (decompression-alone group), a posterior decompression that preserved the midline structures (supraspinous–interspinous ligament complex) was used. The approach could be bilateral, ipsilateral, or ipsilateral with crossover to the contralateral side. For patients who were assigned to undergo decompression with instrumented fusion (fusion group), an optional technique for posterior decompression (with or without preservation of midline structures) was used, followed by implantation of pedicle screws with rods and bone grafting across the level of spondylolisthesis, with optional use of an intervertebral fusion device. The selection of implants was at the discretion of the surgeons. All participating surgeons routinely performed the procedures used in the trial. Microsurgical decompression (i.e., the use of a microscope or magnifying glasses during decompression) was recommended in both treatment groups. Fusion after 2 years was independently evaluated on computed tomography by two experienced surgeons (the first and eighth authors) and one experienced radiologist (the twelfth author). A conclusion of fusion or not was based on majority ratings.

At admission for surgery, patients completed questionnaires on demographic characteristics, coexisting conditions, and patient-reported outcome measurements. The local trial coordinators collected this information and registered complications during the hospital stay. The surgeons recorded data on diagnosis and treatment. The coordinating center distributed patient follow-up forms for outcomes, which were returned to local trial coordinators at follow-up visits at 3 months and 2 years and by mail to the coordinating center at 1 year. The data were entered into the MedInsight database by the coordinating center, stored at the Clinical Trial Unit at OUS, and were inaccessible to the research group until February 5, 2020. Details of the data storage and the date on which the data were made available to the

Table 1. Characteristics of the Patients at Baseline (Modified Intention-to-Treat Population).*

Characteristic	Decompression-Alone Group (N=133)	Fusion Group (N=129)
Age — yr	66.0±7.4	66.5±7.9
Female sex — no. (%)	92 (69.2)	88 (68.2)
≥3 Yr of higher education — no./total no. (%)	30/129 (23.3)	36/125 (28.8)
Married or has domestic partner — no./total no. (%)	91/129 (70.5)	99/127 (78.0)
Smoker — no./total no. (%)	24/130 (18.5)	21/127 (16.5)
Body-mass index†	27.7±4.4	27.9±4.3
Previous surgery — no./total no. (%)‡	4/130 (3.1)	4/127 (3.1)
Leg pain for >1 yr — no./total no. (%)	91/125 (72.8)	95/127 (74.8)
Back pain for >1 yr — no./total no. (%)	107/130 (82.3)	112/129 (86.8)
Use of analgesics — no./total no. (%)	103/130 (79.2)	107/126 (84.9)
ASA score — no./total no. (%)§		
1	16/129 (12.4)	11/124 (8.9)
2	97/129 (75.2)	88/124 (71.0)
3	16/129 (12.4)	25/124 (20.2)
Coexisting conditions — no. (%)		
Hypertension	46 (34.6)	44 (34.1)
Diabetes	9 (6.8)	9 (7.0)
Cardiovascular disease	23 (17.3)	26 (20.2)
Lung disease	13 (9.8)	10 (7.8)
Rheumatoid disease	5 (3.8)	8 (6.2)
Anxiety or depression	6 (4.5)	7 (5.4)
Other musculoskeletal diseases	13 (9.8)	11 (8.5)
HSCL-25 score¶	1.6±0.4	1.6±0.4
ODI score‖	39.3±14.0	39.4±12.4
Score on ZCQ for symptom severity**	3.3±0.6	3.4±0.6
Score on ZCQ for physical function††	2.5±0.5	2.5±0.5
Score on NRS for leg pain‡‡	6.7±2.1	6.7±1.8
Score on NRS for back pain‡‡	6.7±2.0	6.6±2.0
EQ-5D-3L score§§	0.4±0.3	0.4±0.3

* Plus–minus values are means ±SD. Patients were randomly assigned to undergo decompression surgery (decompression-alone group) or decompression surgery with instrumented fusion (fusion group). The modified intention-to-treat population consisted of all the participants who received the trial treatment in accordance with the randomization assignment and had available data after randomization. Percentages may not total 100 because of rounding.

† The body-mass index is the weight in kilograms divided by the square of the height in meters.

‡ Shown are patients with previous lumbar spine surgery but not at the level of spondylolisthesis.

§ An American Society of Anesthesiologists (ASA) score of 1 indicates no disease, 2 mild systemic disease, and 3 severe systemic disease that is not life-threatening.

¶ The 25-item Hopkins Symptom Checklist (HSCL-25) is a patient-administered questionnaire for the assessment of symptoms of anxiety and depression. Scores range from 1 to 4, with lower scores indicating less severe symptoms.

‖ Scores on the Oswestry Disability Index (ODI) range from 0 to 100, with higher scores indicating more severe disability.

** Scores on the Zurich Claudication Questionnaire (ZCQ) for the assessment of symptom severity range from 1 to 5, with lower scores indicating less symptom severity.

†† Scores on the ZCQ for the assessment of physical function range from 1 to 5, with lower scores indicating less impairment.

‡‡ Scores on the Numerical Rating Scale (NRS) for leg pain and for back pain range from 0 to 10, with lower scores indicating less pain.

§§ The 3-level version of the EuroQol Group 5-Dimension (EQ-5D-3L) questionnaire is a generic instrument to evaluate health-related quality of life that includes the domains of mobility, self-care, usual activity, pain or discomfort, and anxiety or depression. Scores range from –0.59 to 1.0, with higher scores indicating better health-related quality of life.

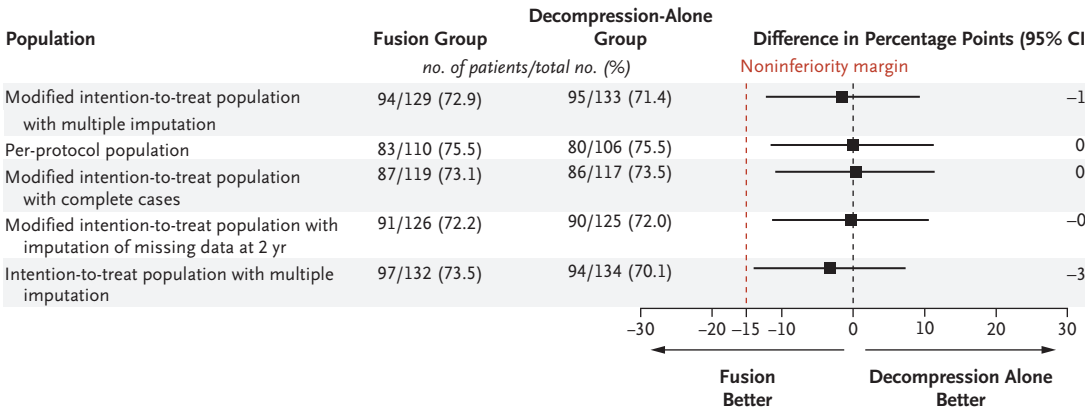

Figure 2. Results of the Primary Outcome.

The number of patients refers to the number who had a reduction of at least 30% in the score on the Oswestry Disability Index (range, 0 to 100, with higher scores indicating more impairment) from baseline to 2-year follow-up (primary outcome). The between-group differences and 95% confidence intervals (CIs) were calculated as decompression alone minus decompression with instrumented fusion. The modified intention-to-treat population consisted of all the patients who received the trial treatment in accordance with the randomization assignment and had available data after randomization. Multiple imputation was performed if data were missing at baseline (4 patients) or at the 2-year follow-up (22 patients). The per-protocol population consisted of all the patients in the modified intention-to-treat population who did not undergo reoperation during the follow-up period and who had available data on the primary outcome. Patients in the modified intention-to-treat population with complete cases had available data on the primary outcome. For patients in the modified intention-to-treat population with imputation of missing data at 2 years, the missing values were replaced by the values at 1 year, when available. A post hoc analysis in the intention-to-treat population included all the patients who underwent randomization, apart from 1 patient who had withdrawn consent before the operation.

research group are provided in Section 5.1 in the Supplementary Appendix. Details regarding imaging assessments are provided in Section 5.3 in the Supplementary Appendix.

OUTCOMES

The primary outcome was a reduction in the score on the Oswestry Disability Index (ODI),[27] version 2.0, of 30% or greater from baseline to 2-year follow-up.[28,29] This index comprises 10 questions with a total score ranging from 0 (no impairment) to 100 (maximum impairment) and has been used and validated for the assessment of treatment outcomes in patients with lumbar degenerative spondylolisthesis.[30] The primary outcome was assessed for noninferiority of decompression alone by 15 percentage points, as described in the Statistical Analysis section. The mean change in the ODI score was a secondary outcome. Other secondary outcomes were the scores on the Zurich Claudication Questionnaire (ZCQ),[31] which assesses symptom severity (range, 1 to 5, with lower scores indicating less severity), functional impairment (range, 1 to 5, with lower scores indicating less impairment), and satisfaction with treatment (range, 1 to 4, with lower scores indicating higher satisfaction); scores on the Numeric Rating Scale (NRS) for leg pain and for back pain (range, 0 to 10, with lower scores indicating less pain), which assessed pain experienced during the past week[32]; and the score on the 3-level version of the EuroQol Group 5-Dimension (EQ-5D-3L) questionnaire (range, −0.59 to 1.0, with higher scores indicating better quality of life).[33] The questionnaires have been translated and validated for a Norwegian population; further information is provided in Section 5.2 in the Supplementary Appendix.

The duration of surgery and length of hospital stay were recorded in order to indicate the use of resources. The incidences of complications and reoperations and the percentage of patients who responded that their condition was "much worse" or "worse than ever" on the Global Perceived Effect scale[34] (7-point Likert scale) were used to evaluate safety and patient-reported deterioration in their condition.

STATISTICAL ANALYSIS

The null hypothesis was that the percentage of patients who had a reduction of at least 30% in

Table 2. Secondary Outcomes (Modified Intention-to-Treat Population).*

Outcome	Mean Value at Baseline (95% CI)	Mean Value at 3 Mo (95% CI)	Mean Value at 2 Yr (95% CI)	Mean Change from Baseline to 2 Yr (95% CI)	Mean Difference in Change, Decompression Alone vs. Fusion (95% CI)
ODI score					
Decompression-alone group	39.2 (36.7 to 41.8)	17.3 (14.9 to 19.8)	18.6 (16.1 to 21.2)	−20.6 (−23.1 to −18.1)	0.7 (−2.8 to 4.3)
Fusion group	39.7 (37.0 to 42.3)	19.6 (17.1 to 22.1)	18.3 (15.8 to 20.9)	−21.3 (−23.8 to −18.8)	
Score on ZCQ for symptom severity					
Decompression-alone group	3.32 (3.20 to 3.45)	2.21 (2.10 to 2.33)	2.32 (2.20 to 2.45)	−1.00 (−1.13 to −0.87)	−0.02 (−0.20 to 0.16)
Fusion group	3.42 (3.29 to 3.54)	2.25 (2.13 to 2.37)	2.44 (2.31 to 2.56)	−0.98 (−1.11 to −0.85)	
Score on ZCQ for physical function					
Decompression-alone group	2.52 (2.42 to 2.63)	1.65 (1.55 to 1.75)	1.68 (1.58 to 1.78)	−0.85 (−0.95 to −0.74)	−0.03 (−0.19 to 0.12)
Fusion group	2.52 (2.41 to 2.62)	1.62 (1.52 to 1.72)	1.71 (1.60 to 1.81)	−0.81 (−0.92 to −0.70)	
Score on ZCQ for patient satisfaction†					
Decompression-alone group	—	1.71 (1.59 to 1.83)	1.78 (1.65 to 1.91)	0.08 (−0.04 to 0.20)	—
Fusion group	—	1.73 (1.61 to 1.85)	1.73 (1.60 to 1.86)	0.00 (−0.12 to 0.12)	
Score on NRS for leg pain					
Decompression-alone group	6.65 (6.22 to 7.09)	2.67 (2.27 to 3.08)	2.81 (2.38 to 3.24)	−3.84 (−4.34 to −3.34)	−0.22 (−0.93 to 0.49)
Fusion group	6.70 (6.27 to 7.14)	2.48 (2.07 to 2.89)	3.08 (2.65 to 3.51)	−3.62 (−4.12 to −3.12)	
Score on NRS for back pain					
Decompression-alone group	6.73 (6.31 to 7.14)	3.38 (2.99 to 3.76)	3.33 (2.92 to 3.74)	−3.39 (−3.87 to −2.92)	−0.45 (−1.13 to 0.23)
Fusion group	6.61 (6.19 to 7.03)	3.25 (2.86 to 3.64)	3.67 (3.25 to 4.08)	−2.94 (−3.42 to −2.46)	
EQ-5D-3L score					
Decompression-alone group	0.44 (0.39 to 0.48)	0.72 (0.67 to 0.76)	0.70 (0.65 to 0.75)	0.26 (0.21 to 0.32)	−0.08 (−0.15 to 0.00)
Fusion group	0.38 (0.33 to 0.43)	0.70 (0.65 to 0.74)	0.72 (0.67 to 0.76)	0.34 (0.29 to 0.39)	

* Shown are estimated mean values and mean changes for continuous repeated outcome measurements. Estimated values are based on linear mixed models. The confidence intervals for differences between groups were not adjusted for multiple comparisons, so no definite conclusions can be made from these data.
† Scores on the ZCQ for the assessment of satisfaction with surgery range from 1 to 4, with lower scores indicating higher satisfaction.

the ODI score (i.e., had a clinically important improvement in functioning) would be at least 15 percentage points lower in the decompression-alone group than in the fusion group. This margin was based on information that decompression alone has some advantages over decompression with instrumented fusion (less complex, less invasive, cheaper, and possibly safer)[5,35] that would justify an "acceptable" loss of efficacy. The margin of –15 percentage points corresponds to a number needed to benefit with fusion of 7 (number needed to treat, $1 \div 0.15 = 6.67$)[36] — that is, at least 7 patients would need instrumented fusion for 1 additional patient to have a clinically important improvement in functioning. To reject the null hypothesis and show the noninferiority of decompression alone, the participation of 232 patients was required to be 80% certain (i.e., power of 80%) that the lower limit of a 95% two-sided confidence interval would exclude a difference in the frequency of a clinically important improvement in functioning of more than 15 percentage points.[37] To account for a possible dropout rate of 10%, 128 patients were required in each group.

The main analysis was conducted in the modified intention-to-treat population, which consisted of all the patients who received the trial treatment in accordance with the randomization assignment and had available data after randomization. For the modified intention-to-treat analyses, missing data on the primary and secondary outcomes were imputed with the use of 50 imputations by chained equations. A list of the 29 variables that were used in the imputation model and other details pertaining to the imputations are provided in Section 5.4.1 in the Supplementary Appendix. The per-protocol population consisted of all the patients in the modified intention-to-treat population who did not undergo reoperation during the follow-up period and had available data on the primary outcome. To declare noninferiority of decompression alone, noninferiority was to be shown for the primary outcome in both the modified intention-to-treat population with multiple imputation of missing data and in the per-protocol population. Three sensitivity analyses were performed: one in the modified intention-to-treat population with complete cases (without imputation for missing data); one in which missing values at 2 years were replaced by values at 1 year, when available; and a post hoc intention-to-treat analysis in which the sample was defined by all the patients who underwent randomization, regardless of treatment adherence. Analyses of the percentage of patients with a clinically important improvement from inclusion to 2 years after surgery as assessed by the ZCQ, NRS for leg pain, and NRS for back pain were performed.[28,31]

The primary outcome and all categorical secondary outcomes were analyzed with the use of Newcombe hybrid score confidence intervals.[38] We analyzed all repeated-measures continuous outcomes (scores on the ODI, ZCQ, NRS for leg pain, NRS for back pain, and EQ-5D-3L) with linear mixed models. The models contained fixed effects for treatment, time, interaction between treatment and time, and trial center. Time was modeled as piecewise linear with a knot at 3 months. A random intercept at the patient level was used. On the basis of the fitted models, we estimated mean values with 95% confidence intervals at baseline (inclusion), 3 months, and 2 years after surgery and the change from baseline to 2 years within each treatment group. We also estimated the between-group difference (with 95% confidence interval) in change from baseline to 2 years.

The mean between-group differences in the duration of surgery and length of hospital stay were analyzed with 95% confidence intervals that were based on the t distribution with adjustment for unequal variance. We assessed the assumption of normally distributed data with visual inspection of histograms and descriptive statistics. No major deviations from normality were observed. There was no method for adjustment of confidence intervals for multiple comparisons of secondary outcomes, and these results are presented as point estimates with unadjusted confidence intervals from which no definite conclusions can be made. The analyses were performed with the use of Stata/SE software, version 15.0 (StataCorp).

RESULTS

PATIENT CHARACTERISTICS

From February 12, 2014, to December 18, 2017, a total of 267 patients from 16 surgical departments were enrolled; 134 were assigned to undergo decompression alone and 133 to undergo decompression with instrumented fusion, and 240 patients (89.9%) had available data at the

Table 3. Additional Secondary Outcomes (Modified Intention-to-Treat Population).*

Outcome	Decompression-Alone Group (N=133)	Fusion Group (N=129)	Difference (95% CI)†
Duration of surgery — minutes	104±4	174±6	−70 (−84 to −55)
Length of hospital stay — days	3.3±0.2	5.0±0.2	−1.8 (−2.4 to −1.2)
Clinically important improvement			
As assessed by the ZCQ — no. (%)‡	99 (74.4)	99 (76.7)	−2.3 (−12.6 to 8.1)
As assessed by the NRS for leg pain — no. (%)§	85 (63.9)	83 (64.3)	−0.4 (−11.9 to 11.1)
As assessed by the NRS for back pain — no. (%)§	88 (66.2)	75 (58.1)	8.0 (−3.7 to 19.5)
As assessed by the ODI, according to decompression technique — no./total no. (%)¶			
Midline structures preserved	91/129 (70.5)	40/58 (69.0)	—
Midline structures not preserved	3/3 (100)	53/70 (75.7)	—
Complications			
Incidental dural tear — no./total no. (%)	7/132 (5.3)	17/128 (13.3)	−8.0 (−15.5 to −0.9)
Blood loss during surgery — ml	141±134	429±278	−292 (−348 to −235)
Blood transfusion — no./total no. (%)	0/132	4/128 (3.1)	−3.1 (−7.8 to 0.3)
Surgery on the wrong side or level — no./total no. (%)‖	1/132 (0.8)	1/128 (0.8)	0.0 (−3.6 to 3.5)
Hematoma resulting in reoperation during hospital stay — no./total no. (%)	1/132 (0.8)	1/128 (0.8)	0.0 (−3.6 to 3.5)
Wound infection — no./total no. (%)			
During hospital stay	0/132	0/128	0.0 (−2.9 to 2.8)
From hospital discharge to 3 mo**	3/129 (2.3)	6/125 (4.8)	−2.5 (−8.0 to 2.5)
Cardiovascular complications — no./total no. (%)††			
During hospital stay	3/132 (2.3)	0/128	2.3 (−1.0 to 6.5)
From hospital discharge to 3 mo	1/129 (0.8)	0/125	0.8 (−2.3 to 4.3)
Venous thromboembolism — no./total no. (%)			
During hospital stay	0/132	0/128	0.0 (−2.9 to 2.8)
From hospital discharge to 3 mo	0/129	0/125	0.0 (−3.9 to 2.9)
Urologic complication — no./total no. (%)			
During hospital stay‡‡	4/132 (3.0)	6/128 (4.7)	−1.7 (−7.1 to 3.5)
From hospital discharge to 3 mo§§	2/129 (1.6)	5/125 (4.0)	−2.5 (−7.6 to 2.1)
Respiratory complication — no./total no. (%)¶¶			
During hospital stay	0/132	2/128 (1.6)	−1.6 (−5.5 to 1.5)
From hospital discharge to 3 mo	1/129 (0.8)	0/125	0.8 (−2.3 to 4.3)
Deterioration			
Neurologic deterioration — no./total no. (%)	15/132 (11.4)	24/128 (18.8)	−7.4 (−16.2 to 1.4)
During hospital stay‖‖	1/132 (0.8)	2/128 (1.6)	−0.8 (−4.8 to 2.8)
From hospital discharge to 3 mo***	3/129 (2.3)	7/125 (5.6)	−3.3 (−9.0 to 1.9)
From 3 mo to 2 yr†††	11/120 (9.2)	15/121 (12.4)	−3.2 (−11.3 to 4.8)
Substantial deterioration according to GPE scale — no./total no. (%)‡‡‡	7/119 (5.9)	6/120 (5.0)	0.9 (−5.4 to 7.2)
Reoperation — no./total no. (%)§§§	15/120 (12.5)	11/121 (9.1)	3.4 (−4.6 to 11.5)

Table 3. (Continued.)

* Plus–minus values are means ±SD. Confidence intervals for differences between groups were not adjusted for multiple comparisons, so no definite conclusions can be made from these data.
† Between-group differences are in percentage points except for duration of surgery, length of hospital stay, and blood loss during surgery.
‡ A clinically important improvement was defined as meeting at least two of three predefined criteria at the 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 in the score on the physical-function scale of 0.42 or more, and a score on the patient-satisfaction scale of 2.42 or less.[31]
§ The predefined criteria for a clinically important improvement were a decrease from baseline of 40% or more in the score on the NRS for leg pain and a decrease from baseline of 33% or more in the score on the NRS for back pain, both at 2-year follow-up.[28]
¶ A clinically important improvement was defined as a reduction in the ODI score of 30% or greater from baseline to 2-year follow-up.
‖ The frequency of surgeries on the wrong side or level was somewhat higher than that reported in the literature.
** All 3 wound infections in the decompression-alone group were reported to be superficial. Of 6 wound infections in the fusion group, 3 were reported to be superficial and 3 were deep (subfascial).
†† Of 3 cardiovascular complications during the hospital stay, 2 were reported to be atrial fibrillation and 1 an ischemic stroke. The cardiovascular complication from hospital discharge to 3 months was reported to be an ischemic stroke.
‡‡ Of 4 urologic complications in the decompression-alone group, 2 were reported to be urinary retention and 2 were not specified. Of 6 urologic complications in the fusion group, 3 were reported to be urinary retention and 3 were not specified.
§§ The 2 urologic complications in the decompression-alone group were not specified. Of 5 urologic complications in the fusion group, 1 was reported to be an upper urinary tract infection and 4 were not specified.
¶¶ The 3 respiratory complications were reported to be pneumonia.
‖‖ The 1 neurologic deterioration in the decompression-alone group was reported to be a sensory deficit. Of 2 neurologic deteriorations in the fusion group, 1 was reported to be a motor deficit and 1 a combined sensory and motor deficit.
*** Of 3 neurologic deteriorations in the decompression-alone group, 1 was reported to be a sensory deficit, 1 a motor deficit, and 1 a combined sensory and motor deficit. Of 7 neurologic deteriorations in the fusion group, 5 were reported to be sensory deficits, 1 a motor deficit, and 1 a combined sensory and motor deficit.
††† Of 11 neurologic deteriorations in the decompression-alone group, 4 were reported to be sensory deficits, 4 motor deficits, and 3 combined sensory and motor deficits. Of 15 neurologic deteriorations in the fusion group, 6 were reported to be sensory deficits, 3 motor deficits, and 6 combined sensory and motor deficits.
‡‡‡ Shown are patients who responded that their condition was "much worse" or "worse than ever" on the Global Perceived Effect (GPE) scale (7-point Likert scale).
§§§ Shown are patients who underwent one or more subsequent operations from the time of the primary operation to 2-year follow-up.

2-year follow-up. Five patients were not included in the modified intention-to-treat population: 1 withdrew consent before randomization, and 4 were not treated according to the randomization assignment (Fig. 1). Of 46 patients who were not included in the per-protocol population, 26 were excluded owing to a reoperation (15 patients in the decompression-alone group and 11 in the fusion group). Table 1 shows the demographic and clinical characteristics of the patients at baseline, which were similar in the two groups. Approximately 75% of the patients in each group had had leg pain for more than a year, and more than 80% in each group had had back pain for more than a year. Table S3 in the Supplementary Appendix shows surgical and imaging data, including the assessment of fusion, which was achieved with certainty in 86 of 100 patients (86.0%) and was uncertain or absent in the remaining 14 patients (14.0%) with available imaging at 2 years. Figure S1 shows the distribution of the observed scores on the continuous repeated outcome measurements. The ODI scores were imputed for 26 patients (15 in the decompression-alone group and 11 in the fusion group) who had missing data; of these patients, 4 had missing data at baseline.

PRIMARY OUTCOME

In the modified intention-to-treat population, 95 of 133 patients (71.4%) in the decompression-alone group and 94 of 129 patients (72.9%) in the fusion group had a reduction of at least 30% in the ODI score (difference, –1.4 percentage points; 95% confidence interval [CI], –12.2 to 9.4). In the per-protocol population, 80 of 106 patients (75.5%) and 83 of 110 patients (75.5%), respectively, had a reduction of at least 30% in the ODI score (difference, 0.0 percentage points; 95% CI, –11.4 to 11.4). For both analyses, the

lower boundary of the 95% confidence interval for the between-group difference was within the noninferiority margin of –15 percentage points. Results of sensitivity analyses of the primary outcome, including the post hoc intention-to-treat analysis, were generally in the same direction as those of the primary analysis (Fig. 2).

SECONDARY OUTCOMES

Results of the secondary outcomes, with confidence intervals that were not adjusted for multiple comparisons and from which no definite conclusions can be drawn, are shown in Tables 2 and 3. The mean change in the ODI score from baseline to 2-year follow-up was –20.6 in the decompression-alone group and –21.3 in the fusion group (mean difference, 0.7; 95% CI, –2.8 to 4.3). Similar small between-group differences in point estimates of mean change were found for scores on the ZCQ, NRS for leg pain, NRS for back pain, and EQ-5D-3L. Results of analyses of the percentage of patients with a clinically important improvement from inclusion to 2 years after surgery as assessed by the ZCQ, NRS for leg pain, and NRS for back pain were generally in the same direction as those of the analysis of the primary outcome.

The mean duration of surgery was 104 minutes in the decompression-alone group and 174 minutes in the fusion group (mean difference, –69 minutes; 95% CI, –83 to –56); the mean length of hospital stay was 3.3 days and 5.0 days, respectively (mean difference, –1.8 days; 95% CI, –2.4 to –1.2); dural tears occurred in 7 of 132 patients (5.3%) and 17 of 128 patients (13.3%) (difference, –8.0 percentage points; 95% CI, –15.5 to –0.9); and 7 of 119 patients (5.9%) and 6 of 120 patients (5.0%) responded that their condition was "much worse" or "worse than ever" on the Global Perceived Effect scale (difference, 0.9 percentage points; 95% CI, –5.4 to 7.2). Table S4 shows additional responses on this scale. During the 2-year follow-up period, reoperation was performed in 15 of 120 patients (12.5%) in the decompression-alone group and in 11 of 121 patients (9.1%) in the fusion group (difference, 3.4 percentage points; 95% CI, –4.6 to 11.5).

DISCUSSION

In this randomized trial involving 267 patients with lumbar spinal stenosis and degenerative spondylolisthesis, surgery with decompression was noninferior to surgery with decompression and instrumented fusion with respect to the percentage of patients who had a reduction of at least 30% in the ODI score at 2 years after surgery within a margin of –15 percentage points. Results for secondary outcomes with respect to pain, disability, symptom severity, functional status, and satisfaction with treatment were generally in the same direction as those for the primary outcome.

Two previous trials, both with a superiority design, have challenged the widespread use of instrumented fusion in the surgical treatment of degenerative spondylolisthesis. Our results were in accordance with the findings of one of these trials, which involved 133 patients with spondylolisthesis at one or two levels but did not include information about dynamic motion.[11] The second trial, which involved 66 patients with single-level spondylolisthesis without dynamic instability, showed less improvement in the physical-component summary score of the Medical Outcomes Study 36-Item Short-Form Health Survey and more reoperations in the decompression-alone group than in the fusion group.[12] The entry criteria and design of both trials differed from those of our trial.

Spine surgeons may presume that slippage and dynamic instability at the level of spondylolisthesis are better treated with fusion.[2] In our trial, approximately 20% of the patients had slippage of at least 3 mm, or at least 10 degrees of angulation, as assessed by dynamic standing radiographs. Nevertheless, the incidence of reoperation in our trial was lower than the incidence in an aforementioned trial involving patients with no instability.[12] We hypothesize that these differences in outcomes may be the result of differences in follow-up time and variations in medical practice. The present trial was not powered to compare the incidence of reoperation between the two treatment groups, and we cannot rule out the possibility that decompression alone may require a subsequent fusion. However, patients who are treated with decompression and fusion may also require further secondary surgeries, such as hardware removal or surgery at the adjacent lumbar level.[39] Patients and surgeons may weigh a potentially higher risk of reoperation against the complexity and costs of surgical alternatives.[5,35]

In the present trial, 38% of instrumented fu-

sions were supported by an interbody cage, as compared with 17%, 6%, and 0% in somewhat similar trials.[3,11,12] Owing to a lack of evidence for the superiority of one fusion method over the other[2,12,40] and the high frequency of successful fusion as assessed by computed tomography in our trial, we consider the risk of the fusion methods having biased our results to be low.

This trial has limitations. The eligibility criteria limit the generalizability of the results. We did not include patients with degenerative scoliosis, those with radicular pain related to extensive foraminal stenosis, those who had previous surgery at the level of spondylolisthesis, and those with spondylolisthesis at more than one level. For these patients, fusion surgery is considered established practice. Conclusions cannot be drawn from data regarding subgroups, including the approximately 20% of patients who had dynamic instability at the level of spondylolisthesis. Well-powered, high-quality studies are warranted to investigate whether patient characteristics and radiologic variables can predict the appropriate treatment for subgroups of patients. This was an open-label trial, and only the data analyst but not the outcome assessors were unaware of the treatment-group assignments. An evidence-based margin of noninferiority for this research question does not exist, and we chose –15 percentage points empirically. There was possible selection bias in the per-protocol analysis because of the exclusion of patients who had undergone reoperation, which could have favored the decompression-alone group. Because these patients were included in the modified intention-to-treat analysis and the three sensitivity analyses of the primary outcome, we consider the conclusion of noninferiority to be valid.

In this multicenter, single-country, randomized trial involving patients with degenerative single-level lumbar spondylolisthesis, decompression alone was noninferior to decompression with instrumented fusion at 2 years of follow-up. Reoperation occurred somewhat more often in the decompression-alone group.

Supported by grants from the Western Norway Regional Health Authority and the Møre and Romsdal Hospital Trust. The funding authorities were not involved in the trial design, collection or analysis of data, interpretation of the results, or drafting of the manuscript.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

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, Anne-Charlotte Fosse Haug, Elisabeth Lilleholdt Muller, Kjartan Krogedal, Bettina Timenes, Vidar Opland, Merete Finjarn, Håvard Furunes, 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 K. Ebbs, Marie Skovli Pettersen, and Ørjan Nesse Vigdal) for meticulous work with data collection; Eira K. Ebbs (FORMI) for providing linguistic assistance in writing the manuscript; and Alf I. Hellevik (Hellevik Studio) for graphic design of earlier versions of Figures 1 and 2.

APPENDIX

The authors' full names and academic degrees are as follows: Mr. Ivar M. Austevoll, M.D., Ph.D., Erland Hermansen, M.D., Ph.D., Morten W. Fagerland, Ph.D., Kjersti Storheim, P.T., Ph.D., Jens I. Brox, M.D., Ph.D., Tore Solberg, M.D., Ph.D., Frode Rekeland, M.D., Eric Franssen, M.D., Clemens Weber, M.D., Ph.D., Helena Brisby, M.D., Ph.D., Oliver Grundnes, M.D., Ph.D., Knut R.H. Algaard, M.D., Tordis Böker, M.D., Hasan Banitalebi, M.D., Kari Indrekvam, M.D., Ph.D., and Christian Hellum, M.D., Ph.D.

The authors' affiliations are as follows: Kysthospitalet in Hagevik, Orthopedic Department, Haukeland University Hospital (I.M.A., E.H., F.R., K.I.), and the Department of Clinical Medicine, University of Bergen (E.H., F.R., K.I.), Bergen, Møre and Romsdal Hospital Trust, Ålesund Hospital, Orthopedic Department, Ålesund (E.H.), the Oslo Center for Biostatistics and Epidemiology, Research Support Services (M.W.F.), the Research and Communication Unit for Musculoskeletal Health, Division of Clinical Neuroscience (K.S.), and the Department of Physical Medicine and Rehabilitation (J.I.B.), Oslo University Hospital, the Department of Physiotherapy, Oslo Metropolitan University (K.S.), the Medical Faculty (J.I.B.), the Division of Radiology and Nuclear Medicine, Institute of Clinical Medicine, Faculty of Medicine (T.B.), and the Institute of Clinical Medicine (H. Banitalebi), University of Oslo, Akershus University Hospital, Orthopedic Department (O.G.), Radiology, Unilabs Radiology (K.R.H.A.), and the Department of Radiology and Nuclear Medicine (T.B.) and the Division of Orthopedic Surgery (C.H.), Oslo University Hospital Ullevål, Oslo, the Institute of Clinical Medicine, University of Tromsø–the Arctic University of Norway, and the Norwegian Registry for Spine Surgery, University Hospital of North Norway, Tromsø (T.S.), the Orthopedic Department (E.F.) and the Department of Neurosurgery (C.W.), Stavanger University Hospital, and the Department of Quality and Health Technology, Faculty of Health Sciences, University of Stavanger (C.W.), Stavanger, and the Department of Diagnostic Imaging, Akershus University Hospital, Lorenskog (H. Banitalebi) — all in Norway; and the Spine Surgery Team, Department of Orthopedics, Sahlgrenska University Hospital, and the Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden (H. Brisby).

REFERENCES

1. Fitzgerald JA, Newman PH. Degenerative spondylolisthesis. J Bone Joint Surg Br 1976;58:184-92.
2. Matz PG, Meagher RJ, Lamer T, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. Spine J 2016;16:439-48.
3. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007;356:2257-70.
4. 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;39:1584-9.
5. Kim CH, Chung CK, Choi Y, et al. Increased proportion of fusion surgery for degenerative lumbar spondylolisthesis and changes in reoperation rate: a nationwide cohort study with a minimum 5-year follow-up. Spine (Phila Pa 1976) 2019;44:346-54.
6. Martin BI, Mirza SK, Spina N, Spiker WR, Lawrence B, Brodke DS. Trends in lumbar fusion procedure rates and associated hospital costs for degenerative spinal diseases in the United States, 2004 to 2015. Spine (Phila Pa 1976) 2019;44:369-76.
7. Weiss AJ, Elixhauser A, Andrews RM. Characteristics of operating room procedures in U.S. hospitals, 2011. HCUP statistical brief no. 170. Rockville, MD: Agency for Healthcare Research and Quality, February 2014.
8. Lønne G, Fritzell P, Hägg O, et al. Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries. Spine J 2019;19:41-9.
9. Austevoll IM, Gjestad R, Solberg T, et al. Comparative effectiveness of microdecompression alone vs decompression plus instrumented fusion in lumbar degenerative spondylolisthesis. JAMA Netw Open 2020;3(9):e2015015.
10. Overdevest GM, Moojen WA, Arts MP, Vleggeert-Lankamp CLAM, Jacobs WCH, Peul WC. Management of lumbar spinal stenosis: a survey among Dutch spine surgeons. Acta Neurochir (Wien) 2014;156:2139-45.
11. Försth P, Ólafsson G, Carlsson T, et al. A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. N Engl J Med 2016;374:1413-23.
12. Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. N Engl J Med 2016;374:1424-34.
13. Peul WC, Moojen WA. Fusion for lumbar spinal stenosis — safeguard or superfluous surgical implant? N Engl J Med 2016;374:1478-9.
14. Försth P, Michaëlsson K, Sandén B. More on fusion surgery for lumbar spinal stenosis. N Engl J Med 2016;375:1806-7.
15. Pisano AJ, Butler JS, Sebastian A, Wagner SC, Wanderman N. Does surgically managed grade I degenerative lumbar spondylolisthesis require fusion? Clin Spine Surg 2019;32:133-6.
16. Dijkerman ML, Overdevest GM, Moojen WA, Vleggeert-Lankamp CLA. Decompression with or without concomitant fusion in lumbar stenosis due to degenerative spondylolisthesis: a systematic review. Eur Spine J 2018;27:1629-43.
17. Liang H-F, Liu S-H, Chen Z-X, Fei Q-M. Decompression plus fusion versus decompression alone for degenerative lumbar spondylolisthesis: a systematic review and meta-analysis. Eur Spine J 2017;26:3084-95.
18. Samuel AM, Moore HG, Cunningham ME. Treatment for degenerative lumbar spondylolisthesis: current concepts and new evidence. Curr Rev Musculoskelet Med 2017;10:521-9.
19. Major surgical question has no answer. Back Lett 2019;34:39-46.
20. Pearson A. 2018 Year in review. 2018 (https://journals.lww.com/spinejournal/blog/SpineBlog/pages/post.aspx?PostID=570).
21. Mauri L, D'Agostino RB Sr. Challenges in the design and interpretation of noninferiority trials. N Engl J Med 2017;377:1357-67.
22. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH harmonised guideline: integrated addendum to ICH E6(R1): guideline for good clinical practice: E6(R2). November 2016 (https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf).
23. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA 2012;308:2594-604.
24. 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;20:7.
25. ClinicalTrials.gov. The NORDSTEN study — degenerative spondylolisthesis (NORDSTEN/DS). January 27, 2020 (https://clinicaltrials.gov/ct2/show/NCT02051374?term=austevoll&draw=2&rank=2).
26. Lee S, Lee JW, Yeom JS, et al. A practical MRI grading system for lumbar foraminal stenosis. AJR Am J Roentgenol 2010;194:1095-8.
27. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry Low Back Pain Disability Questionnaire. Physiotherapy 1980;66:271-3.
28. Austevoll IM, Gjestad R, Grotle M, et al. Follow-up score, change score or percentage change score for determining clinical important outcome following surgery? An observational study from the Norwegian registry for spine surgery evaluating patient reported outcome measures in lumbar spinal stenosis and lumbar degenerative spondylolisthesis. BMC Musculoskelet Disord 2019;20:31.
29. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 2008;33:90-4.
30. Watters WC III, Bono CM, Gilbert TJ, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. Spine J 2009;9:609-14.
31. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. Spine (Phila Pa 1976) 2006;31:1276-80.
32. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. J Pain Symptom Manage 2011;41:1073-93.
33. EuroQol Group. EuroQol — a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
34. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010;63(7):760.e1-766.e1.
35. 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. JAMA 2010;303:1259-65.
36. Katz N, Paillard FC, Van Inwegen R. A review of the use of the number needed to treat to evaluate the efficacy of analgesics. J Pain 2015;16:116-23.
37. Blackwelder WC, Chang MA. Sample size graphs for "proving the null hypothesis." Control Clin Trials 1984;5:97-105.
38. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res 2015;24:224-54.
39. Wang T, Ding W. Risk factors for adjacent segment degeneration after posterior lumbar fusion surgery in treatment for degenerative lumbar disorders: a meta-analysis. J Orthop Surg Res 2020;15:582.
40. Levin JM, Tanenbaum JE, Steinmetz MP, Mroz TE, Overley SC. Posterolateral fusion (PLF) versus transforaminal lumbar interbody fusion (TLIF) for spondylolisthesis: a systematic review and meta-analysis. Spine J 2018;18:1088-98.