ORIGINAL ARTICLE

Surgical versus Nonsurgical Therapy for Lumbar Spinal Stenosis

James N. Weinstein, D.O., M.S., Tor D. Tosteson, Sc.D., Jon D. Lurie, M.D., M.S., Anna N.A. Tosteson, Sc.D., Emily Blood, M.S., Brett Hanscom, M.S., Harry Herkowitz, M.D., Frank Cammisa, M.D., Todd Albert, M.D., Scott D. Boden, M.D., Alan Hilibrand, M.D., Harley Goldberg, D.O., Sigurd Berven, M.D., and Howard An, M.D., for the SPORT Investigators*

ABSTRACT

BACKGROUND

Surgery for spinal stenosis is widely performed, but its effectiveness as compared with nonsurgical treatment has not been shown in controlled trials.

METHODS

Surgical candidates with a history of at least 12 weeks of symptoms and spinal stenosis without spondylolisthesis (as confirmed on imaging) were enrolled in either a randomized cohort or an observational cohort at 13 U.S. spine clinics. Treatment was decompressive surgery or usual nonsurgical care. The primary outcomes were measures of bodily pain and physical function on the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36) and the modified Oswestry Disability Index at 6 weeks, 3 months, 6 months, and 1 and 2 years.

RESULTS

A total of 289 patients were enrolled in the randomized cohort, and 365 patients were enrolled in the observational cohort. At 2 years, 67% of patients who were randomly assigned to surgery had undergone surgery, whereas 43% of those who were randomly assigned to receive nonsurgical care had also undergone surgery. Despite the high level of nonadherence, the intention-to-treat analysis of the randomized cohort showed a significant treatment effect favoring surgery on the SF-36 scale for bodily pain, with a mean difference in change from baseline of 7.8 (95% confidence interval, 1.5 to 14.1); however, there was no significant difference in scores on physical function or on the Oswestry Disability Index. The as-treated analysis, which combined both cohorts and was adjusted for potential confounders, showed a significant advantage for surgery by 3 months for all primary outcomes; these changes remained significant at 2 years.

CONCLUSIONS

In the combined as-treated analysis, patients who underwent surgery showed significantly more improvement in all primary outcomes than did patients who were treated nonsurgically. (ClinicalTrials.gov number, NCT00000411.)

From the Departments of Orthopedics (J.N.W., E.B., B.H.), Community and Family Medicine (T.D.T., J.D.L., A.N.A.T.), and Medicine (J.D.L., A.N.A.T.), Dartmouth Medical School, Hanover, NH, and Dartmouth–Hitchcock Medical Center, Lebanon, NH; William H. Beaumont Hospital, Royal Oak, MI (H.H.); Hospital for Special Surgery, New York (F.C.); Rothman Institute at Thomas Jefferson University, Philadelphia (T.A., A.H.); Emory Spine Center, Emory University, Atlanta (S.D.B.); Kaiser Permanente, San Francisco (H.G.); University of California at San Francisco, San Francisco (S.B.); and Rush–Presbyterian–St. Luke's Medical Center, Chicago (H.A.). Address reprint requests to Dr. Weinstein at the Dartmouth Institute for Health Policy and Clinical Practice, Department of Orthopedics, Dartmouth Medical School, 1 Medical Center Dr., Lebanon, NH 03756, or at sport@dartmouth.edu.

*Investigators in the Spine Patient Outcomes Research Trial (SPORT) are listed in the Appendix.

N Engl J Med 2008;358:794-810.
Copyright © 2008 Massachusetts Medical Society.

Spinal stenosis is a narrowing of the spinal canal with encroachment on the neural structures by surrounding bone and soft tissue. Patients typically present with radicular leg pain or with neurogenic claudication (pain in the buttocks or legs on walking or standing that resolves with sitting down or lumbar flexion). Spinal stenosis is the most common reason for lumbar spine surgery in adults over the age of 65 years.[1,2] Indications for surgery appear to vary widely, and rates of procedures vary by at least a factor of 5 across geographic areas.[3,4] Radiographic evidence of stenosis is frequently asymptomatic; thus, careful clinical correlation between symptoms and imaging is critical.[5,6]

A 2005 Cochrane review found that the paucity and heterogeneity of evidence limited conclusions regarding surgical efficacy for spinal stenosis. The trials comparing surgical with nonsurgical treatment were generally small and involved patients both with and without degenerative spondylolisthesis.[7-12] We know of no randomized trials of isolated spinal stenosis without degenerative spondylolisthesis.

In the Spine Patient Outcomes Research Trial (SPORT), we report on the 2-year outcomes of patients with spinal stenosis without degenerative spondylolisthesis to analyze the relative efficacy of surgical versus nonsurgical treatment.

METHODS

STUDY DESIGN

SPORT was an investigator-initiated study conducted in 11 states at 13 U.S. medical centers with multidisciplinary spine practices. The study included both a randomized cohort and a concurrent observational cohort of patients who declined to undergo randomization.[13-16] This design allowed for improved generalizability of the findings.[17] The ethics committee at each participating institution approved a standardized protocol. An independent data and safety monitoring board evaluated interim safety and efficacy outcomes at 6-month intervals.[13-16,18] Stopping rules were provided on the basis of the alpha spending function of DeMets and Lan.[19]

PATIENT POPULATION

All patients had a history of neurogenic claudication or radicular leg symptoms for at least 12 weeks and confirmatory cross-sectional imaging showing lumbar spinal stenosis at one or more levels; all patients were judged to be surgical candidates. Patients with degenerative spondylolisthesis were studied separately.[16] Patients with lumbar instability (which was defined as translation of more than 4 mm or 10 degrees of angular motion between flexion and extension on upright lateral radiographs) were excluded. The type of nonsurgical care before enrollment was not prespecified but included physical therapy (68% of patients), epidural injections (56%), chiropractic (28%), the use of antiinflammatory drugs (55%), and the use of opioid analgesics (27%).

Research nurses at each site verified eligibility. Patients were offered enrollment in either cohort. To aid in obtaining written informed consent, patients viewed evidence-based videotapes with standardized information regarding alternative treatments.[20,21] Patients in the randomized cohort received treatment assignments with the use of randomly permuted blocks with variable block sizes stratified according to center. Patients in the observational cohort chose their treatment at enrollment with their physician. Enrollment began in March 2000 and ended in March 2005.

STUDY INTERVENTIONS

The protocol surgery was standard posterior decompressive laminectomy.[13] The nonsurgical protocol was "usual care," which was recommended to include at least active physical therapy, education or counseling with home exercise instruction, and the administration of nonsteroidal antiinflammatory drugs, if tolerated.[13,18]

STUDY MEASURES

Primary outcomes were measures of bodily pain and physical function on the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36)[22-25] and on the modified Oswestry Disability Index (American Academy of Orthopaedic Surgeons–MODEMS [Musculoskeletal Outcomes Data Evaluation and Management Systems] version),[26] measured at 6 weeks, 3 months, 6 months, and 1 and 2 years. (SF-36 scores range from 0 to 100, with higher scores indicating less severe symptoms. The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms.)

If surgery was delayed beyond 6 weeks, additional follow-up data were obtained at 6 weeks and at 3 months after surgery. Secondary out-

comes included patient-reported improvement, satisfaction with current symptoms and care,[27] and the bothersomeness of both stenosis[7,28] and low back pain.[7] The effect of treatment was defined as the difference in the mean change from baseline between the surgical group and the nonsurgical group.

STATISTICAL ANALYSIS

For the randomized cohort, we determined that a sample size of 185 per group was needed to detect a 10-point difference in bodily pain and physical function on the SF-36 or a similar effect on the Oswestry Disability Index[13] on the basis of a t-test, with a two-sided significance level of 0.05 and a power of 85%. Standard deviations for changes from baseline were derived from pilot data on repeated visits. The sample-size calculation allowed for 20% missing data but did not account for any specific levels of nonadherence.

Initial analyses compared the baseline characteristics of patients in the randomized cohort with those in the observational cohort and between study groups in the combined cohorts. The extent of missing data and the percentage of patients undergoing surgery were calculated according to study group for each scheduled follow-up. Baseline predictors of the time until surgical treatment (including treatment crossovers) in both cohorts were determined through a stepwise proportional-hazards regression model with an inclusion criterion of $P<0.1$ to enter and $P>0.05$ to exit. Predictors of missing follow-up visits at 1 year were determined through stepwise logistic regression.

Primary analyses compared surgical and nonsurgical treatments with the use of 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 intention-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. For the intention-to-treat analyses, all times are from enrollment. For the as-treated analysis, the times are from the beginning of treatment (i.e., the time of surgery for the surgical group and the time of enrollment for the nonsurgical group). Therefore, all changes from baseline before surgery were included in the estimates of the nonsurgical treatment effect. After surgery, changes were assigned to the surgical group, with follow-up measured from the date of surgery. Repeated measures of outcomes were used as the dependent variables, and treatment received was included as a time-varying covariate. Adjustments were made for the time of surgery with respect to the original enrollment date so as to approximate the designated follow-up times.

The randomized and observational cohorts were each analyzed to produce separate as-treated estimates of treatment effect. These results were compared with the use of a Wald test to simultaneously test all follow-up visit times for differences in estimated treatment effects between the two cohorts.[29] 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.[29] 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 in SAS software, version 9.1 (SAS Institute). Statistical significance was defined as $P<0.05$ on the basis of a two-sided hypothesis test with no adjustments made for multiple comparisons. Data for these analyses were collected through March 2, 2007.

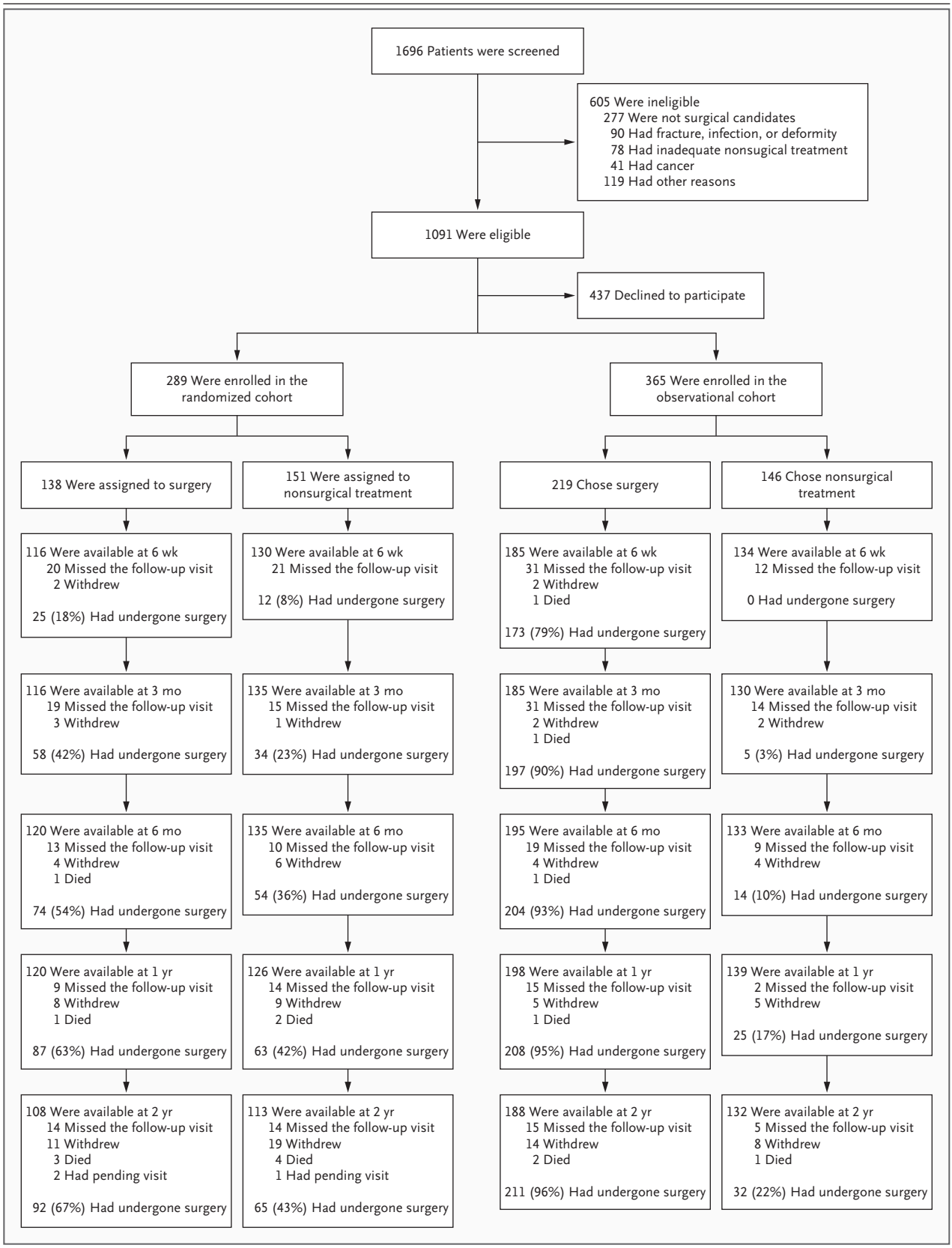
RESULTS

PATIENTS

A total of 654 patients were enrolled out of 1091 who were eligible for enrollment: 289 in the randomized cohort and 365 in the observational cohort (Fig. 1). In the randomized cohort, 138 patients were assigned to the surgical group, and 151 were assigned to the nonsurgical group. In the surgery group, 63% had undergone surgery at 1 year and 67% at 2 years. In the nonsurgical group, 42% had undergone surgery at 1 year and

Figure 1 (facing page). Enrollment, Randomization, and Follow-up.

The numbers of patients who withdrew from the study, died, or underwent surgery are cumulative during the 2-year follow-up period.


1696 Patients were screened
605 Were ineligible
277 Were not surgical candidates
90 Had fracture, infection, or deformity
78 Had inadequate nonsugical treatment
41 Had cancer
119 Had other reasons
1091 Were eligible
437 Declined to participate
289 Were enrolled in the randomized cohort
365 Were enrolled in the observational cohort
138 Were assigned to surgery
151 Were assigned to nonsurgical treatment
219 Chose surgery
146 Chose nonsurgical treatment
116 Were available at 6 wk
20 Missed the follow-up visit
2 Withdrew
25 (18%) Had undergone surgery
130 Were available at 6 wk
21 Missed the follow-up visit
12 (8%) Had undergone surgery
185 Were available at 6 wk
31 Missed the follow-up visit
2 Withdrew
1 Died
173 (79%) Had undergone surgery
134 Were available at 6 wk
12 Missed the follow-up visit
0 Had undergone surgery
116 Were available at 3 mo
19 Missed the follow-up visit
3 Withdrew
58 (42%) Had undergone surgery
135 Were available at 3 mo
15 Missed the follow-up visit
1 Withdrew
34 (23%) Had undergone surgery
185 Were available at 3 mo
31 Missed the follow-up visit
2 Withdrew
1 Died
197 (90%) Had undergone surgery
130 Were available at 3 mo
14 Missed the follow-up visit
2 Withdrew
5 (3%) Had undergone surgery
120 Were available at 6 mo
13 Missed the follow-up visit
4 Withdrew
1 Died
74 (54%) Had undergone surgery
135 Were available at 6 mo
10 Missed the follow-up visit
6 Withdrew
54 (36%) Had undergone surgery
195 Were available at 6 mo
19 Missed the follow-up visit
4 Withdrew
1 Died
204 (93%) Had undergone surgery
133 Were available at 6 mo
9 Missed the follow-up visit
4 Withdrew
14 (10%) Had undergone surgery
120 Were available at 1 yr
9 Missed the follow-up visit
8 Withdrew
1 Died
87 (63%) Had undergone surgery
126 Were available at 1 yr
14 Missed the follow-up visit
9 Withdrew
2 Died
63 (42%) Had undergone surgery
198 Were available at 1 yr
15 Missed the follow-up visit
5 Withdrew
1 Died
208 (95%) Had undergone surgery
139 Were available at 1 yr
2 Missed the follow-up visit
5 Withdrew
25 (17%) Had undergone surgery
108 Were available at 2 yr
14 Missed the follow-up visit
11 Withdrew
3 Died
2 Had pending visit
92 (67%) Had undergone surgery
113 Were available at 2 yr
14 Missed the follow-up visit
19 Withdrew
4 Died
1 Had pending visit
65 (43%) Had undergone surgery
188 Were available at 2 yr
15 Missed the follow-up visit
14 Withdrew
2 Died
211 (96%) Had undergone surgery
132 Were available at 2 yr
5 Missed the follow-up visit
8 Withdrew
1 Died
32 (22%) Had undergone surgery

43% at 2 years. In the observational cohort, 219 patients initially chose surgery and 146 patients initially chose nonsurgical care. Of those who initially chose surgery, 95% had undergone surgery at 1 year and 96% at 2 years. Of those who initially chose nonsurgical treatment, 17% had undergone surgery at 1 year and 22% at 2 years. In the two cohorts combined, 400 patients received surgery at some point during the first 2 years, and 254 received nonsurgical treatment.

The proportion of enrollees who supplied data at each follow-up interval ranged from 83 to 89%, with losses due to dropouts, missed visits, or deaths. A total of 634 patients, each with at least one follow-up through 2 years, were included in the analysis, including 278 patients (96%) in the randomized cohort and 356 patients (98%) in the observational cohort.

CHARACTERISTICS OF THE PATIENTS

Characteristics of the patients at baseline in the two cohorts are compared in Table 1. Overall, the cohorts were similar. However, patients in the observational cohort had more signs of nerve-root tension and less lateral recess stenosis and expressed stronger treatment preferences than did patients in the randomized cohort.

Summary statistics for the combined cohorts are also shown in Table 1, according to treatment received. The study population had a mean age of 65 years; a majority were white men who had attended college. Of these patients, 80% had classic neurogenic claudication, and 79% had associated dermatomal pain radiation; 91% had stenosis at L4 or L5, and 61% had more than one level of stenosis. For most patients, the overall stenosis was graded as severe.

At baseline, the group undergoing surgery was younger and more likely to be working than was the group that did not undergo surgery. Patients in the surgical group had more pain, a lower level of function, more psychological distress, and more self-reported disability than did patients in the nonsurgical group. In addition, patients in the surgical group had symptoms that were more bothersome and radiographic evidence of more severe stenosis. The surgical group was more often dissatisfied with their symptoms and more often rated the symptoms as worsening than did patients in the nonsurgical group.

The final models, combining both cohorts, were adjusted for age, sex, coexisting disorders of the stomach or joints, the presence or absence of pain on straight-leg raising or femoral-nerve tension signs, smoking status, patient-assessed health trend, income, other compensation, body-mass index, baseline score for the outcome variable, and center.

NONSURGICAL TREATMENTS

At 2 years, nonsurgical treatments were similar in the two cohorts. However, more patients in the randomized group than in the observational group reported visits to a surgeon (45% vs. 32%, $P=0.02$) and receiving injections (52% vs. 39%, $P=0.02$), whereas more patients in the observational group reported the use of "other" medications, such as gabapentin (60% vs. 73%, $P=0.01$).

SURGICAL TREATMENTS AND COMPLICATIONS

Overall, surgical treatments and complications were similar in the two cohorts (Table 2). Among patients in the surgical group, 89% underwent decompression only. Instrumented fusion was performed in only 6% of patients. The median surgical time was 120 minutes, with a mean blood loss of 314 ml; 10% of patients required transfusions intraoperatively and 5% postoperatively. The most common surgical complication was dural tear, in 9% of patients. At 2 years, reoperation had occurred in 8% of patients; fewer than half of these operations were for recurrent stenosis.

At 2 years, there were seven deaths in the nonsurgical group and six in the surgical group, one of which occurred within 3 months after surgery. The deaths were reviewed and 12 were judged not to be treatment-related. The one death of unknown cause occurred 501 days after surgery.

CROSSOVER

Nonadherence to treatment assignment affected both study cohorts: some patients in the surgical group chose to delay or decline surgery, and some in the nonsurgical group crossed over to undergo surgery (Fig. 1). The characteristics of crossover patients that differed significantly from patients who did not cross over are shown in Table 3. Patients in the nonsurgical group who crossed

Table 1. Demographic Characteristics, Coexisting Illnesses, and Measures of Health Status of the Patients.*

Variable	SPORT Study Cohort			Combined Randomized and Observational Cohorts		
	Randomized Group (N=278)	Observational Group (N=356)	P Value	Surgical Group (N=394)	Nonsurgical Group (N=240)	P Value
Age — yr	65.5±10.5	63.9±12.5	0.10	63.6±12.2	66.3±10.5	0.004
Female sex — no. (%)	106 (38)	143 (40)	0.66	152 (39)	97 (40)	0.71
Race or ethnic background — no. (%)†						
Non-Hispanic	259 (93)	346 (97)	0.03	378 (96)	227 (95)	0.55
White	238 (86)	295 (83)	0.41	332 (84)	201 (84)	0.95
Attended college — no. (%)	176 (63)	225 (63)	0.96	245 (62)	156 (65)	0.53
Married — no. (%)	197 (71)	249 (70)	0.87	288 (73)	158 (66)	0.06
Employment status — no. (%)			0.12			0.05
Full-time or part-time	88 (32)	128 (36)		144 (37)	72 (30)	
Disabled	24 (9)	36 (10)		40 (10)	20 (8)	
Retired	144 (52)	152 (43)		167 (42)	129 (54)	
Other	22 (8)	40 (11)		43 (11)	19 (8)	
Disability compensation — no. (%)‡	21 (8)	27 (8)	0.89	30 (8)	18 (8)	0.92
Body-mass index§	29.8±5.6	29.3±5.6	0.31	29.3±5.3	29.9±6.1	0.25
Current smoker — no. (%)	34 (12)	28 (8)	0.09	36 (9)	26 (11)	0.58
Coexisting condition — no. (%)						
Hypertension	134 (48)	154 (43)	0.25	168 (43)	120 (50)	0.09
Diabetes	50 (18)	46 (13)	0.10	53 (13)	43 (18)	0.16
Osteoporosis	22 (8)	38 (11)	0.30	30 (8)	30 (12)	0.06
Heart disorder	80 (29)	85 (24)	0.19	95 (24)	70 (29)	0.19
Stomach disorder	60 (22)	79 (22)	0.93	82 (21)	57 (24)	0.44
Bowel or intestinal disorder	36 (13)	50 (14)	0.78	49 (12)	37 (15)	0.35
Depression	36 (13)	34 (10)	0.22	41 (10)	29 (12)	0.60
Joint disorder	158 (57)	188 (53)	0.35	210 (53)	136 (57)	0.46
Other disorder¶	95 (34)	125 (35)	0.87	136 (35)	84 (35)	0.97
Symptom duration >6 mo — no. (%)	158 (57)	210 (59)	0.64	236 (60)	132 (55)	0.26
SF-36 score‖						
Bodily pain	31.9±17.5	31.4±17.4	0.73	28.6±16.2	36.6±18.4	<0.001
Physical function	35.4±22.6	34.3±23.8	0.55	31.7±21.9	39.9±24.5	<0.001
Mental component summary	49.8±12.4	49.1±11.6	0.47	48.5±12.0	50.9±11.7	0.02
Oswestry Disability Index**	42.7±17.9	42.1±19.0	0.70	46.0±17.9	36.4±17.9	<0.001
Stenosis Frequency Index††	13.5±5.7	14.2±5.8	0.13	15.2±5.6	11.8±5.6	<0.001
Stenosis Bothersomeness Index‡‡	13.9±5.7	14.7±5.8	0.08	15.6±5.4	12.3±5.7	<0.001
Low Back Pain Bothersomeness Scale§§	4.0±1.9	4.2±1.8	0.19	4.3±1.8	3.8±1.8	0.002
Leg Pain Bothersomeness Scale¶¶	4.3±1.7	4.4±1.7	0.44	4.6±1.6	3.9±1.8	<0.001
Patient very dissatisfied with symptoms — no. (%)	183 (66)	250 (70)	0.27	309 (78)	124 (52)	<0.001

Table 1. (Continued.)

Variable	SPORT Study Cohort			Combined Randomized and Observational Cohorts		
	Randomized Group (N=278)	Observational Group (N=356)	P Value	Surgical Group (N=394)	Nonsurgical Group (N=240)	P Value
Patient's self-assessed health trend — no. (%)			0.48			<0.001
Problem getting better	18 (6)	28 (8)		14 (4)	32 (13)	
Problem staying about the same	95 (34)	108 (30)		108 (27)	95 (40)	
Problem getting worse	160 (58)	218 (61)		265 (67)	113 (47)	
Treatment preference — no. (%)			<0.001			<0.001
Nonsurgical						
Definitely	37 (13)	86 (24)		36 (9)	87 (36)	
Probably	61 (22)	45 (13)		36 (9)	70 (29)	
Not sure	95 (34)	26 (7)		61 (15)	60 (25)	
Surgical						
Definitely	33 (12)	163 (46)		188 (48)	8 (3)	
Probably	51 (18)	36 (10)		73 (19)	14 (6)	
Signs and symptoms — no. (%)						
Neurogenic claudication	219 (79)	289 (81)	0.51	317 (80)	191 (80)	0.87
Pain on straight-leg raising or femoral-nerve tension sign	41 (15)	91 (26)	0.001	85 (22)	47 (20)	0.62
Dermatomal pain radiation	215 (77)	284 (80)	0.52	310 (79)	189 (79)	0.94
Any neurologic deficit	146 (53)	203 (57)	0.29	210 (53)	139 (58)	0.29
Asymmetric depressed reflexes	76 (27)	92 (26)	0.74	102 (26)	66 (28)	0.72
Asymmetric decrease in sensory reflexes	68 (24)	114 (32)	0.05	116 (29)	66 (28)	0.66
Asymmetric motor weakness	71 (26)	106 (30)	0.28	104 (26)	73 (30)	0.32
Stenosis level — no. (%)						
L2–L3	77 (28)	102 (29)	0.86	121 (31)	58 (24)	0.09
L3–L4	183 (66)	237 (67)	0.91	266 (68)	154 (64)	0.44
L4–L5	255 (92)	324 (91)	0.86	362 (92)	217 (90)	0.63
L5–S1	72 (26)	101 (28)	0.55	100 (25)	73 (30)	0.20
Moderate or severe stenotic levels — no. (%)			0.45			0.19
0	4 (1)	11 (3)		6 (2)	9 (4)	
1	106 (38)	128 (36)		140 (36)	94 (39)	
2	109 (39)	132 (37)		153 (39)	88 (37)	
≥3	59 (21)	85 (24)		95 (24)	49 (20)	
Location of stenosis — no. (%)						
Central	241 (87)	302 (85)	0.58	338 (86)	205 (85)	0.99
Lateral recess	236 (85)	267 (75)	0.003	321 (81)	182 (76)	0.11
Neuroforamen	88 (32)	119 (33)	0.70	119 (30)	88 (37)	0.11

Table 1. (Continued.)

Variable	SPORT Study Cohort			Combined Randomized and Observational Cohorts		
	Randomized Group (N=278)	Observational Group (N=356)	P Value	Surgical Group (N=394)	Nonsurgical Group (N=240)	P Value
Severity of stenosis — no. (%)			0.24			0.006
Mild	4 (1)	11 (3)		6 (2)	9 (4)	
Moderate	131 (47)	151 (42)		161 (41)	121 (50)	
Severe	143 (51)	194 (54)		227 (58)	110 (46)	
Spinal instability	0	0		0	0	

* Plus–minus values are means ±SD. Patients in the combined two cohorts were classified according to whether they received surgical treatment or nonsurgical treatment during the first 2 years of enrollment. Numbers of patients include only those who completed at least one follow-up survey within 2 years after enrollment.
† Race or ethnic group was self-reported. Whites and blacks could be either Hispanic or non-Hispanic.
‡ This category includes patients who were receiving or had applications pending for workers' compensation, Social Security benefits, or other compensation.
§ The body-mass index is the weight in kilograms divided by the square of the height in meters.
¶ Other disorders included problems related to stroke, cancer, lung disorders, fibromyalgia, chronic fatigue syndrome, post-traumatic stress disorder, alcohol or drug dependency, migraine, anxiety, or disorders of the liver, kidney, blood vessels, or nervous system.
‖ Scores on the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36) range from 0 to 100, with higher scores indicating less severe symptoms.
** The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms.
†† The Stenosis Frequency Index ranges from 0 to 24, with lower scores indicating less severe symptoms.
‡‡ The Stenosis Bothersomeness Index ranges from 0 to 24, with lower scores indicating less severe symptoms.
§§ The Low Back Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.
¶¶ The Leg Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.

over to undergo surgery had more self-rated disability, more psychological distress, worse symptoms, and a stronger treatment preference for surgery at baseline than did patients who did not opt for surgery. Patients in the surgical group who crossed over to receive nonsurgical care were more often not white, had less bothersome symptoms, less often rated their symptoms as worsening at enrollment, and had a stronger treatment preference for nonsurgical care at baseline.

MAIN TREATMENT EFFECTS

In the intention-to-treat analysis, a significant treatment effect favoring surgery was seen at 2 years, with a mean difference in change from baseline of 7.8 (95% confidence interval [CI], 1.5 to 14.1) on the SF-36 scale for bodily pain; at earlier times, there was a smaller nonsignificant effect in favor of surgery. However, at 2 years, there were no significant differences between the surgical group and the nonsurgical group on the SF-36 scale for physical function (0.1; 95% CI, −6.4 to 6.5) or on the Oswestry Disability Index (−3.5; 95% CI, −8.7 to 1.7) (Table 4).

In the as-treated analysis, the mean differences in change from baseline in the randomized and observational cohorts were similar at 2 years: bodily pain, 11.7 (95% CI, 6.2 to 17.2) in the randomized group versus 15.3 (95% CI, 10.4 to 20.2) in the observational group; physical function, 8.1 (95% CI, 2.8 to 13.5) in the randomized group versus 13.6 (95% CI, 8.7 to 18.4) in the observational group; and Oswestry Disability Index, −8.7 (95% CI, −13.3 to −4.0) in the randomized group versus −13.1 (95% CI, −16.9 to −9.2) in the observational group (Fig. 2).

The global hypothesis test comparing the as-treated effects in the randomized group and the observational group over all time periods showed no difference between the two cohorts ($P=0.93$ for bodily pain, $P=0.67$ for physical function, and $P=0.60$ for the Oswestry Disability Index).

Results from the intention-to-treat analysis and the as-treated analysis of the two cohorts are compared in Figure 2. The effects shown in the as-treated analysis significantly favored surgery in both cohorts. In the combined analysis, treatment effects were significant in favor of sur-

Table 2. Surgical Treatments, Complications, and Events.*

Variable	Randomized Cohort (N=155)	Observational Cohort (N=239)	P Value
Procedure — no./total no. (%)			0.49
Decompression only	137/154 (89)	209/235 (89)	
Noninstrumented fusion	6/154 (4)	14/235 (6)	
Instrumented fusion	11/154 (7)	12/235 (5)	
Multilevel fusion — no./total no. (%)	5/155 (3)	11/239 (5)	0.68
Decompression level — no./total no. (%)			
L2–L3	53/152 (35)	90/235 (38)	0.57
L3–L4	115/152 (76)	157/235 (67)	0.081
L4–L5	140/152 (92)	218/235 (93)	0.97
L5–S1	60/152 (39)	89/235 (38)	0.83
Levels decompressed — no./total no. (%)			0.92
None	3/155 (2)	4/239 (2)	
1	33/155 (21)	54/239 (23)	
2	47/155 (30)	78/239 (33)	
≥3	72/155 (46)	103/239 (43)	
Operation time — min	128.4±64.7	127.8±66.2	0.93
Blood loss — ml	338.5±527.1	295.6±312.6	0.31
Blood replacement — no./total no. (%)			
Intraoperative transfusion	14/152 (9)	23/238 (10)	0.98
Postoperative transfusion	6/153 (4)	13/238 (5)	0.65
No. of days in hospital	3.5±2.6	3.0±2.2	0.13
Postoperative mortality — no./total no. (%)†			
Within 6 wk	0/155	1/239 (<1)	0.83
Within 3 mo	0/155	1/239 (<1)	0.83
Intraoperative complications — no./total no. (%)‡			
Dural tear or spinal fluid leak	13/155 (8)	23/238 (10)	0.80
Other	1/155 (1)	2/238 (1)	0.71
None	141/155 (91)	213/238 (89)	0.76
Postoperative complications or events — no./total no. (%)§			
Wound hematoma	3/153 (2)	1/238 (<1)	0.34
Wound infection	3/153 (2)	5/238 (2)	0.79
Other	8/153 (5)	13/238 (5)	0.90
None	135/153 (88)	208/238 (87)	0.93

Table 2. (Continued.)

Variable	Randomized Cohort (N=155)	Observational Cohort (N=239)	P Value
Additional surgery — no./total no. (%)¶			
Any surgery			
At 1 yr	6/157 (4)	15/243 (6)	0.29
At 2 yr	10/157 (6)	21/243 (9)	0.39
Recurrent stenosis or progressive spondylolisthesis			
At 1 yr	3/155 (2)	2/241 (1)	
At 2 yr	6/155 (4)	5/241 (2)	
Pseudarthrosis or fusion exploration			
At 1 yr	0/155	0/239	
At 2 yr	0/155	0/239	
Complication or other problem			
At 1 yr	3/155 (2)	10/241 (4)	
At 2 yr	4/155 (3)	11/241 (5)	
New condition			
At 1 yr	0/155	2/241 (1)	
At 2 yr	1/155 (1)	5/241 (2)	

* Plus–minus values are means ±SD. A total of 157 patients in the randomized cohort and 243 patients in the observational cohort underwent surgery. Data are missing for patients in some categories, as indicated by varying denominators.
† One patient had a myocardial infarction.
‡ None of the following were reported: aspiration, nerve-root injury, operation at wrong level, and vascular injury.
§ This category includes all reported complications up to 8 weeks after surgery. None of the following were reported: bone-graft complication, cerebrospinal fluid leak, paralysis, cauda equina injury, wound dehiscence, pseudarthrosis, and nerve-root injury.
¶ Rates of repeated surgery at 1 and 2 years are Kaplan–Meier estimates. P values were calculated with the use of the log-rank test.

gery for all primary and secondary outcome measures at each time point during the 2 years (Table 4).

DISCUSSION

In patients with imaging-confirmed spinal stenosis without spondylolisthesis and leg symptoms persisting for at least 12 weeks, surgery was superior to nonsurgical treatment in relieving symptoms and improving function. In the as-treated analysis, the treatment effect for surgery was seen as early as 6 weeks, appeared to reach a maximum at 6 months, and persisted for 2 years; it is notable that the condition of patients in the nonsurgical group improved only moderately during the 2-year period. The intention-to-treat results must be viewed in the context of the substantial rates of nonadherence to assigned treatment. The pattern of nonadherence was striking because both the surgical and the nonsurgical groups were affected, unlike the results of many studies involving surgical procedures.[30] The mixing of treatments owing to crossover can be expected to create a bias toward the null.[31] The large effects seen in the as-treated analysis and the characteristics of the crossover patients suggest that the intention-to-treat analysis underestimated the true effect of surgery.

This study provides an opportunity to compare

Table 3. Significant Predictors of Treatment Received within 2 Years among Patients in the Randomized Cohort.*

Predictor	Assigned to Surgical Group			Assigned to Nonsurgical Group		
	Surgery (N=89)	No Surgery (N=43)	P Value	Surgery (N=64)	No Surgery (N=82)	P Value
White race — no. (%)	80 (90)	29 (67)	0.003	58 (91)	71 (87)	0.62
Mental Component Summary score†	49.9±12.2	50.4±14.0	0.84	47.0±12.6	51.6±11.2	0.02
Oswestry Disability Index‡	44.6±18.2	38.9±18.8	0.10	46.7±18	39.4±16.2	0.01
Stenosis Frequency Index§	14.8±5.3	11.6±6.2	0.002	14.7±5.5	12.1±5.5	0.005
Stenosis Bothersomeness Index¶	15.0±4.9	12.0±6.0	0.002	15.3±5.5	12.6±6.1	0.005
Leg Pain Bothersomeness Scale‖	4.5±1.6	4.0±1.9	0.09	4.6±1.5	3.9±1.8	0.01
Very dissatisfied with symptoms — no. (%)	65 (73)	25 (58)	0.13	49 (77)	44 (54)	0.007
Problem getting better or worse — no. (%)			0.009			0.31
Getting better	2 (2)	6 (14)		2 (3)	8 (10)	
Staying about the same	27 (30)	18 (42)		22 (34)	28 (34)	
Getting worse	57 (64)	19 (44)		38 (59)	46 (56)	
Treatment preference — no. (%)			0.01			<0.001
Nonsurgical						
Definitely	9 (10)	8 (19)		7 (11)	13 (16)	
Probably	15 (17)	15 (35)		9 (14)	22 (27)	
Not sure	31 (35)	13 (30)		15 (23)	36 (44)	
Surgical						
Definitely	11 (12)	0		17 (27)	5 (6)	
Probably	23 (26)	7 (16)		16 (25)	5 (6)	

* Plus–minus values are means ±SD. Numbers of patients include only those who completed at least one follow-up survey within 2 years after enrollment.
† Scores on the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36) range from 0 to 100, with higher scores indicating less severe symptoms.
‡ The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms.
§ The Stenosis Frequency Index ranges from 0 to 24, with lower scores indicating less severe symptoms.
¶ The Stenosis Bothersomeness Index ranges from 0 to 24, with lower scores indicating less severe symptoms.
‖ The Leg Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.

results involving patients who were willing to participate in a randomized study (randomized cohort) and those who were unwilling to participate in such a study (observational cohort).[13-16] These two cohorts were remarkably similar at baseline. Other than treatment preference, the only significant differences were small ones in signs of nerve-root tension and the location of stenosis. The two cohorts also had similar outcomes, without significant differences in the as-treated analyses. Given these similarities, the combined analyses are well justified. Although these analyses are not based on randomized treatment assignments, the results are strengthened by the use of specific inclusion and exclusion criteria, the sample size, and adjustment for potentially confounding baseline differences.[32]

The characteristics of the patients were similar to those in previous studies, even though the latter involved mixed-cohort patients (i.e., those with or without spondylolisthesis). In our study, the functional status of the patients at baseline was similar to that of patients in the Maine Lumbar Spine Study[7,8] (SF-36 score, 34.8 and 35.0, respectively) but worse than that in the study by Malmivaara et al.[10,11] (Oswestry Disability Index, 42.4 and 35.0, respectively).

In the as-treated analysis, the functional improvement in the surgical group at 1 year was very similar to that in the Maine Lumbar Spine

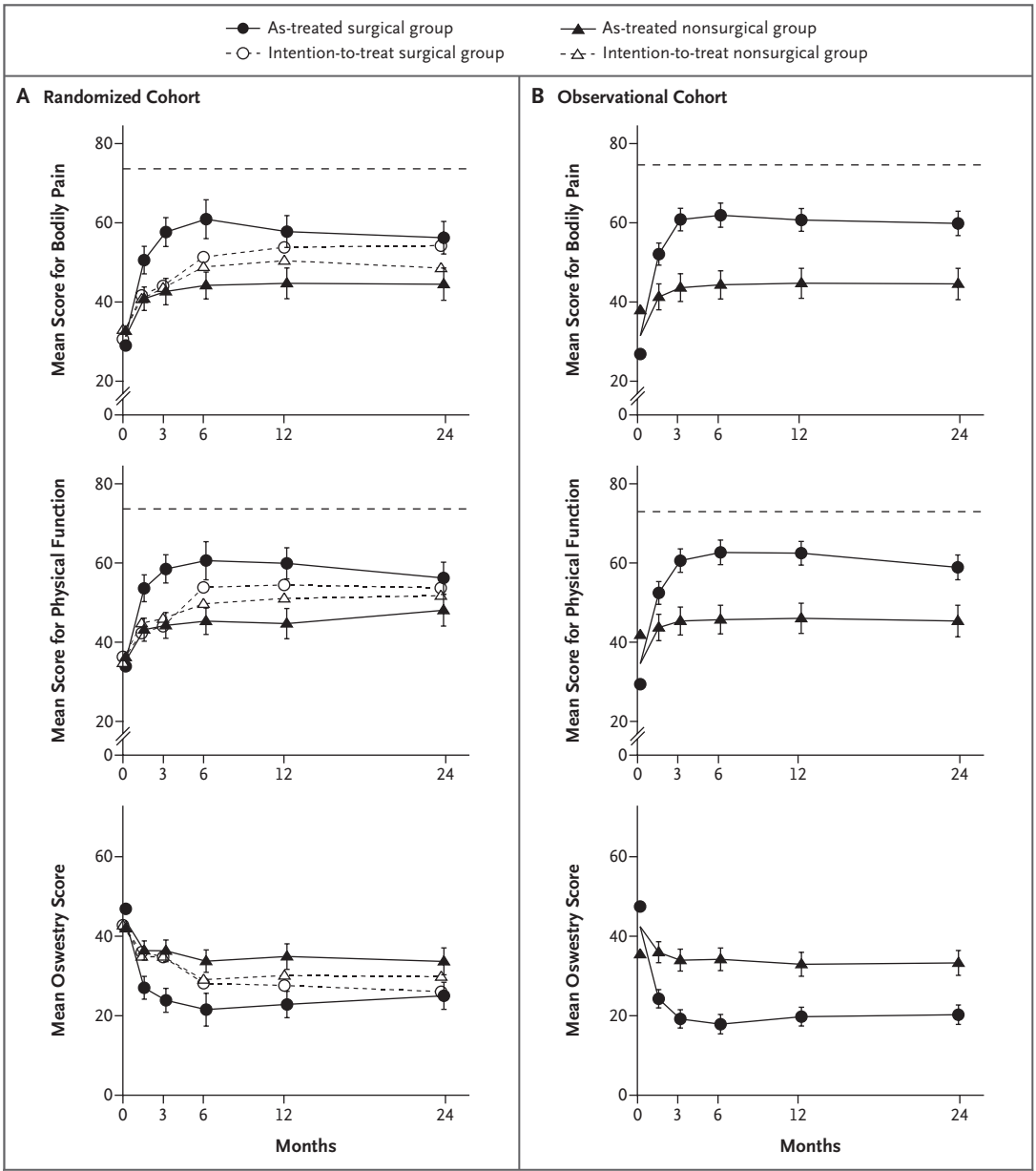


Figure 2. Primary Outcomes in the Randomized and Observational Cohorts during 2 Years of Follow-up.

The graphs show both the intention-to-treat and the as-treated analyses for the randomized cohort (column on left) and the as-treated analysis for the observational cohort (column on right). Results for bodily pain and physical function are scores on the Medical Outcomes Study 36-item Short-Form General Health Survey (SF-36), ranging from 0 to 100, with higher score indicating less severe symptoms. The Oswestry Disability Index (bottom row) ranges from 0 to 100, with lower scores indicating less severe symptoms. The horizontal dashed line in each of the four SF-36 graphs represents normal values adjusted for age and sex. The I bars represent 95% confidence intervals. At 0 months, the floating data points represent the observed mean scores for each study group, whereas the data points on plot lines represent the overall means used in the adjusted analyses.

Table 4. Intention-to-Treat Analysis for the Randomized Cohort and Adjusted Analyses, According to Treatment Received, for the Randomized

Outcome	Baseline	At 6 Wk			At 3 Mo		
	Overall Mean	Surgery	No Surgery	Treatment Effect (95% CI)†	Surgery	No Surgery	Treatment Effect (95% CI)†
		Mean Change			*Mean Change*		
Intention-to-treat analysis of randomized cohort							
No. of patients	278	116	130		116	135	
Primary outcome							
SF-36‡							
Bodily pain	31.9±1.1	11.2±2.4	7.9±2.2	3.3 (–3.2 to 9.8)	13.5±2.5	11.1±2.3	2.4 (–4.2 to 9.1)
Physical function	35.4±1.4	6.0±2.5	10.2±2.3	–4.2 (–10.8 to 2.4)	7.4±2.5	11.6±2.3	–4.2 (–10.9 to 2.6)
Oswestry Disability Index§	42.7±1.1	–6.5±2.0	–7.9±1.8	1.4 (–3.9 to 6.8)	–7.6±2.1	–8.1±1.9	0.5 (–5.0 to 6.0)
Secondary outcome							
Stenosis Bothersomeness Index¶	13.9±0.3	NA	NA	NA	–4.5±0.8	–3.8±0.7	–0.7 (–2.8 to 1.5)
Leg Pain Bothersomeness Index‖	4.3±0.1	NA	NA	NA	–1.5±0.3	–1.2±0.2	–0.3 (–1.0 to 0.4)
Low Back Pain Bothersomeness Index**	4.0±0.1	NA	NA	NA	–0.6±0.2	–1.0±0.2	0.4 (–0.2 to 1.0)
Satisfaction of patient (%)							
Very or somewhat satisfied with symptoms	5.0	22.0	27.9	–5.9 (–18.0 to 6.2)	38.1	33.7	4.3 (–9.6 to 18.2)
Very or somewhat satisfied with care	NA	68.7	71.8	–3.1 (–15.9 to 9.7)	71.6	77.0	–5.4 (–18.0 to 7.2)
Self-rated major improvement in condition	NA	19.7	25.8	–6.1 (–18.0 to 5.9)	30.9	35.4	–4.5 (–18.2 to 9.3)
As-treated analysis of randomized and observational cohorts††							
No. of patients	803	398	370		378	313	
Primary outcome							
SF-36‡							
Bodily pain	31.4±0.6	19.8±1.1	9.8±1.1	10.0 (7.3 to 12.7)	27.9±1.1	11.8±1.2	16.1 (13.1 to 19.1)
Physical function	34.9±0.8	17.8±1.1	8.7±1.1	9.1 (6.5 to 11.8)	24.8±1.2	10.0±1.2	14.8 (11.9 to 17.7)
Oswestry Disability Index§	43.2±0.6	–17.0±0.9	–6.8±0.9	–10.3 (–12.5 to –8.1)	–21.4±0.9	–7.6±1.0	–13.8 (–16.2 to –11.3)
Secondary outcome							
Stenosis Bothersomeness Index¶	14.4±0.2	NA	NA	NA	–8.6±0.4	–2.9±0.4	–5.7 (–6.7 to –4.6)
Leg Pain Bothersomeness Scale‖	4.3±0.1	NA	NA	NA	–2.7±0.1	–0.9±0.1	–1.8 (–2.1 to –1.5)
Low Back Pain Bothersomeness Scale**	4.1±0.1	NA	NA	NA	–2.0±0.1	–0.8±0.1	–1.2 (–1.5 to –0.9)
Satisfaction of patient (%)							
Very or somewhat satisfied with symptoms	5.7	64.0	18.5	45.5 (39.2 to 51.8)	61.9	22.8	39.1 (32.0 to 46.1)
Very or somewhat satisfied with care	NA	90.1	67.5	22.7 (16.8 to 28.5)	90.1	71.3	18.7 (12.6 to 24.9)
Self-rated major improvement in condition	NA	72.0	18.2	53.8 (48.0 to 59.6)	72.0	20.0	52.0 (45.6 to 58.3)

* Plus–minus values are means ±SE. Values in the as-treated analysis have been adjusted for age, sex, the presence or absence of stomach or joint disorders, the presence or absence of pain on straight-leg raising or femoral-nerve tension signs, smoking status, patient-assessed health trend, income, other compensation, body-mass index, baseline score for the outcome variable, and center. NA denotes not available.
† The treatment effect is the difference in the mean change from baseline between the surgical group and the nonsurgical group.
‡ The SF-36 scores range from 0 to 100, with higher scores indicating less severe symptoms.
§ The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms.

and Observational Cohorts Combined.*

At 6 Mo			At 1 Yr			At 2 Yr		
Surgery	No Surgery	Treatment Effect (95% CI)†	Surgery	No Surgery	Treatment Effect (95% CI)†	Surgery	No Surgery	Treatment Effect (95% CI)†
Mean Change			*Mean Change*			*Mean Change*		
120	135		120	126		108	113	
21.0±2.2	16.1±2.1	4.9 (–1.2 to 10.9)	23±2.3	17.5±2.2	5.5 (–0.7 to 11.7)	23.4±2.3	15.6±2.2	7.8 (1.5 to 14.1)
17.6±2.3	15.1±2.2	2.5 (–3.7 to 8.6)	18.0±2.3	16.4±2.2	1.6 (–4.8 to 7.9)	17.1±2.4	17.1±2.3	0.1 (–6.4 to 6.5)
–14.6±1.9	–13.7±1.7	–0.9 (–5.9 to 4.1)	–14.9±1.9	–12.7±1.8	–2.2 (–7.4 to 2.9)	–16.4±1.9	–12.9±1.8	–3.5 (–8.7 to 1.7)
NA	NA	NA	–6.1±0.7	–4.9±0.7	–1.2 (–3.2 to 0.8)	–6.3±0.7	–5.6±0.7	–0.7 (–2.7 to 1.3)
NA	NA	NA	–2.3±0.2	–1.7±0.2	–0.6 (–1.3 to 0)	–2.2±0.2	–1.8±0.2	–0.3 (–1.0 to 0.3)
NA	NA	NA	–1.3±0.2	–1.3±0.2	0 (–0.5 to 0.6)	–1.3±0.2	–1.6±0.2	0.3 (–0.2 to 0.9)
49.9	38.3	11.7 (–0.8 to 24.1)	53.4	40.7	12.8 (–0.3 to 25.8)	53.1	45.0	8.1 (–5.3 to 21.5)
85.0	73.1	11.9 (1.9 to 21.9)	81.4	69.3	12.0 (0.8 to 23.2)	76.2	68.5	7.6 (–4.5 to 19.7)
49.0	46.2	2.8 (–10.0 to 15.5)	55.3	44.2	11.1 (–2.1 to 24.2)	50.7	45.0	5.8 (–7.8 to 19.3)
256	271		302	230		335	198	
29.5±1.3	12.9±1.2	16.6 (13.3 to 19.9)	28.0±1.2	13.5±1.4	14.6 (11.2 to 18)	26.9±1.2	13.3±1.4	13.6 (10.0 to 17.2)
26.9±1.3	10.6±1.3	16.3 (13.1 to 19.6)	26.5±1.2	10.5±1.4	15.9 (12.6 to 19.3)	23.0±1.3	11.8±1.4	11.1 (7.6 to 14.7)
–22.9±1.0	–8.8±1.0	–14.1 (–16.8 to –11.4)	–21.4±1.0	–8.9±1.1	–12.5 (–15.3 to –9.8)	–20.5±1.0	–9.3±1.2	–11.2 (–14.1 to –8.3)
NA	NA	NA	–8.2±0.3	–3.9±0.4	–4.4 (–5.4 to –3.3)	–7.8±0.4	–4.4±0.4	–3.4 (–4.5 to –2.3)
NA	NA	NA	–2.6±0.1	–1.4±0.1	–1.2 (–1.5 to –0.9)	–2.5±0.1	–1.4±0.1	–1.1 (–1.4 to –0.7)
NA	NA	NA	–2.0±0.1	–1.0±0.1	–1.0 (–1.3 to –0.7)	–2.0±0.1	–1.1±0.1	–1.0 (–1.3 to –0.6)
62.8	26.3	36.5 (28.6 to 44.4)	67.6	29.2	38.4 (30.3 to 46.5)	68.2	29.6	38.7 (30.0 to 47.3)
90.8	67.3	23.5 (16.7 to 30.3)	86.1	69.1	17.0 (9.6 to 24.4)	83.1	67.0	16.1 (7.8 to 24.3)
67.0	24.9	42.1 (34.5 to 49.6)	68.8	25.4	43.4 (35.7 to 51.1)	62.9	28.7	34.1 (25.6 to 42.6)

¶ The Stenosis Bothersomeness Index ranges from 0 to 24, with lower scores indicating less severe symptoms.
‖ The Leg Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.
** The Low Back Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.
†† The number of patients in the as-treated analyses reflects the number of patients contributing to the estimate in a given period with the use of the longitudinal-modeling strategy (explained in the Methods section) and may not correspond to the number shown for each visit time in Figure 1.

Study (26.5 and 27.0, respectively) but greater than in the study by Malmivaara et al. (Oswestry Disability Index, –21.4 and –11.3, respectively). Functional improvement in the nonsurgical group was greater in our study than in the previous studies, with a change of 10.5 in the SF-36 physical function score at 1 year, as compared with 1.0 in the Maine Lumbar Spine Study, and a change of 9.3 in the Oswestry Disability Index at 2 years, as compared with 4.5 in the study by Malmivaara et al. The greater improvements in our study, compared with those in the study by Malmivaara et al., may be related to differences in the selection of patients. In the study by Malmivaara et al., patients with moderate spinal stenosis were specifically selected, whereas in our study, we attempted to enroll patients with spinal stenosis who were surgical candidates.

In the as-treated analysis, we can directly compare the estimates of treatment effect with those of the previous studies. The estimated 1-year treatment effects for surgery were smaller in our study than in the Maine Lumbar Spine Study (changes in bodily pain of 14.6 and 30.4, respectively, and in physical function of 15.9 and 25.5, respectively). However, in the Maine Lumbar Spine Study, treatment effects for baseline differences between the study groups were not adjusted, which probably explains these discrepancies. At 1 year, the estimated treatment effects were similar in our study and the study by Malmivaara et al.: Oswestry Disability Index, –12.5 and –11.3, respectively; leg pain, 17% (on a 7-point scale) and 15% (on an 11-point scale); and back pain, 14% (on a 7-point scale) and 21% (on an 11-point scale).

It is interesting that among patients who underwent surgery, the magnitude of the mean changes in patients with spinal stenosis was nearly identical to that in the patients with degenerative spondylolisthesis at 2 years: bodily pain, 26.9 and 29.9, respectively; physical function, 23.0 and 26.6; Oswestry Disability Index, –20.5 and –24.2; and bothersomeness of symptoms, –7.8 and –8.9.[16] The treatment effects in these studies of spinal stenosis were larger than those in the observational study of patients with intervertebral disk herniation because of strong improvements in the nonsurgical group of patients with intervertebral disk herniation that were not seen in either stenosis group.[14-16]

There was little evidence of harm from either treatment. Often patients fear they will get worse without surgery, but this was not the case for the majority of patients in the nonsurgical group, who, on average, showed small improvements in all outcomes. The 1-year rate of reoperation for recurrent stenosis was 1.3%, a rate similar to those reported by Malmivaara et al. (2%) and by the Maine Lumbar Spine Study (1.2%). At 2 years, mortality was nearly the same in the two study groups and was lower than actuarial projections. The postoperative death rate of 0.3% and the overall postoperative complication rate of 12% were slightly better than the reported Medicare rates in patients with spinal stenosis who did not undergo spinal fusion (death rate, 0.8%; rate of complications, 14%).[1] However, higher rates of complications have been reported with increasing age and coexisting medical conditions.[33]

The primary limitation of our study was the marked degree of nonadherence to randomized treatment. This factor reduced the power of the intention-to-treat analysis to show treatment effects, though there was still a significant treatment effect for the measure of bodily pain at 2 years. The as-treated analyses do not share the strong protection from confounding that exists for the intention-to-treat analyses. However, these analyses were carefully adjusted for important baseline covariates and yielded results similar to those of previous studies. The characteristics of the crossover patients were as one might expect: those with severe symptoms and a preference for surgery crossed over into the surgical group, and vice versa.

Another limitation was the heterogeneity of the nonsurgical treatments. Given the limited evidence regarding efficacy of most nonsurgical treatments for spinal stenosis and individual variability in response, the creation of a limited, fixed protocol for nonsurgical treatment was neither clinically feasible nor generalizable. The flexible treatment protocols allowed for individualization of nonsurgical treatment plans, reflect current practice among multidisciplinary spine practices, and were consistent with published guidelines.[34,35] However, we did not assess the effect of surgery versus any specific nonsurgical treatment.

In conclusion, in the as-treated analysis, if we combine the randomized and observational co-

horts, carefully adjusting for potentially confounding baseline factors, patients with spinal stenosis without degenerative spondylolisthesis who underwent surgery showed significantly greater improvement in pain, function, satisfaction, and self-rated progress than did patients who were treated nonsurgically.

Supported by a grant (U01-AR45444-01A1) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Institutes of Health Office of Research on Women's Health, the National Institute of Occupational Safety and Health of the Centers for Disease Control and Prevention, a grant (P60-AR048094-01A1) to the Multidisciplinary Clinical Research Center in Musculoskeletal Diseases from NIAMS, and a Research Career Award (1-K23-AR-048138-01, to Dr. Lurie) from NIAMS.

Dr. Lurie reports receiving grant support from St. Francis Medical Technologies and the American Board of Orthopaedic Surgery and consulting fees from Merck, Ortho-McNeil, Pfizer, Centocor, Myexpertdoctor.com, Pacific Business Group on Health, and the Foundation for Informed Medical Decision Making; Dr. A.N.A. Tosteson, receiving grant support from St. Francis Medical Technologies and Zimmer; Dr. Cammisa, having an equity interest in K2M, Spinal Kinetics, and HealthPoint Capital Partners; Dr. Albert, receiving consulting fees and royalties from DePuy Spine and having an equity interest in K2M; Dr. Boden, receiving consulting fees from Medtronic and lecture fees from Osteotech; and Dr. Berven, receiving grant support from Medtronic. No other potential conflict of interest relevant to this article was reported.

We thank Tamara S. Morgan, Department of Orthopaedic Surgery, Dartmouth Medical School, for graphic design and assistance with the manuscript and the following members of the data and safety monitoring board: Ron Thisted, Ph.D. (chair), University of Chicago, Chicago; Tim Carey, M.D., M.P.H., University of North Carolina at Chapel Hill, Chapel Hill; Peter C. Gerszten, M.D., Presbyterian University Hospital, Pittsburgh; Ed Hanley, M.D., Carolina Health Care, Charlotte, NC; and Bjorn Ryedvik, M.D., Ph.D., Sahlgrenska University Hospital, Gothenburg, Sweden. This study is dedicated to the memory of Brieanna Weinstein.

APPENDIX

In addition to the authors, the following investigators participated in the study, with institutions listed in order from highest to lowest enrollment of patients: *William Beaumont Hospital, Royal Oak, MI:* G. Bradley, M. Lurie, J. Fischgrund, D. Montgomery, L. Kurz, E. Truumees; *Washington University, St. Louis:* L. Lenke, G. Stobbs, A. Margherita, H. Prather, K. Bridwell, K.S. Riew, C. Lauryssen, B. Taylor, J. Metzler; *Dartmouth Medical School, Lebanon, NH:* J. Forman, W. Abdu, B. Butler-Schmidt, J.J. Hebb, P. Ball, P. Bernini, H. Magnadottir, R. Rose, R. Roberts, R. Diegel, S. Banerjee, R. Beasely; *Emory University, Atlanta:* S. Lashley, J. Heller, H. Levy, S.T. Yoon, M. Schaufele, W. Horton; *Rothman Institute at Thomas Jefferson Hospital, Philadelphia:* C. Simon, M. Freedman, O'Brien, S. Dante, T. Conliffe; *University Hospitals of Cleveland and Case Western Reserve University, Cleveland:* S. Emery, C. Furey, K. Higgins, J.X. Yoo, H. Bohlman, E.B. Marsolais, R.S. Krupkin; *Hospital for Special Surgery, New York:* B. Green, O. Boachie-Edjei, J. Farmer; *Nebraska Foundation for Spinal Research, Omaha:* M. Longley, N. Fullmer, A.M. Fredericks, J. Fuller, R. Woodward, J. McClellan, E. Phillips, T. Burd, P. Bowman; *University of California at San Francisco, San Francisco:* P. Malone, D. Bradford, S. Deviren, P. Weinstein, T. Smith; *Hospital for Joint Diseases, New York:* T. Errico, A. Lee, J. Goldstein, J. Spivak, R. Perry, J. Bendo, R. Moskovich; *Rush–Presbyterian–St. Luke's Medical Center, Chicago:* G. Andersson, M. Hickey, E. Goldberg, F. Phillips, R. Massimino, S. Petty; *Kaiser Permanente, Oakland, CA:* H. Goldberg; *Maine Spine and Rehabilitation, Scarborough:* R. Keller.

REFERENCES

1. 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;18:1463-70.

2. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. Spine 2005;30:1441-5.

3. 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.

4. Weinstein J, Birkmeyer J. The Dartmouth atlas of musculoskeletal health care. Chicago: American Hospital Association Press, 2000.

5. Boden SD, McCowin PR, Davis DO, Dina TS, Mark AS, Wiesel S. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects: a prospective investigation. J Bone Joint Surg Am 1990;72:403-8.

6. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331:69-73.

7. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part III: 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. Spine 1996;21:1787-94.

8. 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.

9. Johnsson KE, Uden A, Rosen I. The effect of decompression on the natural course of spinal stenosis: a comparison of surgically treated and untreated patients. Spine 1991;16:615-9.

10. Malmivaara A, Slatis P, Heliovaara M, et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine 2007;32:1-8.

11. Malmivaara A, Statis P, Heliovaara M, et al. Surgical treatment for moderate lumbar spinal stenosis: a randomized controlled trial. In: Proceedings of the International Society for Study of the Lumbar Spine, Porto, Portugal, May 30–June 5, 2004.

12. Mariconda M, Fava R, Gatto A, Longo C, Milano C. Unilateral laminectomy for bilateral decompression of lumbar spinal stenosis: a prospective comparative study with conservatively treated patients. J Spinal Disord Tech 2002;15:39-46.

13. Birkmeyer NJ, Weinstein JN, Tosteson AN, et al. Design of the Spine Patient Outcomes Research Trial (SPORT). Spine 2002; 27:1361-72.

14. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA 2006;296: 2451-9.

15. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA 2006;296:2441-50.

16. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonsurgical treat-

ment for lumbar degenerative spondylolisthesis. N Engl J Med 2007;356:2257-70.
17. 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.
18. Cummins J, Lurie JD, Tosteson TD, et al. Descriptive epidemiology and prior healthcare utilization of patients in the Spine Patient Outcomes Research Trial's (SPORT) three observational cohorts: disc herniation, spinal stenosis, and degenerative spondylolisthesis. Spine 2006;31:806-14.
19. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. Stat Med 1994;13:1341-52.
20. Phelan EA, Deyo RA, Cherkin DC, et al. Helping patients decide about back surgery: a randomized trial of an interactive video program. Spine 2001;26:206-11.
21. Weinstein JN. Partnership: doctor and patient: advocacy for informed choice vs. informed consent. Spine 2005;30:269-72.
22. 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.
23. Stewart AL, Greenfield S, Hays RD, et al. Functional status and well-being of patients with chronic conditions: results from the Medical Outcomes Study. JAMA 1989;262:907-13.
24. 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.
25. Ware JE Jr. SF-36 Health survey: manual and interpretation guide. Boston: Nimrod Press, 1993.
26. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine 2000;25:2940-52.
27. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. Spine 1986;11:28-30.
28. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine 1995;20:1899-908.
29. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. Philadelphia: Wiley–Interscience, 2004.
30. Kuppermann M, Varner RE, Summitt RL Jr, et al. Effect of hysterectomy vs medical treatment on health-related quality of life and sexual functioning: the medicine or surgery (Ms) randomized trial. JAMA 2004;291:1447-55.
31. Meinert CL. Clinical trials: design, conduct, and analysis. New York: Oxford University Press, 1986.
32. 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.
33. 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.
34. Acute low back problems in adults. Bethesda, MD: Agency for Health Care Policy and Research, 1994. (AHCPR publication no. 95-0642.)
35. Herniated disc. In: North American Spine Society phase III clinical guidelines for multidisciplinary spine care specialists. LaGrange, IL: North American Spine Society, 2000.

Copyright © 2008 Massachusetts Medical Society.

***JOURNAL* EDITORIAL FELLOW**

The *Journal*'s editorial office invites applications for a one-year research fellowship beginning in July 2009 from individuals at any stage of training. The editorial fellow will work on *Journal* projects and will participate in the day-to-day editorial activities of the *Journal* but is expected in addition to have his or her own independent projects. Please send curriculum vitae and research interests to the Editor-in-Chief, 10 Shattuck St., Boston, MA 02115 (fax, 617-739-9864), by September 30, 2008.