

What Are Long-term Predictors of Outcomes for Lumbar Disc Herniation? A Randomized and Observational Study

Dana Kerr MD, Wenyan Zhao PhD,
Jon D. Lurie MD, MS

Published online: 24 July 2014

© The Association of Bone and Joint Surgeons® 2014

Abstract

Background Although previous studies have illustrated improvements in surgical cohorts for patients with intervertebral disc herniation, there are limited data on predictors of long-term outcomes comparing surgical and nonsurgical outcomes. **Questions/purposes** We assessed outcomes of operative and nonoperative treatment for patients with intervertebral disc herniation and symptomatic radiculopathy at 8 years from the Spine Patient Outcomes Research Trial. We specifically examined subgroups to determine whether certain populations had a better long-term outcome with surgery or nonoperative treatment.

Funding: The National Institute of Arthritis and Musculoskeletal and Skin Diseases (U01-AR45444 and P60AR062799).

One of the authors (JDL) certifies that he or she, or a member of his or her immediate family, has received or may receive payments or benefits, during the study period, of USD 10,000 to USD 100,000 from Fziomed (San Luis Obispo, CA, USA).

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

D. Kerr, W. Zhao, J. D. Lurie (✉)
The Dartmouth Institute for Health Policy and Clinical Practice,
Lebanon, NH 03756, USA
e-mail: Jon.D.Lurie@dartmouth.edu

J. D. Lurie
Department of Medicine, One Medical Center Drive, Lebanon,
NH 03756, USA

J. D. Lurie
Department of Orthopaedics, One Medical Center Drive,
Lebanon, NH 03756, USA

Methods Patients with symptomatic lumbar radiculopathy for at least 6 weeks associated with nerve root irritation or neurologic deficit on examination and a confirmed disc herniation on cross-sectional imaging were enrolled at 13 different clinical sites. Patients consenting to participate in the randomized cohort were assigned to surgical or nonoperative treatment using variable permuted block randomization stratified by site. Those who declined randomization entered the observational cohort group based on treatment preference but were otherwise treated and followed identically to the randomized cohort. Of those in the randomized cohort, 309 of 501 (62%) provided 8-year data and in the observational group 469 of 743 (63%). Patients were treated with either surgical discectomy or usual nonoperative care. By 8 years, only 148 of 245 (60%) of those randomized to surgery had undergone surgery, whereas 122 of 256 (48%) of those randomized to nonoperative treatment had undergone surgery. The primary outcome measures were SF-36 bodily pain, SF-36 physical function, and Oswestry Disability Index collected at 6 weeks, 3 months, 6 months, 12 months, and then annually. Further analysis studied the following factors to determine if any were predictive of long-term outcomes: sex, herniation location, depression, smoking, work status, other joint problems, herniation level, herniation type, and duration of symptoms.

Results The intent-to-treat analysis of the randomized cohort at 8 years showed no difference between surgical and nonoperative treatment for the primary outcome measures. Secondary outcome measures of sciatica bothersomeness, leg pain, satisfaction with symptoms, and self-rated improvement showed greater improvement in the group randomized to surgery despite high levels of crossover. The as-treated analysis of the combined randomized and observational cohorts, adjusted for potential

confounders, showed advantages for surgery for all primary outcome measures; however, this has the potential for confounding from other unrecognized variables. Smokers and patients with depression or comorbid joint problems had worse functional outcomes overall (with surgery and nonoperative care) but similar surgical treatment effects. Patients with sequestered fragments, symptom duration greater than 6 months, those with higher levels of low back pain, or who were neither working nor disabled at baseline showed greater surgical treatment effects.

Conclusions The intent-to-treat analysis, which is complicated by high rates of crossover, showed no difference over 8 years for primary outcomes of overall pain, physical function, and back-related disability but did show small advantages for secondary outcomes of sciatica bothersomeness, satisfaction with symptoms, and self-rated improvement. Subgroup analyses identified those groups with sequestered fragments on MRI, higher levels of baseline back pain accompanying radiculopathy, a longer duration of symptoms, and those who were neither working nor disabled at baseline with a greater relative advantage from surgery at 8 years.

Level of Evidence Level II, therapeutic study. See Instructions for Authors for a complete description of levels of evidence.

Introduction

Lumbar disc herniation is a common low back disorder and the most common cause of sciatica in adults [12]. Lifetime incidence of sciatica is approximately 13% to 40% [13]. According to the National Hospital Discharge Survey in 2007 [16], there were 320,000 inpatient discharges for intervertebral disc disease. Although inpatient hospital discharge data are available, nationwide data on outpatient procedures are not widely available. From 1996 to 2000, there was an increasing trend of outpatient discectomies with a concomitant increase in both inpatient and outpatient lumbar spine procedures. This implies a substantially larger number than 320,000 discectomies per year for intervertebral disc herniation (IDH); however, exact figures are unavailable [15].

Lumbar discectomy improves short-term pain and function, but there are limited data on long-term outcomes. The Lumbar Disc Herniation study by Weber [3, 25] and the Maine Lumbar Spine Study (MLSS) [3, 25] are two of the largest published studies with long-term results. The Weber study [25] was a randomized, prospective study examining outcomes of patients with sciatica at a single institution with 10-year followup; the main outcome was the author's interpretation of patient statements of their current well-being. Weber's study [25] found greater improvement in the surgical cohort at

1 year; however, this difference diminished at 4 years and there was no difference between the groups at 10 years, although 26% of the nonoperative group crossed over to surgery. The MLSS was a multicenter, prospective, observational study comparing operative and nonoperative cohorts with outcomes including severity of patient symptoms, Roland-Morris disability scale, and the SF-36 [3]. In the MLSS, the operative cohort showed significantly greater improvement at 1 year, which maintained significance out to 10 years [3].

The Spine Patient Outcome Research Trial (SPORT) is a multicenter, prospective study with observational and randomized study arms comparing nonoperative and operative treatments for patients with IDH using validated outcome measures [20, 22, 26, 27]. In SPORT analyses out to 8 years, the intention-to-treat analysis, which was limited by high rates of crossover, found no statistically significant surgical treatment effects for the primary outcomes of general pain, physical function, or pain-related disability but did show statistically significant advantages for sciatica bothersomeness, satisfaction with symptoms, and self-rated improvement; the as-treated analysis combining the randomized and observational cohorts, which carefully controlled for potentially confounding baseline factors but could still be vulnerable to unmeasured confounding, showed significantly greater improvements with surgery for all outcomes except work status compared with patients treated nonoperatively [20]. We summarize those 8-year results and investigate the long-term outcomes in specific patient subgroups defined by sex, work status, smoking status, presence of a comorbid joint problem, history of depression, severity of low back pain, presence of a motor deficit, herniation level, herniation morphology, duration of symptoms at enrollment, and herniation location [4, 19, 21]. Our primary research question is whether there are specific characteristics at baseline that are able to identify subgroups of patients more likely to benefit from surgery compared with nonoperative treatment over the long term.

Patients and Methods

SPORT enrolled patients from March 2000 until November 2004 at 13 medical centers with multidisciplinary spine practice in 11 US states. The methods for this trial have been described previously [19–22, 27–29] and is briefly summarized subsequently. Each center obtained institutional review board approval.

Patient Population

Men and women who had symptomatic and imaging-confirmed lumbar radiculopathy were eligible for this study; patients were required to be surgical candidates with

persistent symptoms for at least 6 weeks. Patients consenting to participate in the randomized cohort were assigned by chance to surgical or nonoperative treatment using variable permuted block randomization stratified by site. Those who declined randomization entered the observational cohort group based on treatment preference but were otherwise treated and followed identically to the randomized cohort. Surgical and nonoperative treatments were compared using primary analyses using the change from baseline at each followup. The final analysis combined the two cohorts. Randomization was performed by computer-generated random treatment assignment based on permuted blocks (randomly generated blocks of six, eight, 10, and 12) within sites, which occurred immediately after enrollment. Neither subjects nor treatment providers were blinded as a result of the nature of the treatment.

Sample Size

An original sample size of 250 patients in each treatment group was originally determined to be sufficient (with a two-sided significance level of 0.05 of 85% power) to detect a 10-point difference in the SF-36 bodily pain and physical functions scales or a similar effect size in the Oswestry Disability Index (ODI).

In total, of the 1244 patients with IDH enrolled into the SPORT, 1195 (96%) had at least one followup and were included in the analysis (Table 1). Of these 1195 patients, 474 were in the randomized cohort and 721 were in the observational cohort. In the randomized cohort, 245 were assigned to surgical treatment and 256 to nonoperative treatment. Of those randomized to surgery, 57% had surgery by 1 year and 60% by 8 years. In the group randomized to nonoperative care, 41% of patients had surgery by 1 year and 48% by 8 years. In the observational cohort, 521 patients initially chose surgery and 222 patients initially chose nonoperative care. Of those initially choosing surgery, 95% received surgery by 1 year; at 8 years, 12 additional patients had undergone primary surgery. Of those choosing nonoperative treatment, 20% had surgery by 1 year and 25% by 8 years [20].

Overall, the randomized and observational patients were similar with the randomized patients showing somewhat lower mean disability (ODI 46.9 versus 51.1, $p < 0.001$), a lower likelihood of worsening symptoms at baseline (34% versus 43%, $p < 0.001$), and a lower preference for surgery (27% versus 66%, $p < 0.001$). In the combined randomized and observational cohorts, 803 of 1195 (67%) underwent surgery by 8 years and 392 of 1195 (33%) remained nonoperative. Comparing groups by treatment received, the nonoperative group was, on average, older (43.8 versus 40.7 years, $p < 0.001$), more likely to be

working (66% versus 58%, $p = 0.007$), less likely to be very dissatisfied with symptoms (65% versus 88%, $p < 0.001$), better physical function (SF-36 physical function score 48.4 versus 32.6, $p < 0.006$), and lower disability on ODI (38.6 versus 54.7, $p < 0.001$) at baseline.

Interventions

The surgical treatment was a standard open discectomy with examination of the involved nerve root by a qualified spine surgeon. The nonoperative group received usual care recommended to include at least active physical therapy, education/counseling with home exercise instruction, and, if tolerated, nonsteroidal antiinflammatory drugs with participating sites encouraged to use all appropriate nonoperative therapies [7]. Additional treatments were individualized to the patient by the treating provider(s) and tracked prospectively.

Study Measures

The primary research question was to determine the predictors of long-term outcomes of operative and nonoperative treatment in patients with lumbar disc herniation. The primary outcome measures were the ODI and SF-36 bodily pain and physical function scores. Secondary outcomes included the Sciatica Bothersomeness Index, leg and back pain bothersomeness, patient satisfaction with care and with symptoms, patient global perceived improvement, and work status. These outcomes were measured at 6 weeks, 3, 6, and 12 months, and annually thereafter.

Statistical Methods

Preliminary analyses compared means and proportions of baseline characteristics between the randomized and observational cohorts as described in previous SPORT papers [20, 26–28]. The primary analyses compared surgical and nonoperative treatments on an intent-to-treat basis by using the change from baseline at each followup. Baseline predictors of time until surgical treatment (including treatment crossover) in both cohorts were determined through a stepwise proportional hazards regression model with an inclusion criterion of p values of < 0.1 to enter and p values of > 0.05 to exit. Predictors of missing followup visits at yearly intervals up to 8 years were separately determined through stepwise logistic regression. Baseline characteristics that predicted surgery or a missed visit at any time point were then entered into

Table 1. Patient baseline demographic characteristics, comorbidities, and health status measures according to study cohort and treatment received

IDH	SPORT study cohorts			Combined cohorts		
	Randomized cohort (n = 474)	Observational cohort (n = 721)	p value	Treatment received		p value
				Surgery (n = 803)	Nonoperative (n = 392)	
Mean age (years; SD)	42.3 (11.6)	41.4 (11.2)	0.18	40.7 (10.8)	43.8 (12.3)	< 0.001
Female	194 (41%)	313 (43%)	0.43	346 (43%)	161 (41%)	0.55
Race, white	401 (85%)	635 (88%)	0.10	707 (88%)	329 (84%)	0.061
Work status			0.71			0.007
Full- or part-time	292 (62%)	433 (60%)		467 (58%)	258 (66%)	
Disabled	58 (12%)	100 (14%)		122 (15%)	36 (9%)	
Other	124 (26%)	187 (26%)		213 (27%)	98 (25%)	
Smoker	108 (23%)	175 (24%)	0.60	201 (25%)	82 (21%)	0.13
Comorbidities						
Depression	62 (13%)	79 (11%)	0.31	94 (12%)	47 (12%)	0.96
Joint Problem	98 (21%)	124 (17%)	0.15	130 (16%)	92 (23%)	0.003
Other	221 (47%)	305 (42%)	0.16	334 (42%)	192 (49%)	0.019
Bodily pain score	28.3 (19.9)	26.4 (20.3)	0.13	23.4 (18)	34.8 (22.1)	< 0.001
Physical functioning score	39.5 (25.3)	36.7 (25.7)	0.066	32.6 (23.5)	48.4 (26.4)	< 0.001
Mental Component Summary score	45.9 (12)	44.7 (11.2)	0.081	44.7 (11.4)	46.2 (11.8)	0.035
Oswestry Disability Index	46.9 (20.9)	51.1 (21.4)	< 0.001	54.7 (19.6)	38.6 (20.5)	< 0.001
Sciatica Frequency Index (0–24)	15.6 (5.5)	16.1 (5.3)	0.19	16.7 (5.1)	14.2 (5.6)	< 0.001
Sciatica Bothersome Index (0–24)	15.2 (5.2)	15.8 (5.3)	0.057	16.4 (4.9)	13.8 (5.6)	< 0.001
Symptom satisfaction, very dissatisfied	371 (78%)	585 (81%)	0.25	705 (88%)	251 (64%)	< 0.001
Problem getting better or worse			< 0.001			< 0.001
Getting better	90 (19%)	89 (12%)		66 (8%)	113 (29%)	
Staying about the same	221 (47%)	315 (44%)		348 (43%)	188 (48%)	
Getting worse	162 (34%)	311 (43%)		383 (48%)	90 (23%)	
Treatment preference			< 0.001			< 0.001
Preference for nonsurgery	193 (41%)	202 (28%)		130 (16%)	265 (68%)	
Not sure	154 (32%)	43 (6%)		114 (14%)	83 (21%)	
Preference for surgery	127 (27%)	473 (66%)		556 (69%)	44 (11%)	

SPORT = Spine Patient Outcomes Research Trial; IDH = intervertebral disc herniation.

longitudinal regression models of primary outcomes. Those that remained significant in the longitudinal models were included as adjusting covariates in all subsequent longitudinal regression models to adjust for potential confounding resulting from treatment selection bias and missing data patterns [10]. In addition, center, age, sex, and, when applicable, baseline outcome score were included in all longitudinal outcome models.

Given the substantial crossover in both directions between the randomized treatment groups, an as-treated analysis was also performed. In this analysis, the nonoperative cohort treatment time started at the time of trial enrollment and surgery was a time-varying covariate adjusting for the variable time of surgery. The sciatica

scales and all binary outcomes were analyzed with longitudinal models based on generalized estimating equations with linear and logit link functions. The operative and nonoperative cohorts were then compared using a Wald test to simultaneously test all followup visit times for differences in estimated treatment effects between the cohorts.

Subgroups were defined based on baseline characteristics, including sex, work status, smoking status, presence of a comorbid joint problem, history of depression, severity of low back pain, presence of a motor deficit, herniation level, herniation morphology, duration of symptoms at enrollment, and herniation location. Herniation level was analyzed in three groups: L2–L3/L3–L4, L4–L5, and L5–

S1. Herniation morphology was divided into three groups: protruding, extruded, or sequestered. Duration of symptoms at enrollment was divided into \leq versus $>$ 6 months. Herniation location was characterized as posterolateral versus all other locations.

To limit the number of statistical comparisons in the subgroup analyses, the ODI was used as the outcome measure and was calculated as a single time-weighted average, ie, area under the curve, over the entire 8 years. All of these subgroup cohorts were adjusted for age, sex, race, marital status, compensation, smoking status, herniation location, working status, stomach comorbidity, depression, diabetes, other comorbidities, self-rated health trend, duration of most recent episode, treatment preference, center, and baseline score except when that variable was the basis of the subgroup being compared.

Computations were completed with SAS® procedures PROC MIXED for continuous data and PROC GENMOD for binary and nonnormal secondary outcomes (SAS® Version 9.1 for Windows® XP Professional; SAS Institute, Inc, Cary, NC, USA). Statistical significance was defined as p values of < 0.05 based on a two-sided hypothesis test with no adjustments made for multiple comparisons.

Study Attrition

By 8 years only 63% of initial enrollees supplied data with losses resulting from dropouts, missed visits, or deaths [20]: 92 missed visits, 360 withdrawals, and 14 deaths. We compared those retained at 8 years and those lost to followup. Those retained in the study were, at baseline, somewhat older (mean age, 4.2 [SD 11.2] versus 40.7 [SD 11.7] years, $p = 0.039$); were more likely to be female (369 [45%] versus 138 [36%], $p = 0.005$), white (725 [89%] versus 311 [82%], $p = 0.002$), college-educated (625 [77%] versus 262 [69%], $p = 0.007$), and working (536 [66%] versus 189 [50%] $p < 0.001$); were less likely to be receiving compensation (115 [14%] versus 93 [25%], $p < 0.001$) or a smoker (163 [20%] versus 120 [32%], $p < 0.001$); had better SF-36 bodily pain score (mean) (28.1 [SD 20.6] versus 25.1 [SD 19], $p = 0.015$); ODI mean score (48.4 [SD 21] versus 51.7 [21.9], $p = 0.011$); and had less bothersome sciatica per Sciatic Bothersome Index (mean 15.3 [SD 5.2] versus 16.1 [SD 5.3], $p = 0.022$).

Results

Surgery versus Nonoperative Treatment: Overall Results (Table 2)

As previously reported, in the intent-to-treat analysis, no significant benefit to surgery was observed in the primary

outcome measures at 8 years (SF-36 bodily pain, 41.1 [SE 1.6] versus 38.7 [SE 1.6], treatment effect 2.8, 95% confidence interval [CI], -1.6 to 7.2 , $p = 0.22$; SF-36 physical function 35.3 [SE 1.6] versus 33.7 [SE 1.6], treatment effect 1.6, 95% CI, -2.8 to 6 , $p = 0.47$; ODI -30.3 [SE 1.4] versus -27 [SE 1.4], treatment effect -3.3 , 95% CI, -7.2 to 0.6 , $p = 0.096$) (Table 2). Despite high levels of crossover, which are expected to bias any comparison toward the null, sciatica bothersomeness, leg pain, patient satisfaction with symptoms, and self-rated improvement all showed significant treatment effects in favor of surgery in the intent-to-treat analysis out to 8 years [20].

When the randomized and observational cohorts were combined in the as-treated analysis, the surgical cohort had greater improvement across all primary outcomes than the nonoperative cohort at 8 years: SF-36 bodily pain, 45.3 (SE 0.6) versus 33.3 (SE 0.8), treatment effect 12.1 (95% CI, 10.3–13.9; $p < 0.001$); SF-36 physical function, 41.8 (SE 0.6) versus 30.5 (SE 0.8), treatment effect 11.3 (95% CI, 9.6–13; $p < 0.001$); and ODI, -36.1 (SE 0.5) versus -25 (SE 0.6), treatment effect -11 (95% CI, -12.4 to -9.6 ; $p < 0.001$) (Table 2). However, although adjusted for numerous covariates, an as-treated analysis is essentially an observational comparison that remains potentially vulnerable to unmeasured confounding. No difference in work status over time was seen between the treatment groups.

Surgery versus Nonoperative Treatment: Analysis by Subgroups

Sex, baseline motor deficit, and herniation level had no effect on outcome for either surgical or nonoperative treatment (Table 3). The subgroups that had less improvement in ODI across both treatment groups but no difference in treatment effect were smokers, herniation location other than posterolateral, baseline history of depression, and baseline comorbid joint problem. These patients did worse independent of treatment. The subgroups found to have an interaction with treatment were disc morphology, baseline work status, baseline back pain, and duration of symptoms. For disc morphology, those with sequestered disc fragments had the greatest improvements with both surgery and nonoperative treatment and also had the highest relative surgical treatment effect; those with extruded fragments had the smallest improvement with surgery and very good outcomes from nonoperative treatment, leading to the smallest relative surgical treatment effect (-8.7 [95% CI, -10.3 to -7] versus -13.6 [95% CI, -18.2 to -9] for sequestered fragment; $p = 0.05$); and protruding discs showed the smallest relative improvement with nonoperative treatment and an intermediate-sized treatment effect (Table 3).

Table 2. Time-weighted average of primary analysis results at 8 years (area under the curve)*

IDH	Baseline overall mean	Mean change in score compared with baseline (SE)		Treatment effect (95% CI) [‡]	p value
		Surgery	Nonoperative		
RCT intent-to-treat					
Primary outcomes		(n = 233)	(n = 241)		
SF-36 bodily pain (0–100) (SE) [§]	28.3 (0.92)	41.4 (1.6)	38.7 (1.6)	2.8 (–1.6 to 7.2)	0.22
SF-36 physical function (0–100) (SE) [§]	39.5 (1.2)	35.3 (1.6)	33.7 (1.6)	1.6 (–2.8 to 6)	0.47
Oswestry Disability Index (0–100) (SE)	46.9 (0.96)	–30.3 (1.4)	–27 (1.4)	–3.3 (–7.2 to 0.6)	0.096
Secondary outcomes					
Sciatica Bothersomeness Index (0–24) (SE) [¶]	15.2 (0.24)	–9.9 (0.4)	–8.3 (0.4)	–1.5 (–2.5 to –0.5)	0.005
Leg pain (0–6) (SE) ^{**}	4.6 (0.1)	–3.2 (0.1)	–2.7 (0.1)	–0.5 (–0.8 to –0.2)	0.006
Low back pain bothersomeness (0–6) (SE) ^{††}	3.9 (0.1)	–1.8 (0.1)	–1.7 (0.1)	–0.1 (–0.5 to 0.3)	0.42
Very/somewhat satisfied with symptoms (%)	3.4 (1.8)	68.1	60.3	7.7 (1.6–13.8)	0.013
Very/somewhat satisfied with care (%)		85.1	82.5	2.6 (–1.8 to 7)	0.25
Self-rated progress: major improvement (%)		68.7	61.3	7.4 (1.6–13.2)	0.013
Work status: working (%)	64.3 (4.8)	73.6	73.6	0 (–6.5 to 6.5)	0.99
RCT/OC as-treated					
Primary outcomes		(n = 803)	(n = 392)		
SF-36 bodily pain (0–100) (SE) [§]	28.3 (0.6)	45.3 (0.6)	33.3 (0.8)	12.1 (10.3–13.9)	< 0.001
SF-36 physical function (0–100) (SE) [§]	38.3 (0.7)	41.8 (0.6)	30.5 (0.8)	11.3 (9.6–13)	< 0.001
Oswestry Disability Index (0–100) (SE)	49.1 (0.6)	–36.1 (0.5)	–25 (0.6)	–11 (–12.4 to –9.6)	< 0.001
Secondary outcomes					
Sciatica Bothersomeness Index (0–24) (SE) [¶]	15.5 (0.1)	–10.3 (0.2)	–8.8 (0.2)	–1.6 (–2 to –1.2)	< 0.001
Leg pain (0–6) (SE) ^{**}	4.7 (0)	–3.4 (0)	–2.7 (0.1)	–0.7 (–0.9 to –0.5)	< 0.001
Low back pain bothersomeness (0–6) (SE) ^{††}	3.8 (0)	–2 (0)	–1.5 (0.1)	–0.5 (–0.7 to –0.3)	< 0.001
Very/somewhat satisfied with symptoms (%)	4.6 (2.1)	74.5	48.2	26.3 (22.1–30.5)	< 0.001
Very/somewhat satisfied with care (%)		90.2	75.9	14.3 (11–17.6)	< 0.001
Self-rated progress: major improvement (%)		74.1	53.3	20.8 (16.7–24.9)	< 0.001
Work status: working (%)	73.7 (4.4)	82.5	79.8	2.7 (–0.8 to 6.2)	0.13

* Intent-to-treat for the randomized cohort and adjusted †as-treated analyses for the randomized and observational cohorts combined; †adjusted for age, sex, race, marital status, compensation, smoking status, herniation location, working status, stomach comorbidity, depression, diabetes, other^{‡‡‡} comorbidity, self-rated health trend, duration of most recent episode, treatment preference, baseline score (for SF-36, Oswestry Disability Index, and Sciatica Bothersomeness Index), and center; ‡treatment effect is the difference between the surgical and nonoperative mean change from baseline; §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; ¶the Sciatica 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; ‡‡‡other comorbidities include: stroke, diabetes, osteoporosis, cancer, fibromyalgia, chronic fatigue syndrome, posttraumatic stress disorder, alcohol, drug dependency, heart, lung, liver, kidney, blood vessel, nervous system, hypertension, migraine, anxiety, stomach, bowel.; IDH = intervertebral disc herniation; CI = confidence interval; RCT = randomized clinical trial; OC = observational cohort.

For baseline work status, patients who were working at baseline showed the greatest relative improvements with both surgery and nonoperative treatment, but the large nonoperative improvement led to the smallest surgical treatment effect; those disabled at baseline had the smallest improvements with both surgery and nonoperative treatment with an intermediate-sized treatment effect; and those in the “other” category (eg, homemakers, students, unemployed, retired) had very good surgical outcomes but modest nonoperative outcomes, leading to the highest

relative surgical treatment effect (–15.2 versus –9.8 for working; $p = 0.003$) (Fig. 1).

Those with severe baseline back pain (≥ 5 of 6) had worse surgical outcomes (–34.6 versus –36.8; $p = 0.04$) but even worse outcomes with nonoperative treatment (–21.8 versus –27.6; $p < 0.001$), resulting in a greater surgical treatment effect compared with the less severe baseline back pain group (–12.8 versus –9.2; $p = 0.005$) (Fig. 2).

Those with symptoms less than or equal to 6 months in duration had better outcomes with surgical (mean change

Table 3. Time-weighted average of Oswestry Disability Index at 8 years (area under the curve) from adjusted* as-treated analysis according to subgroup

Subgroup	Number	Mean change in score compared with baseline (SE)		Treatment effect† (95% CI)
		Surgical	Nonoperative	
Female	507	-35.7 (0.8)	-24.1 (1)	-11.6 (-13.6 to -9.6)
Male	688	-36.3 (0.7)	-25.8 (0.8)	-10.5 (-12.3 to -8.8)
p value		0.56	0.18	0.4
Smoker	283	-32.3 (1)	-21 (1.3)	-11.3 (-14.1 to -8.6)
Nonsmoker	907	-37 (0.6)	-26.1 (0.7)	-10.9 (-12.4 to -9.4)
p value		< 0.001	< 0.001	0.78
No joint problem	965	-37.3 (0.6)	-26 (0.7)	-11.3 (-12.8 to -9.8)
Joint problem	222	-30.5 (1.2)	-21 (1.3)	-9.4 (-12.1 to -6.7)
p value		< 0.001	< 0.001	0.22
Not depressed	1047	-36.5 (0.5)	-25.4 (0.7)	-11.1 (-12.5 to -9.6)
Depressed	141	-32.6 (1.5)	-21.6 (1.8)	-11 (-14.5 to -7.4)
p value		0.013	0.045	0.97
Baseline LBP ≤ 4	663	-36.8 (0.7)	-27.6 (0.8)	-9.2 (-10.9 to -7.5)
Baseline LBP ≥ 5	525	-34.6 (0.8)	-21.8 (1)	-12.8 (-14.9 to -10.8)
p value		0.04	< 0.001	0.005
Motor deficit	502	-35.8 (0.8)	-24.2 (1)	-11.6 (-13.7 to -9.5)
Motor intact	692	-36.3 (0.7)	-25.6 (0.8)	-10.7 (-12.4 to -9)
p value		0.59	0.25	0.49
Posterolateral herniation	921	-36.7 (0.6)	-25.9 (0.7)	-10.8 (-12.3 to -9.3)
Other herniation location	273	-33.8 (1.1)	-22.2 (1.2)	-11.6 (-14.3 to -9)
p value		0.017	0.008	0.59
Working	749	-37.5 (0.6)	-27.7 (0.8)	-9.8 (-11.4 to -8.2)
Disabled	158	-26.6 (1.5)	-16.2 (2.1)	-10.4 (-14.5 to -6.4)
Other	288	-36.3 (1)	-21 (1.3)	-15.2 (-18 to -12.4)
p value		< 0.001	< 0.001	0.003
L2-L3/L3-L4 herniation	88	-35.7 (2.2)	-23.5 (2.1)	-12.3 (-17.5 to -7.1)
L4-L5 herniation	457	-37.2 (0.8)	-25.9 (1)	-11.3 (-13.3 to -9.3)
L5 S1 herniation	649	-35.2 (0.7)	-25 (0.9)	-10.2 (-12.1 to -8.4)
p value		0.16	0.54	0.61
Extruded	786	-34.9 (0.6)	-26.2 (0.8)	-8.7 (-10.3 to -7)
Protruding	322	-35 (1)	-23.8 (1.1)	-11.2 (-13.4 to -9)
Sequestered	86	-41 (1.9)	-27.4 (2.3)	-13.6 (-18.2 to -9)
p value		0.01	0.15	0.05
Duration of symptoms ≤ 6 months	930	-37.1 (0.6)	-26.7 (0.7)	-10.4 (-12 to -8.9)
Duration of symptoms > 6 months	265	-31.7 (1.1)	-18.2 (1.3)	-13.5 (-16.2 to -10.8)
p value		< 0.001	< 0.001	0.043

* Adjusted for age, sex, race, marital status, compensation, smoking status, herniation location, stomach comorbidity, depression, diabetes, other[§] comorbidity, self-rated health trend, duration of most recent episode, treatment preference, baseline score (for SF-36, Oswestry Disability Index[‡], and Sciatica Bothersomeness Index), and center unless the variable itself defined the subgroup; †treatment effect is the difference between the surgical and nonoperative mean change from baseline; ‡the Oswestry Disability Index ranges from 0 to 100 with lower scores indicating less severe symptoms; §other comorbidities include stroke, diabetes, osteoporosis, cancer, fibromyalgia, CFS, posttraumatic stress disorder, alcohol, drug dependency, heart, lung, liver, kidney, blood vessel, nervous system, hypertension, migraine, anxiety, stomach, bowel; CI = confidence interval; LBP = low back pain; CFS = chronic fatigue syndrome.

in ODI -37.1 [SE 0.6] versus -31.7 [SE 1.1], $p < 0.001$) but also had better outcomes in the nonoperative treatment cohort (mean change in ODI -26.7 [SE 0.7] versus -18.2 [SE 1.3], $p < 0.001$). On balance, however, those with

longer duration of symptoms at baseline (> 6 months) had a greater relative advantage from surgery (significantly greater surgical treatment effect) than those with less than or equal to 6 months of symptoms: -13.5 (95% CI, -16.2

to -10.8) versus -10.4 (95% CI, -12 to -8.9; $p = 0.043$; Table 3; Fig. 3).

Discussion

Although previous studies have consistently shown more rapid improvement in patients with herniated discs with surgery relative to nonoperative treatment [3, 20, 25], the long-term data are conflicting. The randomized, intent-to-treat comparison by Weber showed no long-term benefit [25], although this interpretation is complicated by a high rate of crossover in the nonoperative arm, whereas the Maine study showed persistent benefit from surgery in the long term, although this observational study cannot rule out confounding by unbalanced baseline characteristics as contributing to these differences [3]. Long-term results from SPORT [20], summarized here, are consistent with both of these prior studies, although they lend some support for a small advantage to surgery in the long term. Similar to the Weber study, the intent-to-treat results showed no significant advantage for being randomized to surgery in the primary outcome measures, but crossover rates were high; however, secondary outcomes of sciatica bothersomeness, leg pain,

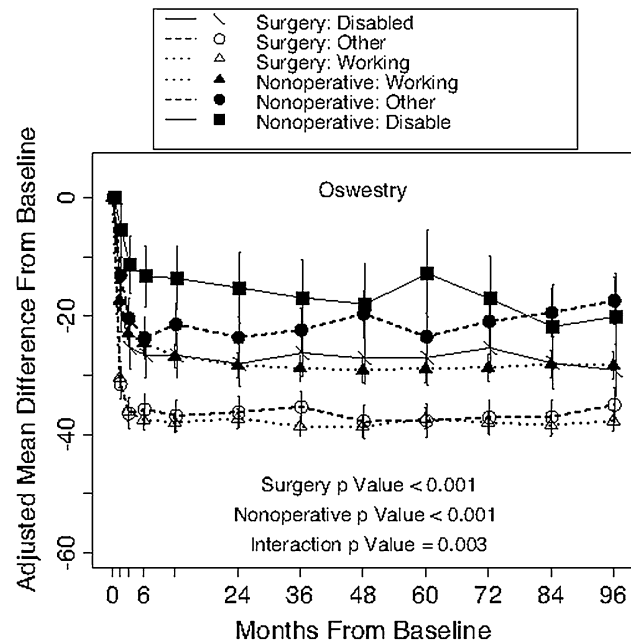


Fig. 1 Change scores in the ODI over time are displayed according to duration of symptoms. The surgery p value compares duration of symptoms ≤ 6 months with duration of symptoms > 6 months. The nonoperative p value compares duration of symptoms ≤ 6 months with duration of symptoms > 6 months. The interaction p value compares treatment effect (surgery versus nonoperative) between duration of symptoms ≤ 6 months and duration of symptoms > 6 months. All p values are time-weighted averages over 8 years (ie, area under the curve).

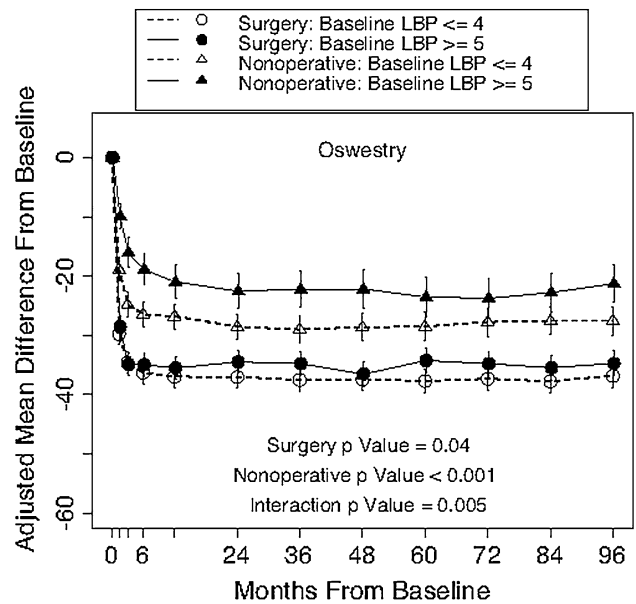


Fig. 2 Change scores in the ODI over time are displayed according to baseline low back pain (LBP). The surgery p value compares baseline LBP ≤ 4 with baseline LBP ≥ 5 among surgery patients. The nonoperative p value compares baseline LBP ≤ 4 with baseline LBP ≥ 5 among nonoperative patients. The interaction p value compares the treatment effect (surgery versus nonoperative) between baseline LBP ≤ 4 and baseline LBP ≥ 5 . All p values are time-weighted averages over 8 years (ie, area under the curve).

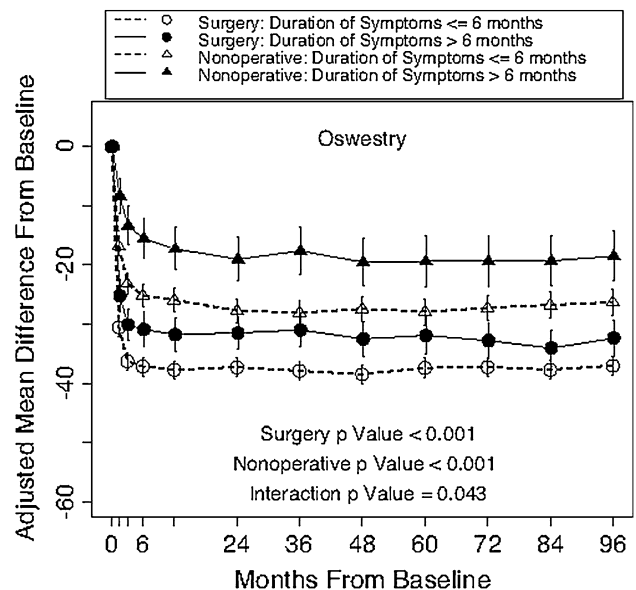


Fig. 3 Change scores in the ODI over time are displayed according to work status. The surgery p value compares working, other, and disabled statuses among surgery patients. The nonoperative p value compares working, other, and disabled statuses among nonoperative patients. The interaction p value compares the treatment effect (surgery versus nonoperative) among working, other, and disabled statuses. All p values are time-weighted averages over 8 years (ie, area under the curve).

satisfaction with symptoms, and self-rated improvement favored the group randomized to surgery in the intent-to-treat analysis despite the high rates of crossover, although these were secondary outcomes and the results were not corrected for multiplicities. Similar to the Maine study, an as-treated comparison based on treatment actually received showed a strong effect in favor of surgery that persisted at long-term followup, but these analyses lose the strong protection against confounding of a randomized comparison.

The main purpose of this report was to investigate possible baseline predictors of a long-term surgical treatment effect. Sex, presence of a motor deficit, and herniation level had no discernible effects on outcomes over 8 years. Herniation level was shown to predict early outcomes with upper-level herniation, having a greater treatment effect at 1 and 2 years compared with L5/S1 herniation [19]; however, over the long term, these differences resolved. The subgroup with a history of depression at baseline had worse outcomes independent of treatment. Chen and Tsai [9] studied patients with lumbar IDH and found depression was correlated with disability in this patient population. In another study examining the effect of surgical microdiscectomy for IDH in patients with concomitant baseline depression and anxiety, Lebow et al [18] found that surgical intervention was correlated with relieving depression symptoms. Chaichana et al [8] performed a prospective trial of patients undergoing discectomy for single-level IDH. At 2 years, they found a correlation between a higher preoperative depression assessed by the Zung scale and failure to achieve postoperative clinical improvement in disability by ODI [8]. Neither the study of Chaichana et al [8] nor that of Lebow et al [18] contained a nonoperative comparison group.

With regard to smoking, previously published papers have correlated smoking with increased low back pain, visual analog scale, ODI, and disc degeneration when compared with not smoking [5, 6, 14]. Akmal et al [2] also demonstrated the toxicity of nicotine on bovine in vitro intervertebral disc cells. We found that nonsmokers had better outcomes than smokers, which is consistent with previous studies [6, 17, 24]. Nonetheless, we found that smokers and nonsmokers had no difference in the surgical treatment effect. In other words, although smokers had, on average, worse ODI scores over 8 years compared with nonsmokers with both surgical and nonoperative treatment, both subgroups had the same amount of relative improvement with operative treatment compared with nonoperative treatment.

Previous studies have found no difference in outcomes in patients with a longer duration of symptoms from lumbar disc herniation when compared with those with a shorter duration of symptoms [11, 23]. Folman et al [11]

performed a retrospective study with short-term followup (average 32.4 months) that showed those with 6 weeks or less of pain history had a larger treatment effect than those who had greater than 12 weeks of preoperative pain. However, in their study, patients with 12 weeks or more of preoperative pain did have more improvement than those with 6 to 12 weeks of pain. In their randomized, controlled trial, Peul et al [23] found that despite short-term differences between patients who underwent early surgery and those who underwent surgery after a course of conservative treatment, there was no significant difference between the cohorts at 52 weeks of followup. In another retrospective study by Akagi et al [1], they found that patients with symptoms greater than 3 months had, on average, greater postoperative improvement for psychological disorders; however, no difference in severity of back pain, leg pain and numbness (visual analog scale), lumbar spine dysfunction, or social life disturbance. In the current study, although those who had a longer duration of symptoms fared worse than those with a shorter duration of symptoms in both treatment groups, the relative benefit from surgery was actually somewhat larger in the group with a longer duration of symptoms at baseline.

Although there was not an advantage to surgery in the as-treated analysis in all the evaluated subgroups, these analyses did identify some interactions with treatment effect that may help inform the choice of treatment in borderline cases. Interestingly, those with higher levels of baseline back pain accompanying their radiculopathy, although they had poorer results overall in both treatment groups, showed a greater relative advantage from surgery. Those who were working at baseline showed relatively less advantage to surgery as a result of relatively better outcomes with nonoperative treatment, whereas those who were disabled at baseline did worse overall with relatively less advantage from surgery. Finally, those with sequestered fragments on MRI showed greater relative advantage from surgery than those with other disc morphologies.

This study had a number of limitations. Overall, SPORT was limited by the amount of crossover within the study. Given the amount of crossover, the interpretation of the intent-to-treat analysis was limited; thus, as-treated analyses were also performed. These as-treated analyses are susceptible to confounding. We attempted to mitigate this as much as possible by adjusting for potential confounders in the longitudinal regression models; however, the potential for residual confounding by unmeasured covariates remains. Another limitation of SPORT was the heterogeneity of the “usual care” in the nonoperative cohort. Although nonoperative treatment was required to include physical therapy, education, home exercises, and nonsteroidal antiinflammatory drugs (if tolerated), the actual protocol and content of these regimens may have varied across centers.

Another important limitation was the limited followup at long-term followup; by 8 years only 63% of the original cohort continued to contribute followup data. As previously reported, those lost to followup had worse early outcomes on average; however, this was true in both the surgical and nonoperative groups with nonsignificant differences in estimated treatment effects. The long-term outcomes are therefore likely to be somewhat overoptimistic on average in both groups, but the comparison between surgical and nonoperative outcomes appears likely to be unbiased despite the long-term loss to follow [20].

In conclusion, in long-term followup, SPORT patients with symptomatic lumbar disc herniation randomized to surgery had similar treatments and similar outcomes as those randomized to nonoperative care; those actually undergoing surgery had better outcomes than those having only nonoperative treatment, which persisted through 8 years. These findings are similar to the long-term results seen in the MLSS; however, nonoperative patients in the SPORT had better outcomes on average than the nonoperative group in the MLSS [3]. Smokers, patients with a history of depression, a history of other joint disease, and a nonposterolateral herniation had poorer outcomes regardless of treatment than those without these conditions. Those with sequestered fragments on MRI, higher levels of baseline back pain accompanying their radiculopathy, those with greater than 6 months of symptoms at baseline, and those who were neither working nor disabled at baseline showed a greater relative advantage from surgery.

References

- Akagi R, Aoki Y, Ikeda Y, Nakajima F, Ohtori S, Takahashi K, Yamagata M. Comparison of early and late surgical intervention for lumbar disc herniation: is earlier better? *J Orthop Sci*. 2010;15:294–298.
- Akmal M, Kesani A, Anand B, Singh A, Wiseman M, Goodship A. Effect of nicotine on spinal disc cells: a cellular mechanism for disc degeneration. *Spine (Phila Pa 1976)*. 2004;29:568–575.
- Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE. 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 (Phila Pa 1976)*. 2005;30:927–935.
- Atlas SJ, Tosteson TD, Blood EA, Skinner JS, Pransky GS, Weinstein JN. The impact of Workers' Compensation on outcomes of surgical and nonoperative therapy for patients with a lumbar disc herniation: SPORT. *Spine (Phila Pa 1976)*. 2010;35:89–97.
- Battie MC, Videman T, Kaprio J, Gibbons LE, Gill K, Manninen H, Saarela J, Peltonen L. The Twin Spine Study: contributions to a changing view of disc degeneration. *Spine J*. 2009;9:47–59.
- Behrend C, Prasarn M, Coyne E, Horodyski M, Wright J, Rehtine GR. Smoking cessation related to improved patient-reported pain scores following spinal care. *J Bone Joint Surg Am*. 2012;94:2161–2166.
- Birkmeyer NJ, Weinstein JN, Tosteson AN, Tosteson TD, Skinner JS, Lurie JD, Deyo R, Wennberg JE. Design of the Spine Patient Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976)*. 2002;27:1361–1372.
- Chaichana KL, Mukherjee D, Adogwa O, Cheng JS, McGirt MJ. Correlation of preoperative depression and somatic perception scales with postoperative disability and quality of life after lumbar discectomy. *J Neurosurg Spine*. 2011;14:261–267.
- Chen HN, Tsai YF. A predictive model for disability in patients with lumbar disc herniation. *J Orthop Sci*. 2013;18:220–229.
- Fitzmaurice G. Adjusting for confounding. *Nutrition*. 2004;20:594–596.
- Folman Y, Shabat S, Catz A, Gepstein R. Late results of surgery for herniated lumbar disk as related to duration of preoperative symptoms and type of herniation. *Surg Neurol*. 2008;70:398–401; discussion 401–492.
- Frymoyer JW. Back pain and sciatica. *N Engl J Med*. 1988;318:291–300.
- Frymoyer JW. Lumbar disk disease: epidemiology. *Instr Course Lect*. 1992;41:217–223.
- Frymoyer JW, Pope MH, Costanza MC, Rosen JC, Goggin JE, Wilder DG. Epidemiologic studies of low-back pain. *Spine (Phila Pa 1976)*. 1980;5:419–423.
- Gray DT, Deyo RA, Kreuter W, Mirza SK, Heagerty PJ, Comstock BA, Chan L. Population-based trends in volumes and rates of ambulatory lumbar spine surgery. *Spine (Phila Pa 1976)*. 2006;31:1957–1963; discussion 1964.
- Hall MJ, DeFrances CJ, Williams SN, Golosinskiy A, Schwartzman A. National hospital discharge survey: 2007 summary. *Natl Health Stat Report*. 2010;29:1–20.
- Jansson KA, Nemeth G, Granath F, Jonsson B, Blomqvist P. Health-related quality of life in patients before and after surgery for a herniated lumbar disc. *J Bone Joint Surg Br*. 2005;87:959–964.
- Lebow R, Parker SL, Adogwa O, Reig A, Cheng J, Bydon A, McGirt MJ. Microdiscectomy improves pain-associated depression, somatic anxiety, and mental well-being in patients with herniated lumbar disc. *Neurosurgery*. 2012;70:306–311; discussion 311.
- Lurie JD, Faucett SC, Hanscom B, Tosteson TD, Ball PA, Abdu WA, Frymoyer JW, Weinstein JN. Lumbar discectomy outcomes vary by herniation level in the Spine Patient Outcomes Research Trial. *J Bone Joint Surg Am*. 2008;90:1811–1819.
- Lurie JD, Tosteson TD, Tosteson AN, Zhao W, Morgan TS, Abdu WA, Herkowitz H, Weinstein JN. Surgical versus nonoperative treatment for lumbar disc herniation: eight-year results for the Spine Patient Outcomes Research Trial. *Spine (Phila Pa 1976)*. 2014;39:3–16.
- Pearson A, Lurie J, Tosteson T, Zhao W, Abdu W, Mirza S, Weinstein J. Who should have surgery for an intervertebral disc herniation? Comparative effectiveness evidence from the Spine Patient Outcomes Research Trial. *Spine (Phila Pa 1976)*. 2012;37:140–149.
- Pearson AM, Blood EA, Frymoyer JW, Herkowitz H, Abdu WA, Woodward R, Longley M, Emery SE, Lurie JD, Tosteson TD, Weinstein JN. Sport lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? *Spine (Phila Pa 1976)*. 2008;33:428–435.
- Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, Thomeer RT, Koes BW. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med*. 2007;356:2245–2256.
- Sorlie A, Moholdt V, Kvistad KA, Nygaard OP, Ingebrigtsen T, Iversen T, Kloster R, Solberg TK. Modic type I changes and recovery of back pain after lumbar microdiscectomy. *Eur Spine J*. 2012;21:2252–2258.
- Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine (Phila Pa 1976)*. 1983;8:131–140.
- Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA. Surgical vs nonoperative treatment for lumbar disk

- herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA*. 2006;296:2451–2459.
27. Weinstein JN, Lurie JD, Tosteson TD, Tosteson AN, Blood EA, Abdu WA, Herkowitz H, Hilibrand A, Albert T, Fischgrund J. Surgical versus nonoperative treatment for lumbar disc herniation: four-year results for the Spine Patient Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976)*. 2008;33:2789–2800.
28. Weinstein JN, Tosteson TD, Lurie JD, Tosteson A, Blood E, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goldberg H, Berven S, An H. Surgical versus nonoperative treatment for lumbar spinal stenosis four-year results of the Spine Patient Outcomes Research Trial. *Spine (Phila Pa 1976)*. 2010;35:1329–1338.
29. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA*. 2006;296:2441–2450.