6-1-2009

Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis

James N. Weinstein
Dartmouth Medical School

Jon D. Lurie
Dartmouth Medical School

Tor D. Tosteson
Dartmouth Medical School

Wenyan Zhao
Dartmouth Medical School

Emily A. Blood
Dartmouth Medical School

See next page for additional authors

Recommended Citation

Authors

James N. Weinstein, Jon D. Lurie, Tor D. Tosteson, Wenyan Zhao, Emily A. Blood, Anna N.A. Tosteson, Nancy Birkmeyer, Harry Herkowitz, Michael Longley, Lawrence Lenke, Sanford Emery, and Serena S. Hu

Surgical Compared with Nonoperative Treatment for Lumbar Degenerative Spondylolisthesis

Four-Year Results in the Spine Patient Outcomes Research Trial (SPORT) Randomized and Observational Cohorts

By James N. Weinstein, DO, MS, Jon D. Lurie, MD, MS, Tor D. Tosteson, ScD, Wenyan Zhao, MS, Emily A. Blood, MS,
Anna N.A. Tosteson, ScD, Nancy Birkmeyer, PhD, Harry Herkowitz, MD, Michael Longley, MD,
Lawrence Lenke, MD, Sanford Emery, MD, and Serena S. Hu, MD

*Investigation performed at Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, William Beaumont Hospital, Royal Oak, Michigan,
Nebraska Foundation for Spinal Research, Omaha, Nebraska, Washington University, St. Louis, Missouri, Case Western Reserve/University Hospitals of
Cleveland, Cleveland, Ohio, University of California, San Francisco, California, Emory University, Atlanta, Georgia, Hospital for Joint Diseases, New York,
NY, The Hospital for Special Surgery, New York, NY, Rothman Institute at Thomas Jefferson University, Philadelphia, Pennsylvania, Kaiser-Permanente,
Oakland, California, Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois, and Maine Spine and Rehabilitation, Scarborough, Maine*

Background: The management of degenerative spondylolisthesis associated with spinal stenosis remains controversial. Surgery is widely used and has recently been shown to be more effective than nonoperative treatment when the results were followed over two years. Questions remain regarding the long-term effects of surgical treatment compared with those of nonoperative treatment.

Methods: Surgical candidates from thirteen centers with symptoms of at least twelve weeks' duration as well as confirmatory imaging showing degenerative spondylolisthesis with spinal stenosis were offered enrollment in a randomized cohort or observational cohort. Treatment consisted of standard decompressive laminectomy (with or without fusion) or usual nonoperative care. Primary outcome measures were the Short Form-36 (SF-36) bodily pain and physical function scores and the modified Oswestry Disability Index at six weeks, three months, six months, and yearly up to four years.

Results: In the randomized cohort (304 patients enrolled), 66% of those randomized to receive surgery received it by four years whereas 54% of those randomized to receive nonoperative care received surgery by four years. In the observational cohort (303 patients enrolled), 97% of those who chose surgery received it whereas 33% of those who chose nonoperative care eventually received surgery. The intent-to-treat analysis of the randomized cohort, which was limited by nonadherence to the assigned treatment, showed no significant differences in treatment outcomes between the operative and nonoperative groups at three or four years. An as-treated analysis combining the randomized and observational cohorts that adjusted for potential confounders demonstrated that the clinically relevant advantages of surgery that had been previously reported through two years were maintained at four years, with treatment effects of 15.3 (95% confidence interval, 11 to 19.7) for bodily pain, 18.9 (95% confidence interval, 14.8 to 23) for physical function, and −14.3 (95% confidence interval, −17.5 to −11.1) for the Oswestry Disability Index. Early advantages (at two years) of surgical treatment in terms of the secondary measures of bothersomeness of back and leg symptoms, overall satisfaction with current symptoms, and self-rated progress were also maintained at four years.

eC A commentary is available with the electronic versions of this article, on our web site (www.jbjs.org) and on our quarterly CD-ROM/DVD (call our subscription department, at 781-449-9780, to order the CD-ROM or DVD).

Conclusions: Compared with patients who are treated nonoperatively, patients in whom degenerative spondylolisthesis and associated spinal stenosis are treated surgically maintain substantially greater pain relief and improvement in function for four years.

Level of Evidence: Therapeutic Level II. See Instructions to Authors for a complete description of levels of evidence.

Lumbar decompression with fusion is commonly performed in the United States for patients with back and leg symptoms due to spinal stenosis with or without degenerative spondylolisthesis[1]. There have been studies comparing surgery with nonoperative treatment for spinal stenosis, but they have included mixed groups of patients with and without degenerative spondylolisthesis[2-4]. There have been several studies comparing surgical techniques in cohorts with degenerative spondylolisthesis, but they have had small sample sizes and limited geographic participation as well as a lack of nonoperative controls and validated outcome measures[5-7].

The special methodological challenges of surgical trials (e.g., compliance with treatment[4-12]) were addressed by the design of the Spine Patient Outcomes Research Trial (SPORT), in which a randomized cohort and a concurrent observational cohort were selected with the identical criteria and evaluated with the same outcomes assessment[8-10]. In the SPORT, as-treated comparisons with careful control for potentially confounding baseline factors showed that patients treated surgically for spinal stenosis and associated degenerative spondylolisthesis had substantially greater pain relief and improvement in function during a two-year period than did patients treated nonoperatively. In prior studies of spinal stenosis, the improvement observed early after surgery was reported to deteriorate and the advantage of surgery compared with nonoperative treatment was found to have narrowed with longer-term follow-up[3-11]. Therefore, in this study, we assessed the stability of pain and functional outcomes out to four years for patients with degenerative spondylolisthesis and associated spinal stenosis.

Materials and Methods

Study Design

The SPORT was conducted at thirteen medical centers with multidisciplinary spine practices in eleven states in the United States. Institutional review board approval was obtained at each center. The trial was registered with ClinicalTrials.gov (NCT00000411). The SPORT included both a randomized cohort and a concurrent observational cohort of patients who declined randomization. This design improves the generalizability of the findings[12]. Additional background information is available in previous publications[8-10,13,14].

Patient Population

All patients had neurogenic claudication or radicular leg pain with associated neurological signs, spinal stenosis seen on cross-sectional imaging, degenerative spondylolisthesis seen on standing lateral radiographs, symptoms that had persisted for at least twelve weeks, and physician confirmation that they were a surgical candidate. Patients with adjacent levels of stenosis were eligible, but those with spondylolysis and isthmic spondylolisthesis were not. Pre-enrollment nonoperative care was not specified but included physical therapy (68%), epidural injections (55%), chiropractic care (25%), anti-inflammatory medications (63%), and opioid analgesics (30%). Enrollment began in March 2000 and ended in February 2005.

Study Interventions

The protocol surgery consisted of a standard posterior decompressive laminectomy with or without bilateral single-level fusion (autogenous iliac crest bone-grafting with or without posterior pedicle screw instrumentation)[8]. The nonoperative protocol was "usual recommended care," which includes, at least, active physical therapy, education and counseling with instructions regarding home exercise, and nonsteroidal anti-inflammatory drugs if the patient can tolerate them[8,15].

Study Measures

The primary end points were the Short Form-36 (SF-36) bodily pain and physical function scores[16-19] and the American Academy of Orthopaedic Surgeons MODEMS (Musculoskeletal Outcomes Data Evaluation and Management System) version of the Oswestry Disability Index[20] measured at six weeks, three months, six months, and yearly up to four years. If surgery was delayed for more than six weeks, additional follow-up data were obtained at six weeks and three months postoperatively. Secondary outcomes included patient self-reported improvement, satisfaction with current symptoms and with care[21], the Stenosis Bothersomeness Index[2,22], the Low Back Pain Bothersomeness Scale[2], and the Leg Pain Bothersomeness Scale[2]. The treatment effect was defined as the difference in the mean changes, as compared with baseline, between the surgical and nonoperative groups (the difference in the difference).

SF-36 scores range from 0 to 100 points, with higher scores indicating less severe symptoms; the Oswestry Disability Index ranges from 0 to 100 points, with lower scores indicating less severe symptoms; the Stenosis Bothersomeness Index ranges from 0 to 24 points, with lower scores indicating less severe symptoms; and the Low Back Pain and Leg Pain Bothersomeness Scales range from 0 to 6 points, with lower scores indicating less severe symptoms.
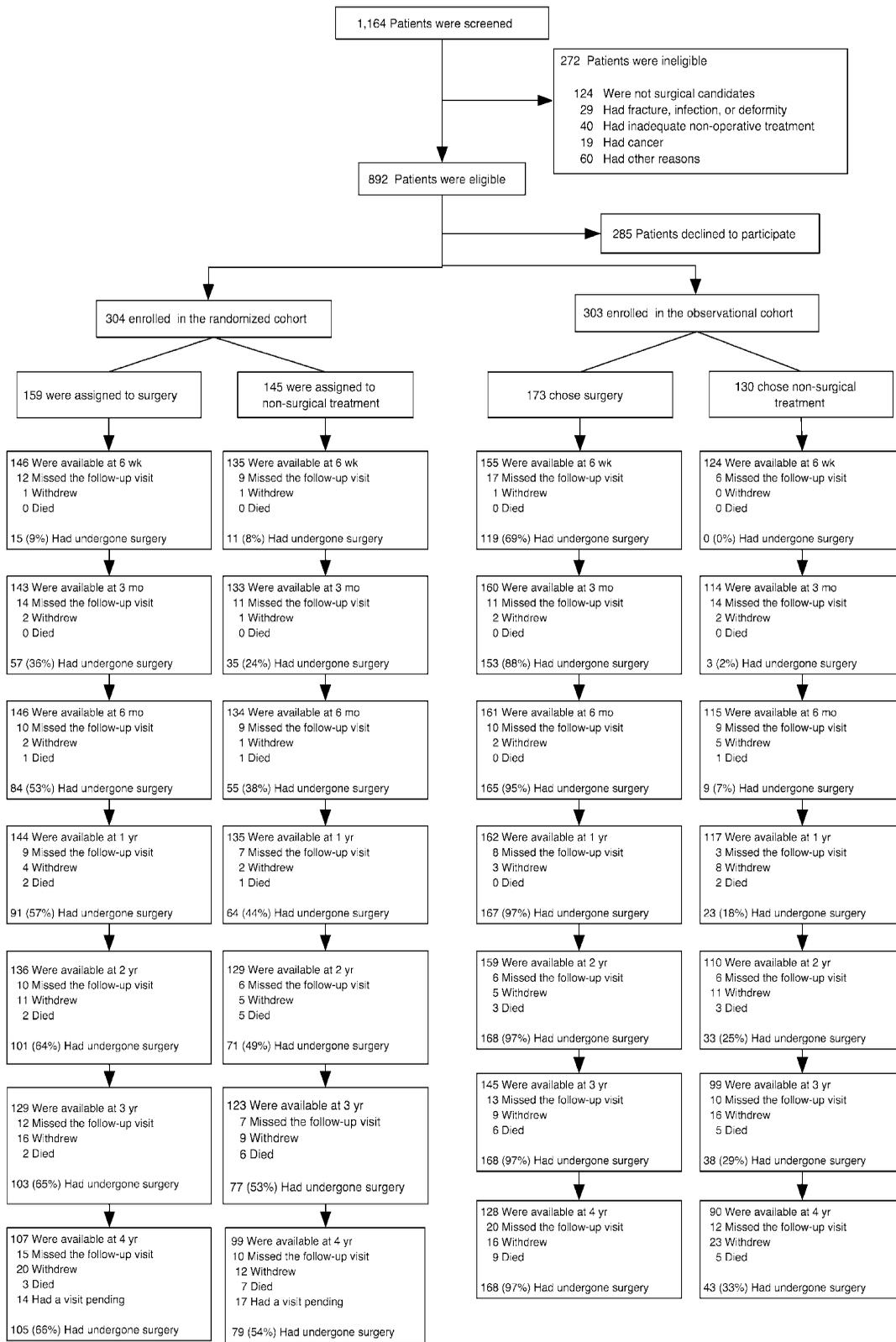
Fig. 1

Exclusion, enrollment, randomization, and follow-up of trial participants. The values for surgery, withdrawal, and death are cumulative over four years. For example, a total of three patients in the group assigned to receive surgery died during the four-year follow-up period.
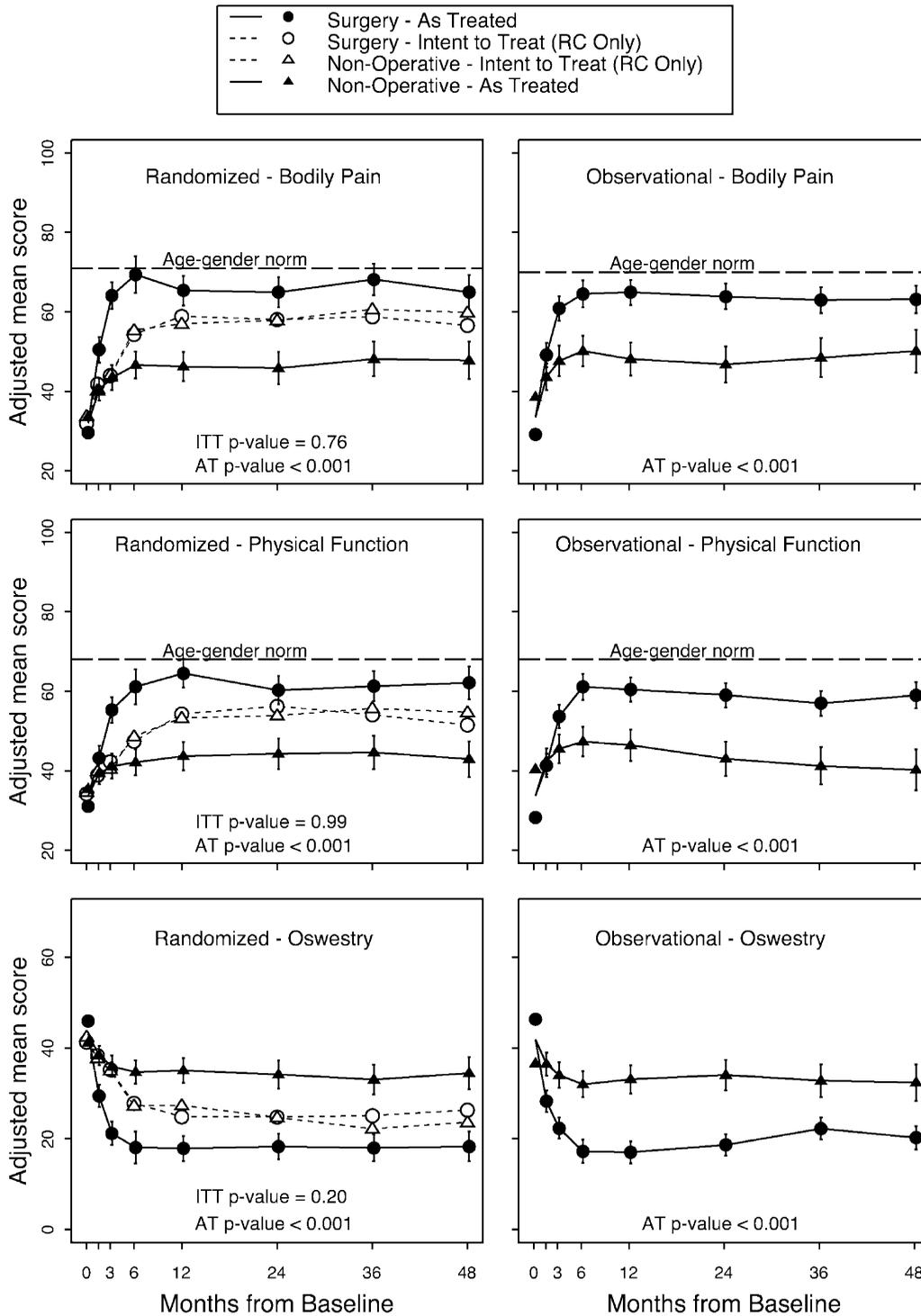
Fig. 2

Intent-to-treat (ITT) compared with as-treated (AT) analysis for the SF-36 bodily pain and physical function scores and the Oswestry Disability Index. RC = randomized cohort. Error bars indicate the 95% confidence intervals.

Statistical Methods

The statistical methods for the analysis of this trial have been reported in previous publications[9,10,13,14], and these descriptions are repeated or paraphrased here as necessary. In the initial analyses, the baseline characteristics were compared between the patients in the randomized cohort and those in the observational cohort and between the surgical and nonoperative groups in the combined cohort. The extent of missing data

and the percentage of patients who had undergone surgery were calculated for each study group at each scheduled follow-up point. Baseline predictors of the time until surgical treatment (including treatment crossovers) in both cohorts were determined through a stepwise proportional-hazards regression model with a criterion of p < 0.1 to enter and p > 0.05 to exit. Predictors of adherence to the assigned treatment and missing follow-up visits at one, two, three, and four years were determined through stepwise logistic regression. The primary analyses consisted of comparisons of surgical and nonoperative treatments, on the basis of the changes from baseline at each follow-up visit, with a mixed-effects model of longitudinal regression that included a random individual effect to account for correlation between repeated measurements. The randomized cohort was initially analyzed on an intent-to-treat basis. Because of crossover, subsequent analyses were based on treatments actually received.

In the as-treated analyses, the treatment indicator was a time-varying covariate, allowing for variable times of surgery. In the intent-to-treat analyses, all times were from enrollment. In the as-treated analyses, the times were from the beginning of treatment (that is, the time of surgery for the surgical group and the time of enrollment for the nonoperative group). Therefore, all changes from baseline before surgery were included in the estimates of the nonoperative treatment effect. Changes after surgery were assigned to the surgical group, with follow-up measured from the date of the surgery.

Repeated measures of outcomes were used as the dependent variables, and the treatment received was included as a time-varying covariate. Adjustments were made for postoperative visit times with respect to the time of surgery in order to approximate the designated follow-up times. Treatment comparisons were performed at designated follow-up times. In addition, a global significance test was based on the time-weighted average/area under the curve analysis over all time periods[23].

As-treated estimates of treatment effect from the randomized and observational cohorts were analyzed to establish comparability. Subsequent analyses combined the two cohorts. To adjust for potential confounding, baseline variables that were associated with missing data or treatment received were included as adjusting covariates in longitudinal regression models. Computations were performed with the use of the PROC MIXED procedure for continuous data and the PROC GENMOD procedure for binary and non-normal secondary outcomes from the SAS software package (version 9.1; SAS Institute, Cary, North Carolina). Significance was defined as p < 0.05 on the basis of a two-sided hypothesis test with no adjustments made for multiple comparisons. The data for these analyses were collected through May 1, 2008.

Source of Funding

Results

Overall, 607 of 892 eligible participants were enrolled in the degenerative spondylolisthesis SPORT trial; 304 were enrolled in the randomized cohort and 303, in the observational cohort. In the randomized cohort, 159 patients were assigned to receive surgery and 145, nonoperative treatment. Of the 159 assigned to receive surgery, 64% (101) underwent surgery by two years and 66% (105), by four years. Of the 145 patients assigned to receive nonoperative care, 49% (seventy-one) underwent surgery by two years and 54% (seventy-nine), by four years (Fig. 1). In the observational cohort, 173 patients initially chose surgery and 130 initially chose nonoperative care. Of the 173 who initially chose surgery, 97% (168) underwent surgery by two years and had had no additional surgery by four years. Of the 130 patients who initially chose nonoperative treatment, 25% (thirty-three) underwent surgery by two years and 33% (forty-three) underwent surgery by four years. In both cohorts combined, 395 patients underwent surgery within four years, with 87% (345) of them undergoing it within the first year. At four years, 35% (212) of the 607 patients had had only nonoperative treatment. A total of 601 patients (301 [99%] of the 304 enrollees in the randomized cohort and 300 [99%] of the 303 enrollees in the observational cohort), each with at least one follow-up visit in the four-year period, were included in the analysis (Fig. 1). The proportion of enrollees who supplied data at each follow-up interval ranged from 70% to 92%, with losses due to dropouts, missed visits, and deaths.

Patient Characteristics

A table in the Appendix shows the baseline characteristics and clinical findings of participants in the randomized and observational cohorts. The cohorts were remarkably similar except for their preferences for surgery (p < 0.001), with more randomized patients unsure of their preference (39% [117] of 301 compared with 7% [twenty] of the 300 patients in the observational cohort) and fewer randomized patients preferring either surgery (13% [thirty-eight of 301] compared with 43% [129 of 300]) or nonoperative treatment (15% [forty-four of 301] compared with 28% [eighty-three of 300]).

The table in the Appendix also shows summary statistics for the combined randomized and observational cohorts according to the treatment received. In the combined cohort, the patients who underwent surgery within four years were, at baseline, significantly younger and more likely to be receiving compensation (for example, Workers' Compensation or Social Security benefits) than those who received nonoperative treatment. They also had significantly worse pain, function, disability, and symptoms than the patients in the nonoperative group. The patients in the surgery group were significantly more dissatisfied with their symptoms and more often rated their symptoms as worsening at the time of enrollment; these patients also definitely preferred surgery.

TABLE I Change in Scores and Treatment Effects for Primary and Secondary Outcomes in the Randomized and Observational Cohorts Combined, According to Treatment Received*					➤
	Mean (Standard Error) at Baseline	3 Yr			
		Op. Group† (N = 326‡)	Nonop. Group† (N = 162‡)	Treatment Effect§	P Value
Primary outcomes#					
SF-36 bodily pain score	32.6 (0.7)	32.4 (1.3)	15.5 (1.6)	17.0 (13, 20.9)	<0.001
SF-36 physical function score	33.7 (0.8)	25.1 (1.2)	9.1 (1.6)	16.1 (12.3, 19.8)	<0.001
Oswestry Disability Index	42.6 (0.6)	−21.9 (0.96)	−9.3 (1.2)	−12.6 (−15.5, −9.7)	<0.001
Secondary outcomes**					
Stenosis Bothersomeness Index	14.9 (0.2)	−9.1 (0.35)	−4.5 (0.46)	−4.6 (−5.8, −3.5)	<0.001
Leg Pain Bothersomeness Scale	4.6 (0.1)	−3 (0.1)	−1.6 (0.1)	−1.3 (−1.7, −1)	<0.001
Low Back Pain Bothersomeness Scale	4.2 (0.1)	−2.1 (0.1)	−1.4 (0.1)	−0.7 (−1, −0.4)	<0.001
Very or somewhat satisfied with symptoms (%)	4.2 (2)	66.1	36.2	29.9 (20.6, 39.3)	<0.001
Very or somewhat satisfied with care (%)		87.8	69.3	18.5 (10, 27)	<0.001
Self-rated progress: major improvement (%)		71	25.2	45.8 (37.3, 54.3)	<0.001

*Scores are adjusted for age, sex, work status, body mass index, neuroforaminal involvement, depression, osteoporosis, joint problems, duration of current symptoms, reflex deficit, number of moderately or severely stenotic levels, hypertension, treatment preference, other comorbidities (including stroke; cancer; fibromyalgia; chronic fatigue syndrome; posttraumatic stress disorder; alcohol or drug dependency; lung, liver, kidney, blood vessel, and nervous system disease; migraine; and anxiety), baseline SF-36 score, baseline Oswestry Disability Index, baseline Stenosis Bothersomeness score, and center. †The values are given as the mean change from baseline (with the standard error in parentheses), except for the values for satisfaction with symptoms and care and for self-rated progress, which are given as the percentages of patients at the time of follow-up. ‡The sample sizes for the as-treated analyses reflect the number of patients contributing to the estimate in a given time-period, with use of the longitudinal modeling strategy explained in the Materials and Methods section, and may not correspond to the counts provided for each visit time in Figure 1. §The treatment effect is the difference between the operative-group and nonoperative-group mean changes from baseline. The 95% confidence interval is given in parentheses. #The SF-36 scores range from 0 to 100 points, with higher scores indicating less severe symptoms. The Oswestry Disability Index ranges from 0 to 100 points, with lower scores indicating less severe symptoms.**The Stenosis Bothersomeness Index ranges from 0 to 24 points, with lower scores indicating less severe symptoms, and the Leg Pain and Low Back Pain Bothersomeness Scales range from 0 to 6 points, with lower scores indicating less severe symptoms.

These observations highlight the need to control for baseline differences in the adjusted models. On the basis of the selection procedure for variables associated with treatment, missing data, and outcomes, the final as-treated models controlled for the following covariates: age, sex, work status, body mass index, neuroforamen involvement, depression, osteoporosis, joint problems, duration of current symptoms, reflex deficit, number of moderately or severely stenotic levels, hypertension, treatment preference, other comorbidities (including stroke; cancer; fibromyalgia; chronic fatigue syndrome; posttraumatic stress disorder; alcohol or drug dependency; lung, liver, kidney, blood vessel, and nervous system disease; migraine; and anxiety), baseline SF-36 score, baseline Oswestry Disability Index, baseline Stenosis Bothersomeness score, and center.

Nonoperative Treatments
Nonoperative treatments used during the SPORT included physical therapy (43% [176 of 412]), epidural steroid injections (47% [192 of 412]), nonsteroidal anti-inflammatory drugs (54% [224 of 412]), and opioids (35% [146 of 412]). Nonoperative treatments were similar in the randomized cohort and the observational cohort, although more patients in the randomized cohort than in the observational cohort reported visits

to a surgeon (48% [122 of 252] compared with 38% [sixty of 160], p = 0.04), receiving injections (51% [128 of 252] compared with 40% [sixty-four of 160], p = 0.04), and opioid use (40% [100 of 252] compared with 29% [forty-six of 160], p = 0.03).

Surgical Treatment and Complications
The median surgical time for the combined cohort was 198 minutes, with a mean blood loss of 583 mL (see Appendix). There was no significant difference between the cohorts with regard to the rates of intraoperative blood replacement, but there was a difference in the postoperative transfusion rates (16% [twenty-nine of 178] in the randomized cohort compared with 24% [fifty-one of 209] in the observational cohort, p = 0.08). The most common surgical complication was a dural tear (11% [forty-one of 387]). The four-year reoperation rate was 15% (fifty-nine of 387), and the rate of recurrent stenosis was 5% (nineteen of 387).

Within four years after enrollment, there were seven deaths in the nonoperative group, compared with twenty-two deaths expected on the basis of age and sex-specific mortality rates[24], and seventeen deaths in the surgery group, compared with twenty-eight expected. The hazard ratio based on a proportional-hazards model adjusted for age was 1.9 (95% confidence interval, 0.76 to 4.6; p = 0.17). All twenty-four deaths

TABLE I (*continued*)			
	4 Yr		
Op. Group† (N = 264‡)	Nonop. Group† (N = 131‡)	Treatment Effect§	P Value
31.1 (1.4)	15.8 (1.8)	15.3 (11, 19.7)	<0.001
26.6 (1.3)	7.7 (1.7)	18.9 (14.8, 23)	<0.001
−23 (1)	−8.6 (1.3)	−14.3 (−17.5, −11.1)	<0.001
−9.2 (0.38)	−3.9 (0.5)	−5.3 (−6.5, −4.1)	<0.001
−3 (0.1)	−1.5 (0.2)	−1.5 (−1.9, −1.1)	<0.001
−2.1 (0.1)	−1.2 (0.1)	−1.0 (−1.3, −0.6)	<0.001
63.5	30.2	33.4 (23.4, 43.4)	<0.001
86.6	67.8	18.7 (9, 28.5)	<0.001
67.1	21	46.1 (37, 55.2)	<0.001

were independently reviewed, and eighteen were judged to be not related to the treatment. Four deaths were of unknown cause but occurred between 621 and 1379 days following the surgery. Two deaths, both in the surgical group, were judged to be potentially related to the treatment; one patient died of respiratory distress thirty-two days after the surgery, and the other died of sepsis eighty-two days after the surgery.

Crossover
Nonadherence to the treatment assignment affected both arms of the SPORT: patients in the surgical arm chose to delay or declined surgery, and patients in the nonoperative arm crossed over to receive surgery (Fig. 1). The characteristics of crossover patients that differed significantly from those of the patients who did not cross over are shown in a table in the Appendix. Patients who crossed over to receive nonoperative care were older, had less pain and disability, were less bothered by their symptoms, and had stronger baseline treatment preferences for nonoperative care as compared with patients who did not cross over. In the group randomized to receive surgery, improvements in pain and function scores at the early follow-up intervals also predicted nonadherence with surgical treatment. In the group randomized to receive nonoperative care, those who crossed over to receive surgery were younger, more often married, and more dissatisfied with their symptoms, and they had a stronger baseline preference for surgery.

Main Treatment Effects
The intent-to-treat analysis of the randomized cohort showed no significant differences between surgery and nonoperative care on the basis of overall global hypothesis tests for differences in mean changes from baseline (Fig. 2). Estimated treatment effects at four years slightly favored nonoperative treatment but were not significant; these effects were −2 for SF-36 bodily pain (95% confidence interval, −8.6 to 4.6, p =

0.56), −3.1 for physical function (95% confidence interval, −9.2 to 3, p = 0.32), and 4.1 for the Oswestry Disability Index (95% confidence interval, −0.8 to 9.1, p = 0.1).

In the as-treated analysis, the treatment effects for the randomized and observational cohorts were similar at four years (Fig. 2). In the randomized and observational groups, respectively, these effects were 17.1 (95% confidence interval, 10.9 to 23.4) and 13 (95% confidence interval, 6.6 to 19.4) for bodily pain, 19.2 (95% confidence interval, 13.4 to 25.1) and 18.8 (95% confidence interval, 12.7 to 24.9) for physical function, and 16.2 (95% confidence interval, −20.7 to −11.6) and −12.2 (95% confidence interval, −16.8 to −7.5) for the Oswestry Disability Index.

The global hypothesis test comparing the as-treated treatment effects between the randomized and observational groups over all time periods showed no difference between the cohorts (p = 0.18 for bodily pain, p = 0.17 for physical function, and p = 0.77 for the Oswestry Disability Index).

Results from the intent-to-treat and as-treated analyses of the two cohorts are compared in Figure 2. The as-treated treatment effects significantly favored surgery in both cohorts. In the combined analysis, the treatment effects were significantly in favor of surgery for all primary and secondary outcome measures at each time point out to four years (Table I).

Subgroup Analyses
Table II shows the results of subgroup analyses comparing the time-weighted average outcomes between patients with and those without neurogenic claudication on baseline clinical examination and between patients with and those without a neurological deficit on baseline clinical examination. Approximately 85% of the subjects had neurogenic claudication, while the remaining 15% had more radicular symptoms with evidence of nerve root irritation. Those with neurogenic claudication had similar overall results of surgery, as compared

TABLE II Time-Weighted Average of Treatment Effects at Four Years According to Whether the Patient Had Neurogenic Claudication or Any Neurological Deficit

	Mean Change in Score Compared with Baseline* (Standard Error)		Treatment Effect (95% Confidence Interval)
	Op. Group	Nonop. Group	
SF-36 bodily pain score			
Neurogenic claudication	31.3 (0.9)	12.9 (1.1)	18.4 (16 to 20.8)
No neurogenic claudication	29.7 (2.2)	19.2 (2.4)	10.4 (5.2 to 15.7)
P value	0.51	0.015	0.006
SF-36 physical function score			
Neurogenic claudication	25.3 (0.9)	8.3 (1)	17 (14.7 to 19.3)
No neurogenic claudication	24.3 (2.2)	12.8 (2.4)	11.5 (6.5 to 16.6)
P value	0.67	0.083	0.047
Oswestry Disability Index			
Neurogenic claudication	−23 (0.7)	−7.4 (0.8)	−15.6 (−17.4 to −13.8)
No neurogenic claudication	−20.6 (1.7)	−11.8 (1.9)	−8.8 (−12.8 to −4.9)
P value	0.20	0.035	0.002
SF-36 bodily pain score			
Any neurological deficit	31.8 (1.2)	13.7 (1.4)	18 (15.1 to 21)
No neurological deficit	30.1 (1.3)	14.3 (1.5)	15.8 (12.7 to 19)
P value	0.37	0.79	0.29
SF-36 physical function score			
Any neurological deficit	26.5 (1.2)	7.6 (1.3)	19 (16.2 to 21.8)
No neurological deficit	23.6 (1.3)	11 (1.5)	12.7 (9.7 to 15.7)
P value	0.11	0.11	0.002
Oswestry Disability Index			
Any neurological deficit	−22.5 (0.9)	−7.2 (1.1)	−15.3 (−17.5 to −13.1)
No neurological deficit	−22.9 (1)	−9.4 (1.2)	−13.5 (−15.9 to −11.2)
P value	0.78	0.20	0.27

*The SF-36 scores range from 0 to 100 points, with higher scores indicating less severe symptoms. The Oswestry Disability Index ranges from 0 to 100 points, with lower scores indicating less severe symptoms.

with those without neurogenic claudication, but showed less improvement after nonoperative care; this resulted in a significantly greater surgical treatment effect in the neurogenic claudication group. The presence of a neurological deficit did not result in significant differences in either the surgical or the nonoperative outcomes, and there was no consistent difference in treatment effects.

Only 7% of the subjects had compensation claims, so this subgroup was too small to allow meaningful comparisons. Ancillary studies from the SPORT have provided information on the effect of baseline radiographic predictors (spondylolisthesis grade, degree of mobility on flexion radiographs, and disc space height) on outcomes[25] as well as the relative outcomes of surgery in subgroups based on the type of fusion performed[26].

Discussion

Our intent-to-treat analysis of patients who had had, for at least twelve weeks, signs and symptoms of degenerative spondylolisthesis with spinal stenosis confirmed by radio-graphic studies demonstrated that surgery had no significant advantage over nonoperative treatment; at the three and four-year follow-up points, nonoperative treatment showed a slight but not a significant advantage. However, these results must be viewed in the context of substantial rates of nonadherence to the assigned treatment. This mixing of treatments generally biases effect estimates toward the null[9,10,13,14,27].

The treatment effect in favor of surgery that was found in the as-treated analysis suggests that the intent-to-treat analysis underestimated the true effect of surgery. The effect was seen as early as six weeks, was maximum by six to twelve months, and persisted over four years. The nonoperative treatment group demonstrated only modest improvement over time. The results in both treatment groups were maintained between two and four years.

This study provided an opportunity to compare the results between patients who were willing to participate in a randomized study (the randomized cohort) and those who were not (the observational cohort). These two cohorts were remarkably similar at baseline. Other than treatment preference,

the only significant differences were small ones in the level and location of the stenosis on baseline imaging. The cohorts also had similar outcomes, with no significant differences between the treatment effects in the as-treated analyses, findings that support the validity of the combined analysis. Although these analyses were not based entirely on randomized treatment assignments, the results were strengthened by the use of specific inclusion and exclusion criteria, the large sample size, and the adjustment for potentially confounding baseline factors[28].

Comparisons with Other Studies

The characteristics of the participants and the short-term outcomes in the SPORT as previously reported[9] are similar to those in studies of degenerative spondylolisthesis and of mixed cohorts of patients who had stenosis with and without degenerative spondylolisthesis.

The surgical outcomes in the SPORT were generally similar to those in previous surgical series. Herkowitz and Kurz[7] reported absolute improvements of 33% in scores for back pain and 55% in scores for leg pain (on 6-point scales) at an average of three years, results that are similar to the improvements of 30% and 43% (on 7-point scales), respectively, seen in the SPORT at four years. Also, the improvements at four years after surgery for degenerative spondylolisthesis in the SPORT were similar to the outcomes of surgery in the Maine Lumbar Spine Study (MLSS)[2] of a mixed cohort of patients who had stenosis with and without degenerative spondylolisthesis. The improvements in the scores for the bothersomeness of stenosis, leg pain, and low-back pain were nearly identical in the two studies: −9.2, −3, and −2.1 points in the SPORT and −9.4, −3.5, and −1.7 points in the MLSS.

Carragee et al. reported on the outcomes of fusion with or without decompression in patients with isthmic spondylolisthesis[29-31]. They found improved outcomes with fusion independent of decompression, particularly among those with instability on flexion-extension radiographs. They also found that circumferential fusion had some early advantage over posterior fusion with instrumentation. However, isthmic spondylolisthesis and degenerative spondylolisthesis, the focus of our study, are quite different disease processes; in particular, decompression was important in our population because all patients had neurogenic claudication or radicular leg pain. In addition, Pearson et al. found that patients with degenerative spondylolisthesis who had had more mobility at baseline actually had better nonoperative outcomes than those with less movement as documented radiographically[25], the opposite of what was found in patients with isthmic spondylolisthesis.

There was little evidence of harm from either treatment in our study. In the interval between two and four years, there were no cases of paralysis in either the surgical or the nonoperative group. The four-year rate of reoperation for recurrent stenosis or spondylolisthesis was 5%, and the overall reoperation rate increased from 12% at two years to 15% at four years (compared with 6.2% at four years in the MLSS[2]). The perioperative mortality rate remained unchanged at 0.5%, which is less than the 1.3% rate seen in Medicare patients after fusion surgery for spondylolisthesis[24].

The four-year mortality rate was similar in the two treatment groups and was lower than actuarial projections. The rate in the nonoperative group was somewhat lower than that in the operative group, although not significantly so. It should be noted that higher rates of complications have been reported with increasing age and coexisting medical conditions[32].

Limitations

A major limitation of this study is the marked degree of non-adherence to the randomized treatment. This reduced the power of the intent-to-treat analysis to demonstrate a treatment effect. Although the as-treated analysis lacked the strong protection from confounding conferred by randomization, these analyses were carefully controlled for important covariates and yielded results similar to those of prior studies[2,7,9]. Another limitation is the heterogeneity of the treatment interventions. The choice of nonoperative therapies as well as the decision regarding whether and how to perform the fusions were at the discretion of the treating physician and the patient. This resulted in a clinically relevant comparison of the effectiveness of current treatment practices but does not allow us to draw direct conclusions regarding the efficacy of one specific surgical technique compared with one specific nonoperative treatment.

Overview

In the as-treated analysis, combining the randomized and observational cohorts of patients with spinal stenosis secondary to degenerative spondylolisthesis, those treated surgically were found to have significantly greater improvement in scores for pain, function, satisfaction, and self-rated progress over four years compared with patients treated nonoperatively. The results in both groups were stable between two and four years.

Appendix

(eA) Tables showing baseline characteristics of the study cohorts, the types of operative treatments and complications, and predictions of adherence to treatment among patients in the randomized cohort are available with the electronic versions of this article, on our web site at jbjs.org (go to the article citation and click on "Supplementary Material") and on our quarterly CD/DVD (call our subscription department, at 781-449-9780, to order the CD or DVD). ■

James N. Weinstein, DO, MS
Jon D. Lurie, MD, MS
Tor D. Tosteson, ScD
Wenyan Zhao, MS
Emily A. Blood, MS
Anna N.A. Tosteson, ScD
Departments of Orthopaedics (J.N.W., W.Z., and E.A.B.), Medicine (J.D.L. and A.N.A.T.), and Community and Family Medicine (T.D.T. and A.N.A.T.), Dartmouth Medical School, One Medical Center Drive,

Lebanon, NH 03756. E-mail address for J.N. Weinstein:
SPORT@dartmouth.edu

Nancy Birkmeyer, PhD
University of Michigan Health System, 2101 Taubman Center,
1500 East Medical Center Drive, Ann Arbor, MI 48109-0346

Harry Herkowitz, MD
Department of Orthopaedic Surgery, William Beaumont Hospital,
3535 West 13 Mile Road, #744, Royal Oak, MI 48073

Michael Longley, MD
Nebraska Foundation for Spinal Research, 11819 Miracle Hills Drive,
Suite 102, Omaha, NE 68154

Lawrence Lenke, MD
Department of Orthopaedic Surgery,
Washington University School of Medicine,
One Barnes-Jewish Hospital Plaza, Suite 11300 West Pavilion,
St. Louis, MO 63110

Sanford Emery, MD
Department of Orthopaedics, West Virginia University,
3700 Robert C. Byrd Health Science Center South,
Morgantown, WV 26506

Serena S. Hu, MD
University of California Medical Center, 400 Parnassus Avenue,
Third Floor, San Francisco, CA 94143

References

1. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES. United States' trends and regional variations in lumbar spine surgery: 1992-2003. Spine. 2006; 31:2707-14.

2. Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, Singer DE. The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. Spine. 1996;21:1787-95.

3. 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. 2000;25:556-62.

4. Malmivaara A, Slätis P, Heliövaara M, Sainio P, Kinnunen H, Kankare J, Dalin-Hirvonen N, Seitsalo S, Herno A, Kortekangas P, Niinimäki T, Rönty H, Tallroth K, Turunen V, Knekt P, Härkänen T, Hurri H; Finnish Lumbar Spinal Research Group. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine. 2007;32:1-8.

5. 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;6:461-72.

6. Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. Spine. 1997;22:2807-12.

7. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Joint Surg Am. 1991;73:802-8.

8. Birkmeyer NJ, Weinstein JN, Tosteson AN, Tosteson TD, Skinner JS, Lurie JD, Deyo R, Wennberg JE. Design of the Spine Patient Outcomes Research Trial (SPORT). Spine. 2002;27:1361-72.

9. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, Birkmeyer NJ, Hilibrand AS, Herkowitz H, Cammisa FP, Albert TJ, Emery SE, Lenke LG, Abdu WA, Longley M, Errico TJ, Hu SS. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med. 2007;356:2257-70.

10. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goldberg H, Berven S, An H; SPORT Investigators. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med. 2008;358:794-810.

11. Katz JN, Lipson SJ, Chang LC, Levine SA, Fossel AH, Liang MH. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. Spine. 1996;21:92-8.

12. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. JAMA. 2003; 290:1624-32.

13. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA. Surgical vs. nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA. 2006;296:2451-9.

14. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA. 2006;296:2441-50.

15. Cummins J, Lurie JD, Tosteson TD, Hanscom B, Abdu WA, Birkmeyer NJ, Herkowitz H, Weinstein J. Descriptive epidemiology and prior healthcare utilization of patients in the Spine Patient Outcomes Research Trials's (SPORT) three

observational cohorts: disc herniation, spinal stenosis, and degenerative spondylolisthesis. Spine. 2006;31:806-14.

16. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care. 1994;32:40-66.

17. Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, McGlynn EA, Ware JE Jr. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA. 1989;262:907-13. Erratum in: JAMA. 1989;262:2542.

18. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473-83.

19. Ware JE Jr. SF-36 Health Survey: manual and interpretation guide. Boston: Nimrod Press; 1993.

20. Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American Spine Society lumbar spine outcome assessment instrument: reliability and validity tests. Spine. 1996;21:741-9.

21. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. Spine. 1986;11:28-30.

22. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine. 1995;20:1899-909.

23. Fitzmaurice G, Laird NM, Ware JH. Applied longitudinal analysis. Hoboken, NJ: Wiley-Interscience; 2004.

24. Deyo RA, Ciol MA, Cherkin DC, Loeser JD, Bigos SJ. Lumbar spinal fusion. A cohort study of complications, reoperations, and resource use in the Medicare population. Spine. 1993;18:1463-10.

25. Pearson AM, Lurie JD, Blood EA, Frymoyer JW, Braeutigam H, An H, Girardi FP, Weinstein JN. Spine Patient Outcomes Research Trial: radiographic predictors of clinical outcomes after operative or nonoperative treatment of degenerative spondylolisthesis. Spine. 2008;33:2759-66.

26. Abdu WA, Spratt KF, Lurie J, Tosteson TD, Tosteson ANA, Weinstein JN. Outcomes for different surgical approaches to degenerative spondylolisthesis: SPORT. Presented at SpineWeek, the Annual Meeting of the International Society for the Study of the Lumbar Spine; 2008 May 26-31; Geneva, Switzerland.

27. Meinert CL. Clinical trials: design, conduct, and analysis. New York: Oxford University Press; 1986.

28. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med. 2000;342:1887-92.

29. Carragee EJ. Single-level posterolateral arthrodesis, with or without posterior decompression, for the treatment of isthmic spondylolisthesis in adults. A prospective, randomized study. J Bone Joint Surg Am. 1997;79:1175-80.

30. Carragee EJ, Lincoln T, Parmar VS, Alamin T. A gold standard evaluation of the ''discogenic pain'' diagnosis as determined by provocative discography. Spine. 2006;31:2115-23.

31. Swan J, Hurwitz E, Malek F, van den Haak E, Cheng I, Alamin T, Carragee E. Surgical treatment for unstable low-grade isthmic spondylolisthesis in adults: a prospective controlled study of posterior instrumented fusion compared with combined anterior-posterior fusion. Spine J. 2006;6:606-14.

32. Ciol MA, Deyo RA, Howell E, Kreif S. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. J Am Geriatr Soc. 1996;44:285-90.