

## Long-Term Outcomes of Surgical and Nonsurgical Management of Lumbar Spinal Stenosis: 8 to 10 Year Results from the Maine Lumbar Spine Study

Steven J. Atlas, MD, MPH,\* Robert B. Keller, MD,† Yen A. Wu, MPH,\*  
Richard A. Deyo, MD, MPH,‡ and Daniel E. Singer, MD\*

**Study Design.** A prospective observational cohort study.

**Objective.** To assess long-term outcomes of patients with lumbar spinal stenosis treated surgically or nonsurgically.

**Summary of Background Data.** The relative benefit of various treatments for lumbar spinal stenosis is uncertain. Surgical treatment has been associated with short-term improvement, but recurrence of symptoms has been documented. Few studies have compared long-term outcomes of surgical and nonsurgical treatments.

**Methods.** Patients recruited from the practices of orthopaedic surgeons, neurosurgeons, and occupational medicine physicians throughout Maine had baseline interviews with follow-up questionnaires mailed at regular intervals over 10 years. Clinical data were obtained at baseline from a physician questionnaire. Most patients initially undergoing surgery had a laminectomy without fusion performed. Outcomes including patient-reported symptoms of leg and back pain, functional status, and satisfaction were assessed at 8- to 10-year follow-up. Primary analyses were based on initial treatment received with secondary analyses examining actual treatment received by 10 years.

**Results.** Of 148 eligible consenting patients initially enrolled, 105 were alive after 10 years (67.7% survival rate). Among surviving patients, long-term follow-up between 8 and 10 years was available for 97 of 123 (79%) patients (including 11 patients who died before the 10-year follow-up but completed a 8 or 9 year survey); 56 of 63 (89%) initially treated surgically and 41 of 60 (68%) initially treated nonsurgically. Patients undergoing surgery had worse baseline symptoms and functional status than those initially treated nonsurgically. Outcomes at 1 and 4 years favored initial surgical treatment. After 8 to 10 years, a similar percentage of surgical and nonsurgical patients reported that their low back pain was improved

(53% vs. 50%,  $P = 0.8$ ), their predominant symptom (either back or leg pain) was improved (54% vs. 42%,  $P = 0.3$ ), and they were satisfied with their current status (55% vs. 49%,  $P = 0.5$ ). These treatment group findings persisted after adjustment for other determinants of outcome in multivariate models. However, patients initially treated surgically reported less severe leg pain symptoms and greater improvement in back-specific functional status after 8 to 10 years than nonsurgically treated patients. By 10 years, 23% of surgical patients had undergone at least one additional lumbar spine operation, and 39% of nonsurgical patients had at least one lumbar spine operation. Patients undergoing subsequent surgical procedures had worse outcomes than those continuing with their initial treatment. Outcomes according to actual treatment received at 10 years did not differ because individuals undergoing additional surgical procedures had worse outcomes than those continuing with their initial treatment.

**Conclusions.** Among patients with lumbar spinal stenosis completing 8- to 10-year follow-up, low back pain relief, predominant symptom improvement, and satisfaction with the current state were similar in patients initially treated surgically or nonsurgically. However, leg pain relief and greater back-related functional status continued to favor those initially receiving surgical treatment. These results support a shared decision-making approach among physicians and patients when considering treatment options for lumbar spinal stenosis.

**Key words:** lumbar spinal stenosis, prospective cohort study, outcomes, lumbar disc surgery, natural history.  
**Spine 2005;30:936–943**

Lumbar spinal stenosis due to degenerative changes is a common cause of low back and leg pain in individuals starting in the fifth and sixth decades of life.<sup>1</sup> Rates of surgery for spinal stenosis have been increasing dramatically in the United States Medicare population.<sup>2,3</sup> In addition to increasing rates of surgery, there is wide variation in the likelihood that surgery will be performed among regions of the United States.<sup>1,2</sup> The rate of surgery for spinal stenosis has been shown to vary 12-fold, with higher rates in parts of the northwest, mountain, midwestern, and southern states.<sup>1</sup> Differences in rates of advanced spinal imaging account for a significant proportion of this variation.<sup>4</sup> A lack of clinical consensus about indications for spinal imaging and choice of treatment options is thought to explain the variation in the management of patients with lumbar spinal stenosis.<sup>1,4</sup> Although surgery is an accepted, commonly performed treatment for lumbar spinal stenosis, there is little evidence to support the relative benefits of surgery

From the \*General Medicine Division and the Clinical Epidemiology Unit, Medical Services, Massachusetts General Hospital, Harvard Medical School, Boston, MA; †the Center for the Evaluative Clinical Sciences, Dartmouth Medical School, Hanover, NH; and ‡the Center for Cost and Outcomes Research and the Departments of Medicine and Health Services, University of Washington, Seattle, WA.

Acknowledgment date: March 11, 2004. First revision date: May 14, 2004. Second revision date: June 23, 2004. Acceptance date: August 2, 2004.

Supported by grants from the Agency for Healthcare Research and Quality (HS-06344, HS-08194, and HS-09804).

The manuscript submitted does not contain information about medical device(s)/drug(s).

Federal funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Steven J. Atlas, MD, General Medicine Division, Massachusetts General Hospital, 50 Staniford Street, Boston, MA 02114. E-mail [satlas@partners.org](mailto:satlas@partners.org)

compared with nonsurgical treatments.<sup>5</sup> A single randomized study compared 13 surgically treated and 18 nonsurgically treated patients over a 10-year period.<sup>6</sup> As a result of the small sample size and high crossover rate from nonsurgical to surgical treatment (56% before the 4-year follow-up), no comparison of the relative benefits of surgical *versus* nonsurgical treatment was performed. Indirect comparisons suggested better short-term outcomes in surgically treated patients. Other randomized studies of patients with spinal stenosis are ongoing.<sup>7,8</sup> Preliminary short-term results favor surgical treatment,<sup>8</sup> but long-term results will not be available for a number of years. For individuals treated in contemporary clinical practice in the United States, the only prospective outcomes comparing surgical and nonsurgical treatment are from an observational study.<sup>9</sup> The Maine Lumbar Spine Study has previously shown that the relative benefit favoring surgical treatment after 1 year narrowed but persisted at 4 years.<sup>10,11</sup> The goal of the current study is to assess the relative benefits of surgical and nonsurgical treatment over a 10-year follow-up period using a broad range of validated patient-reported outcome measures.

## Materials and Methods

Details about the Maine Lumbar Spine Study design and methods, and 1- and 4-year outcomes for patients with lumbar spinal stenosis have been previously published.<sup>9–11</sup> The study prospectively followed patients treated by orthopedic surgeons and neurosurgeons in community-based practices throughout the state of Maine. Treatment, either surgical or nonsurgical, was determined in a routine clinical manner by the patient and the physician. For surgically treated patients, almost all had decompression laminectomy and had findings consistent with a diagnosis of spinal stenosis.<sup>10</sup> Fusion was uncommon, and internal fixation devices were not used. For nonsurgically treated patients, back exercises, bedrest, physical therapy, spinal manipulation, narcotic analgesics, and epidural steroids were most frequently used.<sup>10</sup>

**Study Population.** The diagnosis of lumbar spinal stenosis was based on physician assessment of appropriate symptoms, examination, and radiographic findings. Patients with spinal stenosis on advanced imaging studies could also have a herniated lumbar disc. To restrict the study to patients for whom surgery would be elective and acceptably safe, patients were excluded if they had prior lumbar spine surgery, cauda equina syndrome, developmental spine deformities, vertebral fractures, spine infection or tumor, inflammatory spondylopathy, pregnancy, or severe comorbid conditions. A total of 148 patients were enrolled from 1990 to 1992, with enrollment stratified to obtain roughly equal numbers of surgical and nonsurgical patients. Patients who initially chose nonsurgical treatment but underwent surgery before the first follow-up evaluation at 3 months were included in primary analyses as having been surgically treated ( $n = 10$ , 13%).

**Study Protocol and Patient Follow-up.** For eligible patients, written informed consent was obtained at study entry. Physicians completed a detailed baseline questionnaire including history, physical and neurologic findings, diagnostic proce-

**Table 1. Patient Characteristics and Features of Back Disorder at Baseline Evaluation**

Characteristic*	Surgical (n = 56)	Nonsurgical (n = 41)	P†
Age (yr) [mean (SD)]	65.7 (10.7)	65.5 (12.4)	0.94
Gender, male	37.5	43.9	0.53
Education, college graduate	25	20	0.57
Comorbid illnesses, yes‡	57.1	65	0.44
Work status, % retired	50	40	0.33
Receiving or applying for Workers' Compensation	12.5	20	0.32
Past episodes of back pain, none	19.6	10	0.20
Abnormal exam findings, mean (SD)§	0.86 (0.77)	1.0 (0.95)	0.43
Radiographic image reviewed (n)¶	35	15	
Moderate or severe findings#	80	80	1.0
Length of current episode, > 6 mo	65.5	56.4	0.37
Unilateral leg pain, yes	51.8	51.3	0.96
SF-36 score (0–100), mean (SD)**			
Physical function	30.9 (25.4)	47.8 (28.4)	0.005
Bodily pain	17.4 (16.9)	40.3 (21.8)	<0.001
Role-emotional	50 (46.3)	53.5 (41.4)	0.71
Mental health	66.7 (21.5)	66.8 (20.8)	0.99
General health	75.3 (21)	66.7 (21.2)	0.05

\*The data are expressed as the percentage, except as noted. Denominators differ slightly among variables because not all patients answered each question on the survey.

†P values compare surgical and nonsurgical treatment groups using Fisher's exact test or t test.

‡Any self-reported chronic pulmonary disease, heart disease, stroke, cancer, or diabetes.

§The mean no. of positive physical examination findings for a patient including unilateral strength, sensation or reflex abnormality (range, 0–3 findings).

¶Any computerized tomography, magnetic resonance imaging, or myelogram available for independent review.

#Global rating from normal to severe by study neuroradiologist blinded to treatment group and clinical information.

\*\*Higher scores indicate better function.

dures, and planned treatment. Baseline imaging studies (CT, MRI, or myelogram) were ordered as directed by the treating physicians, and 54% were available for independent review.<sup>10</sup> Baseline interviews were conducted in person with mailed follow-up questionnaires at 3, 6, and 12 months, and then yearly through 10 years. Forty-three patients died during follow-up, and they accounted for 43 of 62 patients (69%) not completing the 10-year survey. To maximize the number of patients included in long-term follow-up analyses, patients with 8- or 9-year follow-up, were included with those completing 10-year follow-up. All study activities were approved by the Institutional Review Boards at the University of Washington, Seattle, ME Medical Center, Portland, ME, and Massachusetts General Hospital, Boston, MA.

**Baseline Variables and Findings.** Baseline findings, including demographic information, employment and disability status, comorbid conditions, past spine history, physical examination and imaging findings, symptoms, and functional status, have been reported previously.<sup>10,11</sup> Patient sociodemographic characteristics were similar among those initially treated surgically or nonsurgically (Table 1). Most patients had symptoms for more than 6 months. Half had unilateral leg symptoms; abnormal physical examination findings were uncommon; and most had moderate or severe findings, all similar between treatment groups. However, there were important differences be-

**Table 2. Change in Symptoms and Functional Status from Baseline to 8- to 10-Yr Follow-up According to Initial Treatment**

Variable, mean score (SD) †	Surgical Treatment*			Nonsurgical Treatment*			P <sup>¶</sup>
	Baseline‡	8- to 10-Yr‡	Change§	Baseline‡	8- to 10-Yr‡	Change§	
Low back pain in the past week							
Frequency score	4.5 (1.9)	2.5 (2.1)	-1.9 (2.6)	3.1 (2.1)	2.3 (1.9)	-0.7 (2.6)	0.92
Bothersome score	4.4 (2.0)	2.5 (2.1)	-1.9 (2.4)	2.9 (2.0)	2.4 (2.1)	-0.6 (2.5)	0.53
Leg pain in the past week							
Frequency score	4.8 (1.6)	1.6 (2.0)	-3.3 (2.3)	3.1 (1.9)	2.4 (2.1)	-0.7 (2.6)	0.06
Bothersome score	5.1 (1.7)	1.6 (2.0)	-3.5 (2.6)	3.0 (1.9)	2.7 (2.3)	-0.3 (2.8)	0.02
Symptom Index							
Frequency	15.7 (5.7)	6.7 (6.8)	-9.7 (6.7)	9.6 (6.3)	8.4 (6.8)	-1.6 (9.0)	0.053
Bothersome	15.8 (6.4)	6.7 (6.9)	-9.6 (7.7)	8.9 (6.1)	8.8 (7.9)	-0.2 (9.3)	0.02
Modified Roland scale	15.9 (4.0)	8.7 (6.9)	-7.3 (6.5)	10.8 (6.5)	9.7 (7.6)	-1.2 (7.8)	0.02
SF-36 general health perceptions	2.3 (1.1)	3.1 (1.0)	0.8 (1.2)	2.7 (0.9)	3.4 (0.9)	0.7 (1.0)	0.25

\*The no. of patients reporting each variable ranged from 47 to 56 for surgical patients and 34 to 40 for nonsurgical patients.

†Low back and leg pain scores range from 0 to 6. Symptom frequency and bothersome indices are the sum of 4 questions (each response scored from 0 to 6): leg pain, leg or foot weakness, leg numbness, and leg pain after walking. Symptom Index scores range from 0 to 24, modified Roland Scale scores range from 0 to 23, and SF-36 general health perceptions range from 1 to 5.

‡Higher mean scores at baseline and 8 to 10 years indicate worse symptoms or function for all variables.

§Change is calculated as score at 8 to 10 years minus score at baseline. Negative values indicate improvement for all variables.

¶P values compare the change between surgical and nonsurgical treatment groups using multiple linear regression models that control for baseline score.

tween those surgically and nonsurgically treated. Surgically treated patients had more severe low back and leg pain, worse back-related function, and worse generic functional status (Tables 1 and 2). Despite worse back-related symptoms and function, surgically treated patients had better SF-36 general health perceptions than nonsurgically treated patients.

Patients not completing 8- to 10-year follow-up ( $n = 51$ ) had similar baseline findings compared with those completing follow-up. Compared with those completing long-term follow-up, nonresponders were more likely to smoke (27% vs. 13%,  $P = 0.02$ ), had older mean age (68.6 vs. 65.6 years,  $P = 0.26$ ), and had more comorbid conditions (71% vs. 60%,  $P = 0.22$ ). Forty-seven of 51 nonresponders (92%) returned at least 1 earlier follow-up survey (most recent follow-up: range, 6–84 months; median, 48 months). Outcomes at the most recent follow-up reported by these patients were similar to those completing 8- to 10-year follow-up.

**Outcome Measures.** On each follow-up questionnaire, patients were asked to describe their low back and leg pain compared with baseline with responses ranging from “much worse” to “completely gone.” The primary symptom outcome was improvement in the patient’s predominant symptom, either back or leg pain, as indicated at baseline. The outcome was considered improved if the response was “better,” “much better,” or “completely gone,” the same if the response was “about the same” or “a little better,” and worse if the response was “a little worse” or “much worse.” To assess for a more definitive positive treatment outcome, symptom responses were also classified as “much better” or “completely gone” compared with other responses. Patients were considered to be satisfied with their current state if they replied that they were “delighted,” “pleased,” or “mostly satisfied” on a 7-point scale. Satisfaction with their treatment decision was assessed by determining whether they would still choose their initial treatment at follow-up visits.

The frequency (from 0 “not at all” to 6 “always”) and bothersomeness (from 0 “not bothersome” to 6 “extremely bothersome”) of low back pain, leg pain, leg or foot weakness, leg numbness, and leg pain after walking in the past week were

assessed at baseline and follow-up. Symptom frequency and bothersome indexes, each with scores ranging from 0 to 24, were created by summing the four leg-related questions.<sup>12</sup> Back-specific functional status was measured using the modified Roland disability scale.<sup>12</sup> General health perceptions were assessed with a question from the Medical Outcomes Study Short Form 36-item questionnaire, “In general, would you say your health is” “excellent” (1) to “poor” (5). For each variable, higher scores indicate more severe symptoms or function. Finally, first operations for patients initially treated nonsurgically and reoperations for those treated surgically were assessed for all patients from physician office records, state hospital discharge data, and patient responses to follow-up surveys and telephone contact. Clinical details of subsequent surgical procedures were not available.

**Analysis.** Patient-reported outcomes were assessed from the most recent long-term follow-up questionnaire completed between 8 and 10 years. The rating of findings at follow-up used categorical responses, so distributions were directly compared between treatment groups using  $\chi^2$  tests or Fisher’s exact tests. Although baseline clinical features of the treatment groups differed, there was considerable overlap. To adjust for baseline differences between the two treatment groups, logistic regression models were used to estimate the marginal effect of surgical compared with nonsurgical treatment for the predominant pain symptom and satisfaction with the current state at long-term follow-up. Change in symptoms and functional status was assessed by subtracting long-term follow-up results from those at baseline, and linear regression models were used to examine the effect of treatment group after controlling for baseline score. In all analyses, the effect of the patient’s initial treatment decision, either surgical or nonsurgical care, was assessed. Reoperation rates among patients initially undergoing surgery and surgical crossover rates among those initially receiving nonsurgical treatment were assessed using survival analysis methods (in this case, time-to-an-event analysis) over 10 years. Secondary analyses of treatment effect considered the actual treatment received at most recent 8- to 10-year follow-up.

To examine the pattern of change over time, repeated-

**Table 3. Patient Reported Improvement in Symptom and Satisfaction Outcomes at 8- to 10-Year Follow-up**

8- to 10-Yr Outcomes (%)	Surgical	Nonsurgical	<i>P</i> *
Low back pain vs. baseline <sup>†</sup>	(n = 53)	(n = 38)	0.82
Improved	52.8	50.0	
Same	20.8	26.3	
Worse	26.4	23.7	
Leg pain compared to baseline <sup>†</sup>	(n = 48)	(n = 37)	0.04
Improved	66.7	40.5	
Same	14.6	18.9	
Worse	18.8	40.5	
Predominant symptom compared to baseline <sup>†‡</sup>	(n = 48)	(n = 36)	0.29
Improved	54.2	41.7	
Completely gone	27.1	13.9	
Much better	14.6	13.9	
Better	12.5	13.9	
Same	25.0	22.2	
Worse	20.8	36.1	
Satisfied with current state, yes <sup>§</sup>	55.4	48.8	0.52
Still choose the same initial treatment, yes <sup>¶</sup>	82.1	63.9	0.05

\**P* values assessed using Fisher's exact or  $\chi^2$  tests for categorical variables.  
<sup>†</sup>Symptom severity was reported to be improved if the response was "better" to "completely gone," the same if the response was "about the same" or "a little better," and worse if the response was "a little worse" or "much worse."

<sup>‡</sup>The predominant symptom, either back or leg pain, as rated by the patient at baseline.

<sup>§</sup>*N* = 56 for surgical and *N* = 41 for nonsurgical cohorts.

<sup>¶</sup>*N* = 56 for surgical and *N* = 36 for nonsurgical cohorts.

measures analysis was performed using data from all follow-ups between 2 and 10 years (*n* = 136, 92%). Mixed-effects models and logistic regression models with Generalized Estimating Equations were used to model the correlation structure of the repeated measures within each patient.<sup>13</sup> The treatment-by-time interaction was used to test how treatment effect differed over time in these models. All analyses were performed using a commercial software package (Statistical Analysis System, SAS Institute, Cary, NC).

## ■ Results

Forty-three of 148 eligible patients enrolled in the study died during follow-up, a 67.7% 10-year survival rate (using survival analysis). Among surviving patients, long-term follow-up between 8 and 10 years was available for 97 of 123 (79%) patients, 56 of 63 (89%) initially treated surgically, and 41 of 60 (68%) initially treated nonsurgically. For 86 (89%) patients, the most recent follow-up questionnaire was at 10 years.

### **Patients' Global Evaluation After 8 to 10 Years**

Patients assessed their current low back and leg pain symptoms at long-term follow-up in relation to their symptoms at baseline. A similar percentage of patients treated surgically reported improved low back pain compared with patients treated nonsurgically (53% vs. 50%, *P* = 0.82), but patients treated surgically reported greater improvement in leg pain (67% vs. 41%, *P* = 0.04) (Table 3). The predominant pain symptom, either low back or leg pain based on the patient's report at entry, was improved in 54% of surgical patients compared with 42% of nonsurgical patients (*P* = 0.29). If the

predominant symptoms was reported to be "much better" or "completely gone," 42% of surgical patients and 28% of nonsurgical patients were improved (*P* = 0.24). Long-term patient satisfaction with the current state was similar regardless of initial treatment. Fifty-five percent of surgically treated patients and 49% of nonsurgically treated patients were satisfied (*P* = 0.52). However, 82% of patients treated surgically would still definitely or probably chose their initial treatment compared with 64% of those treated nonsurgically (*P* = 0.05).

### **Change in Symptoms and Functional Status Over 8 to 10 Years**

The frequency and bothersomeness of low back pain and leg symptoms in the past week were rated by patients at baseline and long-term follow-up (Table 2). There was no statistically significant difference when comparing the change in low back pain symptoms between surgically and nonsurgically treated patients. In contrast, surgically treated patients reported worse leg pain symptoms and symptom indexes at baseline and had better outcomes after 8 to 10 years than nonsurgically treated patients. The change in the bothersomeness of leg pain and the symptom index was statistically significantly better for surgical patients (both *P* < 0.05). Back-specific functional status was assessed with the Roland disability questionnaire. Surgically treated patients reported significantly greater improvement in the modified Roland scale than nonsurgically treated patients (*P* = 0.03). On the other hand, general health perceptions worsened to a similar degree in both treatment groups over time.

### **Long-term Independent Predictors of Symptom Improvement and Satisfaction**

Since initial treatment was not randomly assigned, logistic regression modeling was performed to adjust for differences in baseline features that may be associated with outcomes among patients initially treated surgically or nonsurgically. Outcomes assessed included satisfaction with the current state and improvement in the patients' predominant symptom at 8 to 10 years. Variables in all models included treatment group, age, and sex as well as other baseline variables with *P* values < 0.2 in adjusted models. For satisfaction with the current state, the effect of treatment received was similar in unadjusted (odds ratio [OR], 1.3; 95% confidence interval [CI], 0.6–2.9) and adjusted (OR, 1.4; 95% CI, 0.5–3.9) models (Table 4). For improvement in the patients' predominant symptom ("improved" defined as a patient response of "better," "much better," or "completely gone"), surgical treatment was not a statistically significant predictor, but the OR was more in favor of surgical treatment in adjusted (OR, 2.6; 95% CI, 0.9–8.5) than in unadjusted (OR, 1.7; 95% CI, 0.7–4.0) models. If symptom improvement was restricted to responses of "much better" or "completely gone," the OR favoring surgery was significantly greater in adjusted (OR, 3.6; 95% CI, 1.1–13.9) compared with unadjusted (OR, 1.9; 95% CI, 0.7–4.7) models.

**Table 4. Baseline Predictors of Patient Satisfaction at 8- to 10-Yr Follow-up**

Baseline Features	Satisfied With Current State*	
	OR (95% CI)	P
Initial treatment, surgical		
Unadjusted	1.3 (0.6–2.9)	0.52
Adjusted†	1.4 (0.5–3.9)	0.52
Education, college graduate	2.3 (0.7–7.8)	0.18
Cigarette use, current or quit in past 6 mo	0.2 (0.03–1.0)	0.06
SF-36 Social Function (0–100), per 10-point increment	1.2 (1.0–1.4)	0.06
SF-36 General Health (0–100), per 10-point increment	1.2 (1.0–1.6)	0.08

\* Patients were considered satisfied with their current state if they replied that they were “delighted,” “pleased,” or “mostly satisfied” on a 7-point scale. †Stepwise multiple logistic regression modeling was used to adjust for differences in baseline features for patients treated surgically or nonsurgically. The table presents the final model adjusted for initial treatment, age, gender, and other baseline variables with *P* values of < 0.2 (c-statistic = 0.74). Age and gender are not shown because these variables were not associated with outcome.

The small number of individuals with long-term follow-up limits the study’s ability to identify independent predictors of outcome. No statistically significant independent predictors of satisfaction at 8- to 10-year follow-up assessment were identified (Table 4). However, better baseline SF-36 social function and general health status and higher educational level increased the odds of satisfaction, whereas cigarette smoking decreased the odds of a patient reporting long-term satisfaction. In models examining symptom improvement, SF-36 social function, educational level, and smoking status were also potential independent predictors of outcome (data not shown).

**Time Course of Outcomes**

Mean scores at baseline and at each follow-up over 10 years for the frequency of symptoms and back-specific functional status are shown in Figure 1. The percentage of patients reporting satisfaction with their current state at each follow-up is shown in Figure 2. As previously reported,<sup>11</sup> most of the improvement in outcome due to surgery was seen shortly after patients’ entry into the study followed by a narrowing of the relative benefit of surgical treatment. From years 2 through 10, there was slight worsening of the frequency of symptoms and Roland functional status over time in both treatment groups and to a similar extent. There was a slight increase in satisfaction with the current state for nonsurgically treated patients and a slight decrease for surgically treated patients between 2 and 10 years; however, this narrowing was not statistically significant (*P* = 0.36 for the interaction between time and treatment group).

**Lumbar Spine Surgery After Initial Treatment**

Among patients initially undergoing surgical treatment, the 10-year reoperation rate was 23% (n = 15; median time to reoperation, 60 months). Among patients initially receiving nonsurgical treatment, the crossover rate

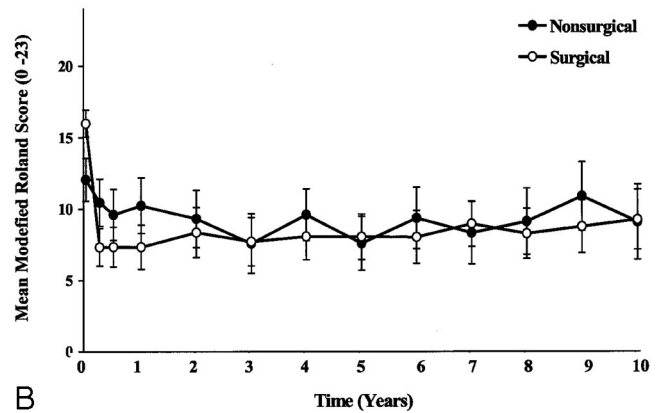
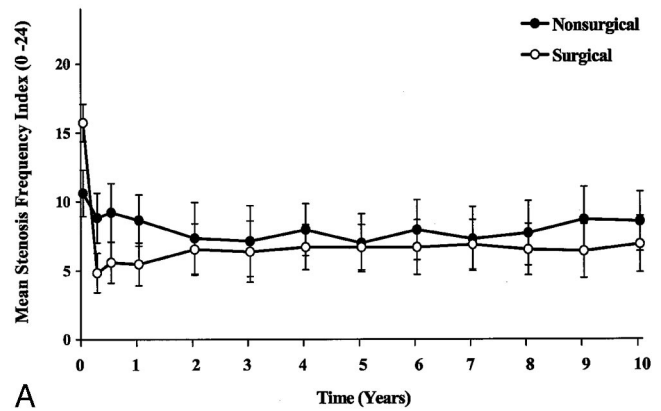


Figure 1. Time course of symptom and functional status outcomes. Assessed at initial evaluation and at 3, 6, and 12 months, and then yearly follow-up for nonsurgical (●) and surgical (○) treatment. Mean ± 2 SE. **A**, Symptom Frequency Index. **B**, Modified Roland score.

to surgery between 3 months and 10 years was 39% (n = 22; median time to crossover, 24 months). Although there were no significant differences in baseline characteristics and findings, surgical patients undergoing a reoperation tended to be younger, male, have fewer comorbid conditions, and have worse back pain than those not undergoing a reoperation. Nonsurgical patients having subsequent surgery had slightly worse mental and

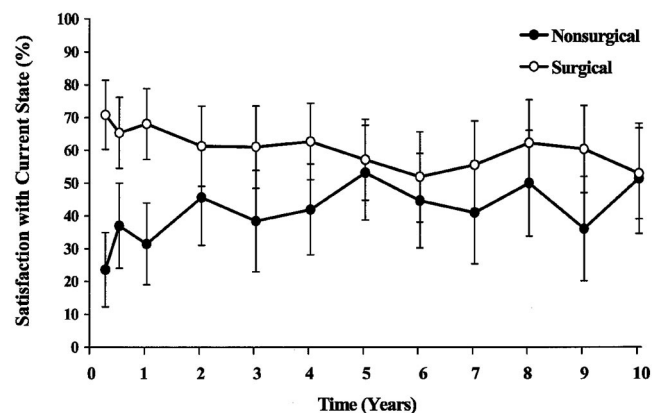


Figure 2. Time course of satisfaction outcome. Assessed at 3, 6, and 12 months, and then yearly follow-up for nonsurgical (●) and surgical (○) treatment. Percent satisfied with current state ± 2 SE.

**Table 5. Eight- to Ten-Yr Outcomes for Patients Based Upon Initial and Final Treatment Status**

	Surgical Treatment				Nonsurgical Treatment		
	No Reoperation	Reoperation		Nonsurgical Crossover			
8- to 10-yr outcomes	(n = 42)	(n = 14)	<i>P</i> *	(n = 15)	<i>P</i> *	(n = 26)	<i>P</i> *
Predominant symptom, % improved <sup>†</sup>	60.0	38.5	0.19	35.7	0.13	45.5	0.29
Satisfied with current state, % yes	59.5	42.9	0.28	40.0	0.20	53.9	0.65
Stenosis Frequency Index, mean change (SD) <sup>‡</sup>	-10.3 (6.3)	-8.3 (7.6)	0.47	3.8 (8.6)	<0.001	-5.1 (7.6)	1.0
Roland scale, mean change (SD) <sup>‡</sup>	-8.3 (6.3)	-4.1 (6.0)	0.08	2.6 (5.5)	<0.001	-3.8 (8.1)	0.60
SF-36 general health perceptions, mean change (SD) <sup>‡</sup>	0.7 (1.1)	1.0 (1.4)	1.0	1.2 (0.8)	0.02	0.4 (1.0)	1.0

\* From multivariate regression models, using surgical with no reoperation group as reference.

<sup>†</sup>The predominant symptom, either back or leg pain, as rated by the patient at baseline. Symptom severity was reported to be improved if the response was "better" to "completely gone."

<sup>‡</sup>Change is calculated as score at 8 to 10 years minus score at baseline. Negative values indicate improvement for all variables. *P* values obtained using analysis of variance with Dunnett's adjustment for multiple comparisons and controlling for baseline score.

general health status and more abnormal examination findings than those not having subsequent surgery, but no difference was statistically significant.

### 10-Year Outcomes According to Actual Treatment Received

For those completing 8- to 10-year follow-up, outcomes were also assessed based on the actual treatment received for the 26 patients treated nonsurgically and the 71 patients treated surgically (including 15 patients initially in the nonsurgical group who crossed over to surgical treatment). There were no clinically important or statistically significant differences according to actual treatment received at 8- to 10-year follow-up for any variable (data not shown).

Eight- to 10-year outcomes were also compared according to the patients' initial and final treatment status: surgery with or without a second spine operation (reoperation) and initial nonsurgical treatment with or without crossover to surgical treatment after 3 months (Table 5). There were no significant differences between surgical and nonsurgical patients who remained in their initial treatment groups throughout the study, although surgically treated patients had slightly better outcomes than nonsurgically treated patients. Surgical patients having a reoperation and nonsurgical patients having subsequent surgery had outcomes that were inferior to those initially receiving surgical treatment. For example, 43% and 40% of those having a reoperation or crossing over to subsequent surgery were satisfied with the current state compared with 60% and 54% of surgical or nonsurgical treatment continuing without subsequent surgery.

### Discussion

In this study, evaluating long-term outcomes of lumbar spinal stenosis was made more difficult by loss to follow-up due to mortality. Among patients completing 8- to 10-year follow-up, low back pain and satisfaction were similar among patients initially treated surgically or nonsurgically. However, surgically treated patients reported greater improvement in leg symptoms and back-related functional status than nonsurgically treated pa-

tients. There was a high rate of additional surgical procedures after initial surgical and nonsurgical treatment over 10 years. Long-term outcomes according to actual treatment received did not differ for those treated surgically or nonsurgically, mainly because of poorer outcomes associated with individuals undergoing additional surgical procedures than among those continuing with their initial treatment.

Despite being an increasingly common diagnosis in aging individuals, the relative risks and benefits of various treatment options for lumbar spinal stenosis are poorly understood. Although rates of surgery for patients with lumbar spinal stenosis have risen dramatically in the United States,<sup>2,3</sup> there are no large randomized studies demonstrating the relative benefit of surgical compared with nonsurgical treatment.<sup>5</sup> Short-term results from a randomized trial presented as an abstract reported improved outcomes associated with surgical treatment.<sup>8</sup> Among published studies, a single small trial randomly assigned 31 of 100 patients with lumbar spinal stenosis to surgical (n = 13) or nonsurgical (n = 18) treatment.<sup>14</sup> Of the 18 individuals randomly assigned to nonsurgical treatment, 10 had subsequent surgery (crossed over) after a median of 3.5 months. Patient outcomes were recorded and judged by the principal investigator (level of agreement only fair, kappa 0.45 after 10 years). Short-term outcomes were reported to favor surgical treatment. Among 28 of 31 patients completing 10-year follow-up, "good" treatment results were reported for 10 of 11 (91%) of surgical patients, 8 of 8 (100%) of nonsurgical patients, and 4 of 9 (44%) patients crossing over to surgery. However, little or no pain was reported by only 5 of 11 (45%) patients treated surgically, 2 of 8 (25%) patients treated nonsurgically, and 2 of 6 (33%) patients crossing over to surgery.

In contrast to similar outcomes among patients with an intervertebral disc herniation in the Maine Lumbar Spine Study and Weber's randomized trial,<sup>15-18</sup> outcomes of patients with lumbar spinal stenosis in the Maine Lumbar Spine Study and Amundsen's study are considerably different, especially in terms of change over

time.<sup>6,10,11</sup> In the Maine study, improvement in the patient's predominant symptom (patient self-reported leg or back pain "completely gone," "much better," or "better") according to initial surgical or nonsurgical treatment was reported by 77% and 44% at 1 year, 70% and 52% at 4 years, and 54% and 42% at 8 to 10 years, respectively.<sup>10,11</sup> In the Amundsen randomized study, the overall outcome was judged "good" according to initial surgical or nonsurgical treatment in 69% and 33% at 1 year, 92% and 47% at 4 years, and 91% and 71% at 10 years, respectively.<sup>6</sup> For observational patients followed by Amundsen, the overall outcome was judged "good" according to initial surgical or nonsurgical treatment in 89% and 64% at 1 year, 84% and 57% at 4 years, and 71% and 73% at 10 years.<sup>6</sup> In the current study, surgical outcomes declined over time while nonsurgical outcomes remained fairly stable. In contrast, except for the surgical patients in the observational cohort, Amundsen's outcomes improved over time regardless of treatment. A number of other observational studies show stable or declining long-term outcomes.<sup>19-22</sup>

Regardless of temporal trends in outcome, there is considerable variation in the percentage of individuals reporting long-term benefit from treatment.<sup>6,21-26</sup> A variety of reasons may account for differences in reported outcomes, such as differences in patient selection, length of follow-up, or the outcome measures used to define improvement. For example, subjects in Amundsen's study were younger, more likely to be male, and more likely to be receiving disability than patients in this study.<sup>10,14</sup> The nature of treatments received also varied. For surgically treated patients, decompressive laminectomy varies in its location and extent, and there is increasing use of fusion, especially with instrumentation. For nonsurgically treated patients, a wide variety of measures are used, including invasive procedures such as epidural and facet blocks. Thus, despite evidence from this and other studies demonstrating at least short-term benefit of surgical treatment, larger randomized trials such as the Spine Patient Outcomes Research Trials are clearly needed to better define the relative benefit of alternative treatments and which patients derive any benefit.<sup>1,5,27,28</sup>

Evaluating long-term outcomes of patients in this study was made more difficult by a high mortality rate. This is expected for a disorder that is most common in the elderly. The survival rate was similar in patients initially undergoing surgical and nonsurgical treatment. Although elderly patients may have comorbid conditions that not only increase the risk of surgery but also may shorten any benefit of the procedure, symptomatic patients with impaired function should not necessarily be excluded from these interventions. However, future studies need to use analytic techniques that can account for the impact of other morbidity and mortality over time.<sup>29</sup>

Patients in this study initially treated nonsurgically had a high rate of subsequent surgery, as noted in other studies.<sup>6</sup> After 5 years, 25% of nonsurgical patients had

undergone surgery, and this increased to 39% after 10 years. Among patients initially treated surgically, there was a lower reoperation rate for spinal stenosis patients than for patients in the Maine Lumbar Spine Study with an intervertebral disc herniation over 5 years (1.3% *vs.* 9% at 1 year and 11% *vs.* 20.5% at 5 years, respectively, for spinal stenosis and disc herniation patients). However, by 10 years, the reoperation rate for spinal stenosis patients was 23%, comparable to disc herniation patients. This long-term reoperation rate is higher than most,<sup>6,19,24</sup> but not all, prior reported studies.<sup>20</sup> Regardless of initial treatment, outcomes of those undergoing subsequent surgery were inferior to those who continued with their initial treatment. Because of this, excluding patients who underwent subsequent surgery ("as treated" analyses) did not change the relative outcomes of intention-to-treat (*i.e.*, comparing all initially nonsurgically treated to all initially surgically treated) analyses.

Strengths of this study include prospective, long-term follow-up in a large percentage of surviving patients from contemporary comparison groups treated in community-based clinical practice employing a range of validated outcome measures using prospective and retrospective perspectives.<sup>30</sup> The study's major limitation is its observational, nonrandomized design.<sup>10,11,17</sup> Important differences at baseline among patients treated surgically or nonsurgically means that one cannot be certain that differences in outcomes among treatment groups are exclusively due to surgery rather than unmeasured confounders. However, the effect of surgical treatment on the predominant symptom and satisfaction was similar after adjusting for potential predictors of outcome in regression models. Other limitations include the relatively small sample size and high loss to follow-up because of death from other illnesses. However, even with a larger sample size, the small differences in low back pain and satisfaction favoring surgical treatment would not be clinically important. The lack of clinical details for patients having a subsequent surgical procedure does not permit knowing why the procedure was performed or what it involved. Finally, since few patients underwent fusion in this study and none received instrumented fusion, it is not possible to compare fusion outcomes to those undergoing decompressive laminectomy alone or to conservative care.

## ■ Conclusion

For patients with lumbar spinal stenosis, the relative benefits of initial surgical compared with nonsurgical treatment in terms of low back pain relief, predominant symptom improvement, and satisfaction with the current state seen at 1- and 4-year follow-up<sup>10,11</sup> were no longer present after 8 to 10 years. Leg pain relief and greater back-related functional status continued to favor those initially receiving surgical treatment. Because a large number of patients initially receiving nonsurgical treatment subsequently had surgery over long-term follow-up, we also examined results according to actual treat-

ment received. Outcomes after 8 to 10 years based on the actual treatment showed no differences. Finally, regardless of initial or subsequent treatment received, favorable long-term outcomes were only reported by about half the patients. Thus, surgery may offer improved outcomes compared with nonsurgical treatment over several years. However, patients reluctant to undergo surgery may elect conservative care, knowing that their symptoms will likely remain stable and long-term outcomes are similar. Until results from larger, controlled studies are available, patients should be offered treatment options that are based on their symptoms, functional impairment, other limiting medical conditions, and personal preferences.<sup>31,32</sup>

### Acknowledgments

The authors thank YuChiao Chang, PhD, for assistance with statistical analyses and Valerie Soucie for assistance with managing long-term patient follow-up.

### Key Points

- Outcomes of 97 patients with lumbar spinal stenosis treated surgically or nonsurgically were evaluated over 8 to 10 years.
- Because treatment was determined in a routine clinical manner by the patient and the physician, those surgically treated had more severe symptoms and worse functional status at baseline.
- Low back pain relief, predominant symptom improvement, and satisfaction were similar among patients initially undergoing surgical or nonsurgical treatment.
- However, leg pain relief and change in back-specific functional status favored patients initially treated surgically.

### References

1. Birkmeyer NJ, Weinstein JN. Medical versus surgical treatment for low back pain: evidence and clinical practice. *Eff Clin Pract* 1999;2:218–27.
2. Ciol MA, Deyo RA, Howell E, et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc* 1996;44:285–90.
3. Taylor VM, Deyo RA, Cherkin DC, et al. Low back pain hospitalization: recent United States trends and regional variations. *Spine* 1994;19:1207–13.
4. Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine* 2003;28:616–20.
5. Gibson JN, Waddell G, Grant IC. Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev* 2000;3:CD001352.
6. Amundsen T, Weber H, Nordal HJ, et al. Lumbar spinal stenosis: conservative or surgical management? A prospective 10-year study. *Spine* 2000;25:1424–36.
7. Birkmeyer NJ, Weinstein JN, Tosteson AN, et al. Design of the Spine Patient outcomes Research Trial (SPORT). *Spine* 2002;27:1361–72.
8. Malmivaara A, Slati P, Heliovaara P, et al. Operative treatment for moderately severe lumbar spinal stenosis: a randomized controlled trial. *Presented at the Annual Meeting of the International Society for the Study of the Lumbar Spine*. Vancouver, British Columbia, Canada, May 2003.
9. Keller RB, Atlas SJ, Singer DE, et al. The Maine Lumbar Spine Study: I. Background and Concepts. *Spine* 1996;21:1769–76.
10. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study: III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. *Spine* 1996;21:1787–95.
11. Atlas SJ, Keller RB, Robson D, et al. Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the Maine Lumbar Spine Study. *Spine* 2000;25:556–62.
12. Patrick DL, Deyo RA, Atlas SJ, et al. Assessing health-related quality of life in patients with sciatica. *Spine* 1995;20:1899–909.
13. Diggle PJ, Liang K, Zeger SL. *Analysis of Longitudinal Data*. New York: Oxford University Press, 1994.
14. Amundsen T, Weber H, Lilleas F, et al. Lumbar spinal stenosis: clinical and radiologic features. *Spine* 1995;20:1178–86.
15. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study: II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine* 1996;21:1777–86.
16. Atlas SJ, Keller RB, Chang Y, et al. Surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: five-year outcomes from the Maine Lumbar Spine Study. *Spine* 2001;26:1179–87.
17. Atlas SJ, Singer DE, Wu YA, et al. Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the Maine Lumbar Spine Study. *Spine* 2005;30:926–934.
18. Weber H. Lumbar disc herniation: a controlled, prospective study with ten years of observation. *Spine* 1983;8:131–40.
19. Herno A, Airaksinen O, Saari T. Long-term results of surgical treatment of lumbar spinal stenosis. *Spine* 1993;18:1471–4.
20. Katz JN, Lipson SJ, Chang LC, et al. Seven to 10-year outcome of decompressive surgery for degenerative lumbar spine disorders. *Spine* 1996;21:92–8.
21. Rompe JD, Eysel P, Zollner J, et al. Degenerative lumbar spinal stenosis: long-term results after undercutting decompression compared with decompressive laminectomy alone or with instrumented fusion. *Neurosurg Rev* 1999;22:102–6.
22. Scholz M, Firsching R, Lanksch WR. Long-term follow up in lumbar spinal stenosis. *Spinal Cord* 1998;36:200–4.
23. Cornefjord M, Byrod G, Brisby H, et al. A long-term (4- to 12-year) follow-up study of surgical treatment of lumbar spinal stenosis. *Eur Spine J* 2000;9:563–70.
24. Hee HT, Wong HK. The long-term results of surgical treatment for spinal stenosis in the elderly. *Singapore Med J* 2003;44:175–80.
25. Hurri H, Slati P, Soini J, et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. *J Spinal Disord* 1998;11:110–5.
26. Iguchi T, Kurihara A, Nakayama J, et al. Minimum 10-year outcome of decompressive laminectomy for degenerative lumbar spinal stenosis. *Spine* 2000;25:1754–9.
27. Benoit M. The natural history of lumbar degenerative spinal stenosis. *Joint Bone Spine* 2002;69:450–7.
28. Turner JA, Ersek M, Herron L, et al. Surgery for lumbar spinal stenosis: attempted meta-analysis of the literature. *Spine* 1992;17:1–8.
29. Tosteson TD, Buonaccorsi JP, Demidenko E. Covariate measurement error and the estimation of random effect parameters in a mixed model for longitudinal data. *Stat Med* 1998;17:1959–71.
30. Fischer D, Stewart AL, Bloch DA, et al. Capturing the patient's view of change as a clinical outcome measure. *JAMA* 1999;282:1157–62.
31. Deyo RA, Cherkin DC, Weinstein J, et al. Involving patients in clinical decisions: impact of an interactive video program on use of back surgery. *Med Care* 2000;38:959–69.
32. Goldberg HI, Deyo RA, Taylor VM, et al. Can evidence change the rate of back surgery? A randomized trial of community-based education. *Eff Clin Pract* 2001;4:95–104.