Who should have surgery for spinal stenosis?:

Treatment Effect Predictors in SPORT

Adam Pearson, MD, MS, Jon Lurie, MD, MS, Tor Tosteson, ScD, Wenyan Zhao, MS, William Abdu, MD, MS, and **James Weinstein, DO, MS**
Dartmouth Medical School & The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, NH

Abstract

Study Design—Combined prospective randomized controlled trial and observational cohort study of spinal stenosis (SpS) with an as-treated analysis.

Objective—To determine modifiers of the treatment effect (TE) of surgery (the difference between surgical and nonoperative outcomes) for SpS using subgroup analysis.

Summary of Background Data—SPORT demonstrated a positive surgical TE for SpS at the group level. However, individual characteristics may affect TE. No prior studies have evaluated TE modifiers in SpS.

Methods—SpS patients were treated with either surgery (n=419) or nonoperative care (n=235) and were analyzed according to treatment received. Fifty-three baseline variables were used to define subgroups for calculating the time-weighted average TE for the Oswestry Disability Index (ODI) over 4 years (TE=$\Delta ODI_{surgery}$-$\Delta ODI_{nonoperative}$). Variables with significant subgroup by* treatment interactions (p<0.05) were simultaneously entered into a multivariate model to select independent TE predictors.

Results—Other than smokers, all analyzed subgroups including at least 50 patients improved significantly more with surgery than with nonoperative treatment (p<0.05). Multivariate analysis demonstrated: baseline ODI 56 (TE −15.0 vs. −4.4 ODI > 56, p<0.001), not smoking (TE −11.7 vs. −1.6 smokers, p<0.001), neuroforaminal stenosis (TE −14.2 vs. −8.7 no neuroforaminal stenosis, p=0.002), predominant leg pain (TE −11.5 vs. −7.3 predominant back pain, p=0.035), not lifting at work (TE −12.5 vs. −8.5 lifting at work, p=0.017), and the presence of a neurological deficit (TE −13.3 vs. −7.2 no neurological deficit, p<0.001) were associated with greater TE.

Conclusions—With the exception of smokers, patients who met strict inclusion criteria improved more with surgery than with nonoperative treatment, regardless of other specific characteristics. However, TE varied significantly across certain subgroups, and these data can be used to individualize shared decision making discussions about likely outcomes. Smoking cessation should be considered prior to surgery for SpS.

Corresponding Author: Adam Pearson, MD, MS Dept. of Orthopaedics, Dartmouth-Hitchcock Medical Center One Medical Center Dr. Lebanon, NH 03756 603-650-8942 adam.m.pearson@hitchcock.org.

INTRODUCTION

The Spine Patient Outcomes Research Trial (SPORT), the Maine Lumbar Spine Study (MLSS), and a recent Finnish randomized controlled trial demonstrated better outcomes for spinal stenosis (SpS) patients treated surgically than those treated nonoperatively.[1-5] However, these findings were all at the group level, and it has been established that individual patient characteristics are strongly related to surgical and nonoperative outcomes.[6-13] In the SPORT SpS cohort, diabetes, the number of stenotic levels, and predominant pain location (leg vs. back) have been specifically evaluated as outcome predictors.[14-16] Both SPORT and prior studies have empirically selected variables for evaluation as potential outcome predictors based on clinical judgment, yet demographic and psychosocial characteristics may be more strongly associated with outcomes than clinical or disease characteristics.[17-19] As such, the current study aims to evaluate a large pool of demographic, psychosocial, and clinical variables to determine which are predictors of surgical and nonoperative outcomes in SpS.

Most prior studies evaluating outcomes predictors in SpS have focused on surgical outcomes, usually due to the lack of a nonoperative cohort.[6,10,13] While it is useful to counsel patients about their likely surgical outcomes, the factor that should drive decision making is their likely benefit from surgery *relative* to nonoperative outcomes, the treatment effect (TE) of surgery. If only surgical outcomes are considered, surgeons could erroneously suggest surgery to patients who may do well with nonoperative treatment and nonoperative treatment to patients who are likely to fail without surgery. In fact, it has been shown that certain predictors of worse surgical outcomes were actually associated with a greater TE in disk herniation patients, indicating that considering only surgical outcome predictors in the decision making process could lead to improper treatment choices.[20] Given the importance of predicting TE in surgical decision making, the goal of this study was to determine which variables were significant, independent TE modifiers in SpS.

MATERIALS AND METHODS

Study Design

The SPORT SpS investigation consisted of a randomized controlled trial with a concurrent observational cohort study conducted in 11 states at 13 institutions with multidisciplinary spine practices.[21] In the first year of follow-up in the randomized trial, 37% of patients assigned to surgery did not have surgery, and 42% of patients assigned to nonoperative treatment did have surgery.[5] By 4 years, 32% of the patients randomized to surgery had still not undergone surgery, and 49% of the patients assigned to nonoperative treatment had received surgery.[4] Given this rate of protocol nonadherence and the consistency of the findings between the randomized and the observational cohorts, the data were combined in an as-treated analysis in this study.[4,5] The rationale for this decision has been discussed previously.[22]

Patient Population

Patients were considered for inclusion in the study if they had neurogenic claudication or radicular pain for at least twelve weeks and a confirmatory cross-sectional imaging study demonstrating stenosis at one or more levels. Exclusion criteria included cauda equina syndrome, malignancy, significant deformity, prior back surgery, instability on flexion-extension radiographs (defined as greater than 4 mm of translation or 10 degrees of rotation), and other established contra-indications to elective surgery.[21] Patients with degenerative spondylolisthesis were studied separately.[23]

Study Interventions

Surgery consisted of a standard open decompressive laminectomy.[21] The nonoperative treatment group received "usual care"--defined as including at least physical therapy, education and counseling with home exercise instruction, and non-steroidal anti-inflammatory drugs if tolerated. Physicians were instructed to individualize nonoperative treatment and explore a wide range of non-operative options.[21]

Imaging Studies

All patients underwent standing x-rays and cross-sectional imaging. The cross-sectional imaging was evaluated to determine which levels were stenotic and the location (central, lateral recess, and/or neuroforamen) and severity of the stenosis (mild, moderate or severe).[24,25] The kappa scores for intra-rater reliability of the location and severity classification have been reported to range from 0.75 to 0.82, while inter-rater reliability ranged from 0.49 to 0.73.[25]

Baseline Characteristics and Primary Outcome Measure

At baseline, patients and clinicians answered questionnaires evaluating demographic, socioeconomic, clinical, and radiographic characteristics.[5,21] Baseline scores on the Short-Form 36 (SF-36),[26] AAOS/Modems version of the Oswestry Disability Index (ODI),[27] Stenosis Bothersomeness Index,[28,29] and Leg and Back Pain Bothersomeness Scale (0-6 point Likert-type scale) were also recorded. The ODI was also recorded at 6 weeks, 3 months, 6 months, 1 year and yearly thereafter out to 4 years. Fifty-three variables were selected as potential TE modifiers. Continuous variables were converted to categorical variables in order to allow comparison of TE between subgroups, which would not have been possible with continuous variables. This was done as follows: ODI defined by quartiles, with comparison between the highest quartile and the other three quartiles combined; Stenosis Bothersomeness and Age by the median; Duration of Symptoms as less than or greater than 6 months; Income as greater than $50,000, less than $50,000 or not working; Body Mass Index (BMI) as less than 30 or greater than or equal to 30; and the SF-36 Mental Component Score (MCS) as less than or equal to 35 or greater than 35.[30] In the initial analyses, all four ODI quartiles were evaluated separately, but TE did not vary across the lowest three quartiles, so these were combined for comparison against the highest quartile. Stenosis level was classified as either involving or not involving a given level. Stenosis location was classified as involving or not involving the central canal, lateral recess or neuroforamen.[24] Predominant pain location was classified as predominant leg pain (baseline leg pain bothersomeness score greater than or equal to back pain pain bothersomeness score) or predominant back pain (back pain bothersomeness score greater than leg pain bothersomeness score).

Statistical Considerations

The primary aim of the analysis was to identify baseline variables that were significant indicators of differential treatment effects of surgery (i.e. variables with significant subgroup by*treatment interactions). The treatment effect (TE) of surgery was defined as:

$$\text{Treatment Effect} = \text{Change in } ODI_{\text{surgery}} - \text{Change in } ODI_{\text{nonoperative}}$$

In all analyses, TE was calculated as the time-weighted average (area under the curve) over 4 years. Note that a negative TE indicated that surgery was more effective than nonoperative treatment. To identify candidate variables, a "minimally-adjusted" analysis (controlling only for age, gender, center, and baseline ODI score) was performed for each of the fifty-three

variables to identify those that had significant TE differences across subgroups. Those that were identified as potential TE modifiers (p<0.05) were then entered along with their treatment interaction terms into a longitudinal regression model that controlled for factors predicting treatment received or missing data in order to control for selection and attrition bias (center, age, gender, baseline ODI score, income, treatment preference, duration of symptoms, compensation status, smoking status, BMI, baseline Stenosis Bothersomeness Index, joint problems, stomach problems, and bowel problems).[4] Variables with significant treatment interaction terms (p<0.05) were identified as independent TE modifiers and included in the final multivariate mixed effects longitudinal regression model including a random individual effect to account for correlation between repeated measurements within individuals. The outcomes were stratified by modifiers and overall comparisons of area-under-the-curve across the four-year follow-up were made using a Wald test.[31] This analytical approach, starting with "minimally adjusted" analyses, was performed in order to ensure that no potential independent modifiers were missed in the first stage. In the second stage, the complete multivariate model was used in order to definitively identify independent TE modifiers.

RESULTS

Patient Distribution and Follow-Up

There were 654 patients enrolled overall, with 289 in the randomized trial and 365 in the observational cohort. By 4 years, 419 (64%) had undergone surgery. Completeness of follow-up ranged from 89% at 6 months to 67% at 4 years. Details can be found in the primary analysis.[4]

Minimally Adjusted Analyses

All examined subgroups including at least fifty patients had a significant TE except for smokers (Table 1, TE = −1.6, 95% CI −6.8 – 3.5). Some smaller subgroups (single patients, those on antidepressants, and those who described their race as non-white or non-black) did not have a significant TE, but this was likely due to the small number of patients in these groups and the large confidence intervals around their estimated TEs. The majority of variables were significantly associated with surgical and/or nonoperative outcomes. However, the direction of association was almost always the same for both treatments, so most variables were not significantly associated with TE (e.g. patients receiving worker's compensation did much worse than those not receiving compensation with both surgical and nonoperative treatment but had similar TE). Ten variables predicted a greater TE: baseline ODI 56 (TE −15.6 vs. −5.2 for baseline ODI > 56, p < 0.001), BMI 30 (TE −13.2 vs. −9.5 for BMI < 30, p = 0.035), being a non-smoker (TE −11.8 vs. −1.6 for smokers, p < 0.001), having diabetes (TE −14.8 vs. −10.2 for non-diabetics, p = 0.043), the presence of any neurological deficit (TE −13.5 vs. − 7.5 for no neurological deficit, p < 0.001), the presence of a sensory deficit (TE −14.9 vs. − 9.4 for no sensory deficit, p = 0.003), L5-S1 stenosis (TE −13.7 vs. −9.7 for no L5-S1 stenosis, p = 0.036), neuroforaminal stenosis (TE −14.4 vs. − 9.0 for no neuroforaminal stenosis, p = 0.003), predominant leg pain (TE −11.8 vs. −7.6 for predominant back pain, p = 0.039), and not lifting at work (TE −12.8 vs. −8.7 for lifting at work, p = 0.015).

Multivariate Analyses

The ten variables that were associated with TE were evaluated for collinearity, and, not surprisingly, having any neurological deficit and having a sensory deficit were strongly related (p < 0.0001) as were BMI > 30 and diabetes (p < 0.0001). As such, sensory deficit and BMI, the weaker TE predictors in the collinear variable pairs, were not included in the multivariate model. The other eight significant TE predictors identified in the minimally

adjusted analyses and their interaction terms were added to a previously published model evaluating the time-weighted average TE over four years of follow-up.[4] Baseline ODI 56, not smoking, neuroforaminal stenosis, predominant leg pain, and not lifting at work were significant independent predictors of greater TE (p < 0.05), while BMI > 30 and L5-S1 stenosis were not (Table 2, Figures 1).

To illustrate the range of TEs across subgroups defined by multiple variables, we compared outcomes between patients with characteristics associated with the greatest TE (ODI 56, non-smokers, neuroforaminal stenosis, neurological deficit, predominant leg pain, and not lifting at work) and smokers with baseline ODI > 56 (Table 3). Patients with the six characteristics associated with greater TE had a predicted TE of −24.1 compared to +3.6 for the smokers with baseline ODI > 56 (p < 0.001). We were unable to compare outcomes to the group with the lowest predicted TE (ODI > 56, smokers, no neuroforaminal stenosis, no neurological deficit, predominant back pain, and lifting at work) since no patients had this combination of characteristics.

DISCUSSION

While SPORT demonstrated a significant treatment effect of surgery for spinal stenosis at the group level, surgical decision making takes place at level of the individual patient.[4] As such, the current study evaluated fifty-three variables to determine which were significant TE predictors and could be used to guide treatment decisions. Somewhat surprisingly, essentially all subgroups other than smokers had a significant TE, indicating that surgery would likely be beneficial for most non-smokers who met the SPORT inclusion criteria (symptoms for at least 12 weeks, neurogenic claudication or radicular pain, and SpS on cross-sectional imaging). Prior studies have identified demographic, psychosocial, and clinical factors associated with surgical outcomes in SpS and used these findings to make recommendations about which patients would likely benefit most from surgery.[6,9,11,12,18,19] However, ignoring the effect of these variables on nonoperative outcomes does not allow for the estimation of the TE, and the current study suggests that most variables affect surgical and nonoperative outcomes similarly and thus do not predict TE.

Consistent with prior studies, the current investigation demonstrated that both patient and disease characteristics predicted surgical and nonoperative outcomes and TE. Lower baseline ODI score and not smoking were the strongest predictors of higher TE, while disease specific characteristics such as predominant leg pain, neuroforaminal stenosis, and neurological deficit were weaker but still significant independent predictors of greater TE. Interestingly, demographic characteristics and comorbidities tended to be strong predictors of surgical and nonoperative outcomes but were often not associated with TE. Similar to prior reports, the current study demonstrated better surgical outcomes associated with male gender, higher educational attainment, higher income, and better baseline mental component score (MCS).[10,12,18,19] These factors also tended to be associated with better nonoperative outcomes so were not significant TE predictors. Katz et al. reported that self-rated health and comorbidity burden were the strongest predictors of SpS surgical outcomes, and the current study also found that diabetes, hypertension, joint problems, and increasing comorbidity burden were associated with worse surgical and nonoperative outcomes.[10] However, only diabetes was a significant TE predictor in the minimally adjusted analyses, and diabetics actually had a higher TE due to the very poor nonoperative outcomes in this group. Diabetes did not remain a significant predictor in the multivariate model. These findings confirm the importance of considering nonoperative outcomes and TE in evaluating outcome predictors.

A baseline ODI score in the highest quartile (above 56) was associated with a greater degree of improvement with surgical and nonoperative treatment but a much smaller TE, primarily due to these patients improving significantly more with nonoperative treatment than patients in the other three quartiles. Given that patients in the highest quartile reported extremely high levels of disability at baseline (greater than the average patient with metastatic spine lesions), regression to the mean may explain a portion of their high level of improvement with both surgery and nonoperative treatment.[32] Cobo Soriano et al. also reported that higher baseline ODI scores were associated with greater improvement in ODI in SpS patients treated with decompression and fusion.[18] Contrary to the conventional wisdom that patients with the worst symptoms should be treated surgically, these findings suggest that nonoperative treatment is a reasonable option in patients with extremely high baseline ODI scores given their high level of improvement with nonoperative treatment.

The other factor most strongly related to TE was smoking, with smokers actually not benefiting significantly from surgery relative to nonoperative treatment. This finding mirrors a recently published study of the Swedish Spine Register that demonstrated smokers improved less on the ODI and most other outcome measures two years following surgery for SpS.[33] While it has been well-established that smoking is associated with a higher pseudarthrosis rate and worse outcomes in lumbar fusion, the vast majority of SpS patients in SPORT and the Swedish Spine Registry did not undergo fusion.[18,34] The data from SPORT and the Swedish Spine Register demonstrate that smoking is a strong predictor of worse surgical outcomes following decompression without fusion in SpS, and the lack of a significant TE for smokers suggests that smoking cessation prior to decompression—even without fusion—should be considered. It is possible that the lack of a significant outcomes difference in smokers treated surgically and nonoperatively may represent a Type II error due to the fact that only 62 smokers were included. However, the TE estimate for smokers was close to zero (−1.6), and the 95% confidence interval was not extremely broad (−6.8 to +3.5). Making definitive conclusions about the effect of smoking on TE in SpS would require a sufficiently powered randomized trial including only smokers.

Four other variables—neuroforaminal stenosis, predominant leg pain, not lifting at work, and neurological deficit—were also associated with greater TE, though not to the same magnitude as baseline ODI score or smoking. Neuroforaminal stenosis had no effect on surgical outcomes but predicted markedly worse nonoperative outcomes. Predominant leg pain—as compared to predominant back pain—was associated with significantly better surgical outcomes and had no effect on nonoperative outcomes. The current study supports multiple prior investigations reporting better surgical outcomes for those with predominant leg pain but also demonstrates that predominant back pain patients still benefit from surgery compared to nonoperative care.[9,11,16,18] Not lifting at work was also associated with a greater TE, primarily due to more improvement following surgery for those who do not lift at work. These results are different from those reported by Mariconda et al. who found worse nonoperative outcomes among those who lift at work.[12] The presence of a baseline neurological deficit was the only variable that modified outcomes in opposite directions for surgery and nonopeartive care, being associated with significantly better surgical outcomes and a trend towards worse nonoperative outcomes. No prior SpS study has evaluated the effect of neurological deficit on outcome, however, Abramowitz et al. actually found that a motor or sensory deficit predicted worse surgical outcomes for disk herniation patients.[35]

While neuroforaminal stenosis, predominant leg pain, not lifting at work, and neurological deficit predicted relatively modest differences in TE, comparing TE between patients with the six characteristics associated with greater TE and smokers with a baseline ODI in the top quartile was impressive: the predicted TEs were −24.1 versus +3.6. In other words, the high TE group improved 24 points more with surgery than with nonoperative treatment (this

group essentially did not improve with nonoperative treatment), while the low TE group actually improved marginally more with nonoperative treatment than with surgery. While these differences demonstrate the extremes of the TE spectrum, they highlight the importance of considering the effect of individual characteristics on outcomes when making treatment decisions.

There are a number of limitations of this study. While SPORT was designed with a randomized cohort, the high rate of protocol non-adherence precluded meaningful interpretation of those data on an intent-to-treat basis alone.[5] This phenomenon underscores the difficulty of conducting and analyzing a randomized trial of an elective surgical procedure that is primarily performed for pain relief. Additionally, the subgroup analyses used to evaluate potential TE modifiers were possibly underpowered for some variables (as indicated by broad confidence intervals around some TE estimates), and this may have resulted in the failure to detect some meaningful modifiers (Type II error). On the other hand, fifty-three potential TE modifiers were evaluated, which put the study at risk for finding spurious associations due to chance alone (Type I error). We intentionally performed the initial analyses with minimally adjusted models to generate a list of potential TE modifiers that could be tested in the complete multivariate model. As such, the findings of these minimally adjusted analyses were subject to confounding and could yield slightly different results from those reported previously using a multivariate model.[15,16] Finally, this study developed a model using group level data with the aim of counseling individual patients. As such, there is uncertainty in the estimates of the TE, and this is reflected in the confidence intervals around the estimates.

Do the results of this study give us any further guidance about who should undergo surgery and who should receive nonoperative care for SpS? The most striking finding was that smokers improved to a similar degree with surgery and nonoperative treatment, indicating that smoking cessation should be considered before surgery. However, this study did not demonstrate a causal relationship between smoking and lower TE, and smoking may simply be a marker for other characteristics responsible for the association. Future studies should evaluate the effect of smoking cessation on TE and determine if there is a duration of cessation necessary to observe a benefit. In addition, a prolonged period of nonoperative treatment should be considered in patients with extremely high baseline ODI scores given that they tended to improve significantly and almost to the same degree with nonoperative care as with surgery. While all large subgroups other than smokers had a significant TE, the TE was relatively small for some subgroups (i.e. patients with predominant low back pain and those with baseline ODI scores above 56). The minimal clinically important difference for the ODI has been estimated to be around 11 points, so some of these subgroups may have a TE that is statistically significant but not clinically important.[36] All patients should be guided through a shared decision making process that considers their individual characteristics in estimating their likely TE—not just their predicted surgical outcome.[37,38] In advising patients, clinicians should remember that the findings of this study apply only to patients who meet the inclusion criteria of SPORT: neurogenic claudication or radicular pain for at least 12 weeks and imaging demonstrating stenosis. Given that the treatment of SpS represents preference sensitive care, the appropriate treatment decision for patients who meet SPORT's inclusion criteria is that which the patient makes after evidence-based, individualized shared decision making. Future work will involve the creation and evaluation of real time computer models that can be used by individual patients with their providers in the clinical setting to predict their likely surgical and nonoperative outcomes.[39]

Acknowledgments

REFERENCES

1. Atlas SJ, Deyo RA, Keller RB, et al. The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. Spine. 1996; 21:1787–94. discussion 94-5. [PubMed: 8855463]

2. Atlas SJ, Keller RB, Robson D, Deyo RA, Singer DE. Surgical and nonsurgical management of lumbar spinal stenosis: four-year outcomes from the maine lumbar spine study. Spine. 2000; 25:556–62. [PubMed: 10749631]

3. Malmivaara A, Slatis P, Heliovaara M, et al. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine. 2007; 32:1–8. [PubMed: 17202885]

4. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis four-year results of the Spine Patient Outcomes Research Trial. Spine (Phila Pa 1976). 2010; 35:1329–38. [PubMed: 20453723]

5. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. The New England journal of medicine. 2008; 358:794–810. [PubMed: 18287602]

6. Airaksinen O, Herno A, Turunen V, Saari T, Suomlainen O. Surgical outcome of 438 patients treated surgically for lumbar spinal stenosis. Spine (Phila Pa 1976). 1997; 22:2278–82. [PubMed: 9346149]

7. Hurri H, Slatis P, Soini J, et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. Journal of spinal disorders. 1998; 11:110–5. [PubMed: 9588466]

8. Iguchi T, Kurihara A, Nakayama J, Sato K, Kurosaka M, Yamasaki K. Minimum 10-year outcome of decompressive laminectomy for degenerative lumbar spinal stenosis. Spine. 2000; 25:1754–9. [PubMed: 10888941]

9. Jonsson B, Annertz M, Sjoberg C, Stromqvist B. A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part II: Five-year follow-up by an independent observer. Spine. 1997; 22:2938–44. [PubMed: 9431630]

10. Katz JN, Stucki G, Lipson SJ, Fossel AH, Grobler LJ, Weinstein JN. Predictors of surgical outcome in degenerative lumbar spinal stenosis. Spine. 1999; 24:2229–33. [PubMed: 10562989]

11. Kleinstuck FS, Grob D, Lattig F, et al. The influence of preoperative back pain on the outcome of lumbar decompression surgery. Spine (Phila Pa 1976). 2009; 34:1198–203. [PubMed: 19407677]

12. Mariconda M, Zanforlino G, Celestino GA, Brancaleone S, Fava R, Milano C. Factors influencing the outcome of degenerative lumbar spinal stenosis. Journal of spinal disorders. 2000; 13:131–7. [PubMed: 10780688]

13. Spratt KF, Keller TS, Szpalski M, Vandeputte K, Gunzburg R. A predictive model for outcome after conservative decompression surgery for lumbar spinal stenosis. Eur Spine J. 2004; 13:14–21. [PubMed: 14658061]

14. Freedman MK, Hilibrand AS, Blood EA, et al. The impact of diabetes on the outcomes of surgical and nonsurgical treatment of patients in the spine patient outcomes research trial. Spine (Phila Pa 1976). 2011; 36:290–307. [PubMed: 21270715]

15. Park DK, An HS, Lurie JD, et al. Does multilevel lumbar stenosis lead to poorer outcomes?: a subanalysis of the Spine Patient Outcomes Research Trial (SPORT) lumbar stenosis study. Spine (Phila Pa 1976). 2010; 35:439–46. [PubMed: 20081560]

16. Pearson A, Blood E, Lurie J, et al. Predominant leg pain is associated with better surgical outcomes in degenerative spondylolisthesis and spinal stenosis: results from the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976). 2011; 36:219–29. [PubMed: 21124260]

17. 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. [PubMed: 20023603]

18. Cobo Soriano J, Sendino Revuelta M, Fabregate Fuente M, Cimarra Diaz I, Martinez Urena P, Deglane Meneses R. Predictors of outcome after decompressive lumbar surgery and instrumented posterolateral fusion. Eur Spine J. 2010; 19:1841–8. [PubMed: 20135333]

19. Sinikallio S, Aalto T, Airaksinen O, Herno A, Kroger H, Viinamaki H. Depressive burden in the preoperative and early recovery phase predicts poorer surgery outcome among lumbar spinal stenosis patients: a one-year prospective follow-up study. Spine (Phila Pa 1976). 2009; 34:2573–8. [PubMed: 19927107]

20. Pearson, A.; Lurie, J.; Tosteson, TD., et al. SPORT Intervertebral Disc Herniation: Indications Matter Most. North American Spine Society Annual Meeting; Orlando, FL. 2010.

21. Birkmeyer NJ, Weinstein JN, Tosteson AN, et al. Design of the Spine Patient outcomes Research Trial (SPORT). Spine. 2002; 27:1361–72. [PubMed: 12065987]

22. Tosteson, TD.; Hanscom, B.; Blood, EA.; Lurie, JD.; Tosteson, ANA.; Weinstein, JN. Statistical methods for cross-over in the SPORT lumbar disc herniation trial. International Society for the Study of the Lumbar Spine Annual Meeting; Hong Kong. 2007.

23. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. The New England journal of medicine. 2007; 356:2257–70. [PubMed: 17538085]

24. Fardon DF, Milette PC. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Spine. 2001; 26:E93–E113. [PubMed: 11242399]

25. Lurie JD, Tosteson AN, Tosteson TD, et al. Reliability of readings of magnetic resonance imaging features of lumbar spinal stenosis. Spine. 2008; 33:1605–10. [PubMed: 18552677]

26. Ware JE Jr. Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care. 1992; 30:473–83. [PubMed: 1593914]

27. Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American spine society lumbar spine outcome assessment Instrument: reliability and validity tests. Spine. 1996; 21:741–9. [PubMed: 8882698]

28. Atlas SJ, Deyo RA, Patrick DL, Convery K, Keller RB, Singer DE. The Quebec Task Force classification for Spinal Disorders and the severity, treatment, and outcomes of sciatica and lumbar spinal stenosis. Spine. 1996; 21:2885–92. [PubMed: 9112713]

29. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine. 1995; 20:1899–908. discussion 909. [PubMed: 8560339]

30. Walsh TL, Homa K, Hanscom B, Lurie J, Sepulveda MG, Abdu W. Screening for depressive symptoms in patients with chronic spinal pain using the SF-36 Health Survey. Spine J. 2006; 6:316–20. [PubMed: 16651227]

31. Fitzmaurice, G.; Laird, N.; Ware, J. Applied Longitudinal Analysis. John Wiley & Sons; Philadelphia, PA: 2004.

32. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine. 2000; 25:3115–24. [PubMed: 11124727]

33. Sanden B, Forsth P, Michaelsson K. SMOKERS SHOW LESS IMPROVEMENT THAN NON-SMOKERS 2 YEARS AFTER SURGERY FOR LUMBAR SPINAL STENOSIS: A study of 4555 Patients from the Swedish Spine Register. Spine (Phila Pa 1976). 2011

34. Andersen T, Christensen FB, Laursen M, Hoy K, Hansen ES, Bunger C. Smoking as a predictor of negative outcome in lumbar spinal fusion. Spine (Phila Pa 1976). 2001; 26:2623–8. [PubMed: 11725245]

35. Abramovitz JN, Neff SR. Lumbar disc surgery: results of the Prospective Lumbar Discectomy Study of the Joint Section on Disorders of the Spine and Peripheral Nerves of the American

Association of Neurological Surgeons and the Congress of Neurological Surgeons. Neurosurgery. 1991; 29:301–7. discussion 7-8. [PubMed: 1886676]

36. Weinstein JN. The missing piece: embracing shared decision making to reform health care. Spine (Phila Pa 1976). 2000; 25:1–4. [PubMed: 10647152]

37. Weinstein JN, Clay K, Morgan TS. Informed patient choice: patient-centered valuing of surgical risks and benefits. Health affairs (Project Hope). 2007; 26:726–30. [PubMed: 17485750]

38. Tosteson, TD.; Pearson, AM.; Lurie, J., et al. Improving predictions of outcomes for surgical and non-operative treatment of intervertebral disc herniation. International Society for the Study of the Lumbar Spine Annual Meeting; Auckland. 2010. 2010

KEY POINTS

- Other than smokers, all patient subgroups improved more with surgery than with nonoperative treatment.

- Baseline ODI score less than 56, not smoking, neuroforaminal stenosis, predominant leg pain, not lifting at work, and baseline neurological deficit predicted a greater treatment effect of surgery.

- Smoking cessation should be considered prior to surgery for SpS.

- These data can be used to help individualize shared decision making discussions about likely outcomes following surgical or nonoperative treatment for SpS.

A
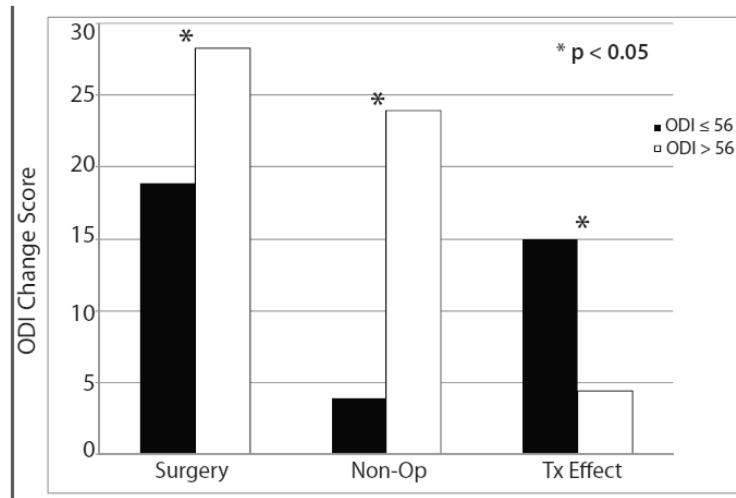
B
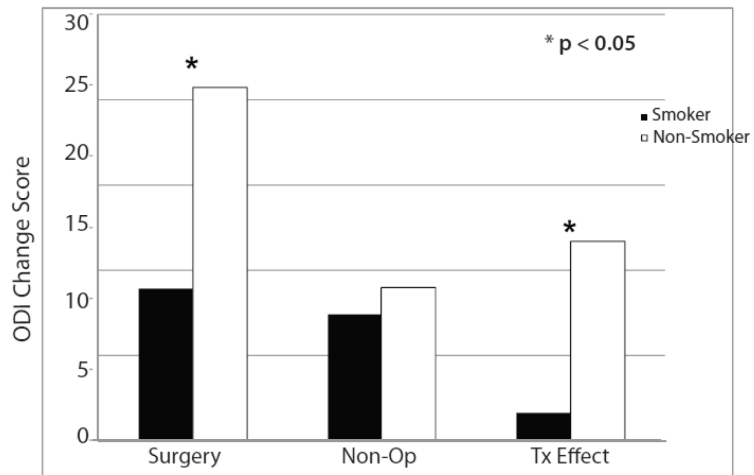
C

D

E

F

Figure 1.
These graphs compare surgical and nonoperative outcomes and treatment effect (TE) among the groups defined by the six independent TE predictors.

Table 1

Area under curve subgroup results for Oswestry Disability Index (ODI) change scores, adjusted for age, gender, center, and baseline ODI score[*] *(Time weighted average 4 years treatment effects).*

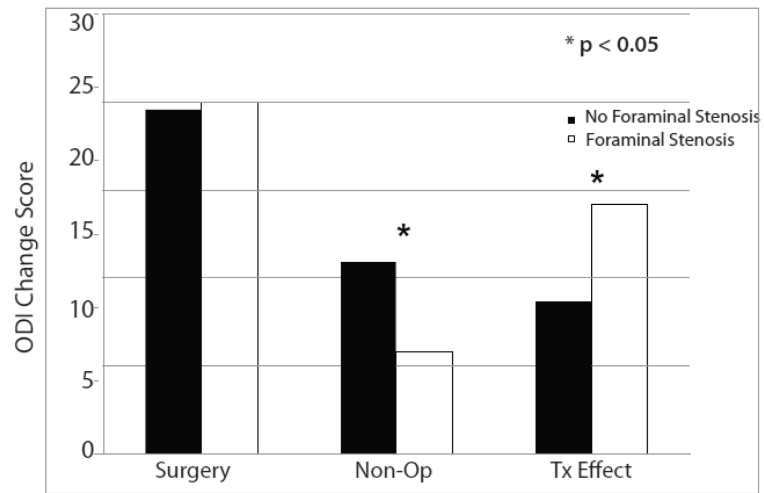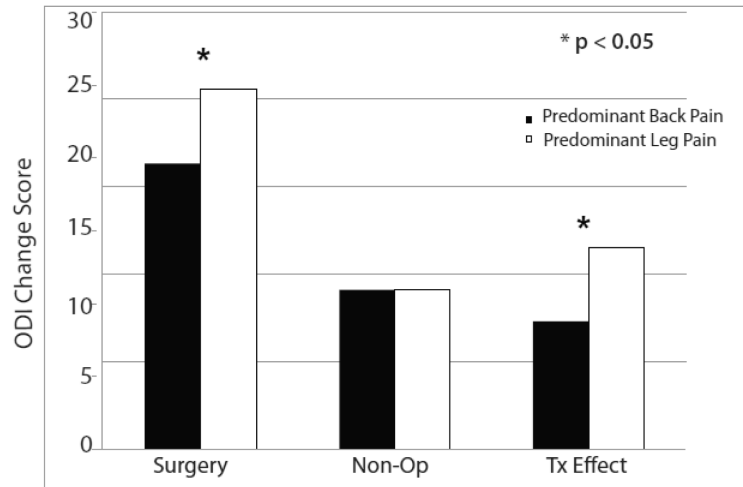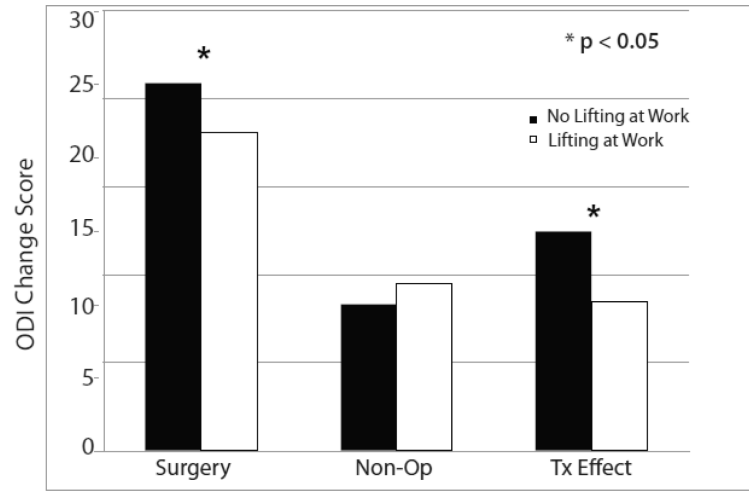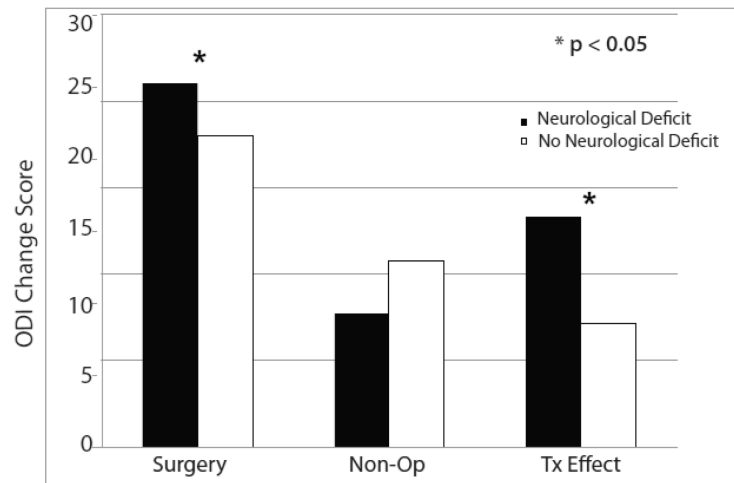| SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
Baseline ODI 56	482	−18.7 (0.8)	−3.1 (0.9)	−15.6 (−17.5, −13.7)
Baseline ODI > 56	150	−28.7 (1.5)	−23.5 (2.2)	−5.2 (−9.5, −0.9)
p-value		<0.001	<0.001	<0.001
Age 65	315	−20.7 (1)	−8.3 (1.2)	−12.4 (−14.9, −9.9)
Age > 65	319	−18.2 (1)	−8.7 (1.1)	−9.5 (−11.8, −7.2)
p-value		0.069	0.83	0.094
Female	249	−16.8 (1.2)	−6.8 (1.3)	−9.9 (−12.6, −7.2)
Male	385	−21 (0.9)	−9.5 (1.1)	−11.5 (−13.7, −9.4)
p-value		0.003	0.11	0.35
White	533	−20 (0.8)	−9.1 (0.9)	−10.9 (−12.8, −9)
Black	52	−19.1 (2.6)	−5.3 (2.7)	−13.7 (−19.4, −8)
Others	49	−13.3 (2.5)	−6.7 (3.9)	−6.6 (−14, 0.9)
p-value		0.039	0.35	0.32
High School or Less	223	−16.7 (1.2)	−4.3 (1.4)	−12.5 (−15.3, −9.7)
At Least Some College or More	401	−20.7 (0.9)	−11 (1)	−9.6 (−11.7, −7.5)
p-value		0.007	<0.001	0.10
Divorced/Widowed	144	−16.5 (1.5)	−5.3 (1.6)	−11.2 (−14.5, −7.8)
Married	446	−20.6 (0.8)	−9.7 (1)	−10.9 (−12.9, −8.8)
Single	38	−16.2 (3)	−9.2 (3.3)	−7 (−14.1, 0.2)
p-value		0.033	0.072	0.56
Not Working	379	−18.2 (1)	−8 (1.1)	−10.2 (−12.4, −8)
Income Over $50,000	129	−25.4 (1.6)	−13.3 (2.1)	−12.2 (−16, −8.3)
Income Under $50,000	122	−16.3 (1.6)	−5.9 (1.9)	−10.5 (−14.1, −6.8)
p-value		<0.001	0.02	0.68
Body Mass Index < 30	373	−20 (0.9)	−10.5 (1.1)	−9.5 (−11.6, −7.3)
Body Mass Index 30	261	−18.7 (1.1)	−5.5 (1.3)	−13.2 (−15.9, −10.4)
p-value		0.36	0.003	0.035
Stenosis Bothersomeness 15	347	−20.9 (1)	−9.9 (1)	−11 (−13.2, −8.8)
Stenosis Bothersomeness > 15	279	−17.7 (1)	−7 (1.4)	−10.7 (−13.5, −7.9)
p-value		0.025	0.10	0.87
Baseline MCS 35	89	−15.6 (1.8)	−8.3 (2.3)	−7.4 (−11.9, −2.8)
Baseline MCS > 35	545	−20.1 (0.8)	−8.7 (0.9)	−11.4 (−13.2, −9.5)
p-value		0.024	0.87	0.11

SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
Symptoms Less Than 6 Months	266	−21.8 (1.1)	−12.5 (1.2)	−9.2 (−11.8, −6.7)
Symptoms More Than 6 Months	368	−17.7 (0.9)	−5.5 (1.1)	−12.3 (−14.5, −10)
p-value		0.003	<0.001	0.076
Smoker	62	−8 (2.3)	−6.4 (2.4)	−1.6 (−6.8, 3.5)
Non-Smoker	566	−20.6 (0.7)	−8.8 (0.9)	−11.8 (−13.6, −10)
p-value		<0.001	0.34	<0.001
No Hypertension	339	−21.3 (0.9)	−10.2 (1.1)	−11.1 (−13.3, −8.8)
Hypertension	288	−17.3 (1.1)	−7.1 (1.2)	−10.2 (−12.7, −7.6)
p-value		0.004	0.053	0.60
No Diabetes	531	−20 (0.8)	−9.8 (0.9)	−10.2 (−12, −8.3)
Diabetes	96	−16.6 (1.8)	−1.8 (2)	−14.8 (−18.9, −10.7)
p-value		0.075	<0.001	0.043
No osteoporosis	567	−19.8 (0.7)	−8.7 (0.9)	−11.1 (−12.9, −9.3)
Osteoporosis	60	−16.5 (2.5)	−7.8 (2.5)	−8.7 (−14, −3.3)
p-value		0.20	0.74	0.40
No Heart Problem	462	−19.8 (0.8)	−10 (1)	−9.9 (−11.9, −7.9)
Heart Problem	165	−18.5 (1.4)	−5.2 (1.5)	−13.3 (−16.5, −10.1)
p-value		0.41	0.009	0.069
No Stomach Problem	488	−20.8 (0.8)	−9.5 (0.9)	−11.3 (−13.3, −9.4)
Stomach Problem	139	−14.7 (1.5)	−6 (1.7)	−8.7 (−12.3, −5.1)
p-value		<0.001	0.069	0.20
No Bowel or Intestinal Problem	541	−20.4 (0.7)	−9 (0.9)	−11.4 (−13.3, −9.5)
Bowel or Intestinal Problem	86	−14 (1.9)	−6.5 (2.1)	−7.5 (−11.9, −3.1)
p-value		0.002	0.28	0.10
No Depression	557	−19.5 (0.7)	−8.6 (0.9)	−11 (−12.8, −9.2)
Depression	70	−19.6 (2.1)	−9 (2.4)	−10.6 (−15.5, −5.6)
p-value		1	0.87	0.88
No Joint Problem	281	−23 (1)	−11.9 (1.2)	−11.1 (−13.7, −8.5)
Joint Problem	346	−16.7 (0.9)	−6.1 (1.1)	−10.7 (−12.9, −8.4)
p-value		<0.001	<0.001	0.79
No Other Comorbidities	414	−20 (0.9)	−9.5 (1)	−10.5 (−12.6, −8.4)
Other Comorbidities [**]	220	−18.3 (1.2)	−6.5 (1.4)	−11.8 (−14.7, −9)
p-value		0.24	0.073	0.44
Zero or One Comorbidity	217	−22.6 (1.1)	−12.9 (1.5)	−9.7 (−12.7, −6.7)
Two or Three Comorbidities	242	−20.9 (1.1)	−8.7 (1.2)	−12.2 (−14.9, −9.6)

SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
Four or More Comorbidities	168	−12.7 (1.4)	−3.8 (1.5)	−9 (−12.1, −5.8)
p-value		<0.001	<0.001	0.23
Neurogenic Claudication	508	−19.5 (0.8)	−8.3 (0.9)	−11.2 (−13.1, −9.3)
No Neurogenic Claudication	126	−19.1 (1.6)	−9.6 (1.8)	−9.5 (−13.3, −5.7)
p-value		0.82	0.53	0.42
Positive Tension Sign	132	−20 (1.6)	−7.5 (1.9)	−12.5 (−16.4, −8.6)
Negative Tension Sign	502	−19.3 (0.8)	−8.8 (0.9)	−10.5 (−12.4, −8.7)
p-value		0.72	0.53	0.37
No Dermatomal Pain Radiation	135	−16.4 (1.5)	−7.4 (1.8)	−9 (−12.5, −5.5)
Some Dermatomal Pain Radiation	499	−20.3 (0.8)	−8.8 (0.9)	−11.5 (−13.5, −9.6)
p-value		0.022	0.49	0.21
Any Neurological Deficit	349	−20.7 (1)	−7.2 (1.1)	−13.5 (−15.7, −11.2)
No Neurological Deficit	285	−17.9 (1.1)	−10.4 (1.3)	−7.5 (−10.1, −4.9)
p-value		0.055	0.062	<0.001
Reflex Asymmetry	168	−18.8 (1.4)	−6.2 (1.6)	−12.6 (−15.8, −9.4)
No Reflex Asymmetry	466	−19.7 (0.8)	−9.4 (1)	−10.3 (−12.3, −8.2)
p-value		0.61	0.086	0.21
Sensory Deficit	182	−19.9 (1.3)	−4.9 (1.6)	−14.9 (−18.2, −11.7)
Sensory Intact	452	−19.3 (0.8)	−9.9 (1)	−9.4 (−11.4, −7.4)
p-value		0.72	0.006	0.003
Motor Deficit	177	−21.2 (1.4)	−9.9 (1.5)	−11.3 (−14.5, −8.2)
Motor Intact	457	−18.7 (0.8)	−8 (1)	−10.8 (−12.8, −8.7)
p-value		0.12	0.31	0.76
No L2-L3 Stenosis	455	−19.7 (0.8)	−8.8 (1)	−10.9 (−12.9, −8.9)
L2-L3 Stenosis	179	−19 (1.3)	−7.6 (1.7)	−11.4 (−14.8, −8.1)
p-value		0.67	0.54	0.78
No L3-L4 Stenosis	214	−21.1 (1.3)	−9.6 (1.4)	−11.5 (−14.4, −8.6)
L3-L4 Stenosis	420	−18.6 (0.9)	−7.9 (1)	−10.7 (−12.8, −8.6)
p-value		0.10	0.32	0.66
No L4-L5 Stenosis	55	−20.2 (2.4)	−10 (2.8)	−10.2 (−16.4, −4)
L4-L5 Stenosis	579	−19.4 (0.7)	−8.4 (0.9)	−11 (−12.7, −9.2)
p-value		0.73	0.58	0.82
No L5-S1 Stenosis	461	−18.7 (0.8)	−9 (1)	−9.7 (−11.8, −7.7)
L5-S1 Stenosis	173	−21.4 (1.4)	−7.7 (1.5)	−13.7 (−16.8, −10.5)
p-value		0.091	0.48	0.036

SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
No Moderate/Severe Stenotic Levels	15	−28.7 (5.9)	−9.4 (4.4)	−19.4 (−31.4, −7.4)
One Moderate/Severe Stenotic Level	234	−19.8 (1.2)	−10.6 (1.4)	−9.2 (−12, −6.4)
Two Moderate/Severe Stenotic Levels	241	−18.5 (1.1)	−7.9 (1.3)	−10.5 (−13.1, −7.9)
Three+ Moderate/Severe Stenotic Levels	144	−20.1 (1.4)	−6.1 (1.8)	−14 (−17.8, −10.3)
p-value		0.32	0.24	0.11
No Central Stenosis	91	−20.2 (1.9)	−8.4 (2.2)	−11.8 (−16.6, −6.9)
Central Stenosis	543	−19.4 (0.8)	−8.4 (0.9)	−10.9 (−12.8, −9.1)
p-value		0.70	0.98	0.75
No Lateral Recess Stenosis	131	−21.7 (1.6)	−7.8 (1.7)	−13.9 (−17.4, −10.4)
Lateral Recess Stenosis	503	−18.9 (0.8)	−8.8 (0.9)	−10.2 (−12.1, −8.2)
p-value		0.11	0.63	0.064
No Neuroforaminal Stenosis	427	−19.6 (0.9)	−10.6 (1)	−9 (−11.1, −6.9)
Neuroforaminal Stenosis	207	−19.5 (1.3)	−5.1 (1.4)	−14.4 (−17.2, −11.5)
p-value		0.96	0.002	0.003
Mild/Moderate Stenosis Severity	297	−20.2 (1.1)	−10.1 (1.1)	−10.1 (−12.5, −7.7)
Severe Stenosis Severity	337	−18.8 (0.9)	−6.9 (1.2)	−11.9 (−14.3, −9.5)
p-value		0.32	0.047	0.28
No Opioid Use	355	−19.2 (0.9)	−9.1 (1.1)	−10.1 (−12.3, −7.9)
Opioid Use	172	−18.7 (1.3)	−5.6 (1.8)	−13.1 (−16.8, −9.5)
p-value		0.76	0.099	0.17
No Prior Injections	285	−20 (1.1)	−8.6 (1.2)	−11.5 (−13.9, −9)
Prior Injections	349	−19 (0.9)	−8.5 (1.1)	−10.5 (−12.8, −8.2)
p-value		0.45	0.95	0.56
No Antidepressant Use	506	−19.4 (0.8)	−8.4 (0.9)	−11 (−13, −9)
Taking Antidepressants	21	−11.5 (4.1)	−3.4 (4.3)	−8.1 (−17.3, 1)
p-value		0.062	0.26	0.55
No NSAID Use	350	−18.9 (0.9)	−7 (1.1)	−11.8 (−14.1, −9.5)
Taking NSAID	177	−19.6 (1.3)	−10.6 (1.6)	−9 (−12.4, −5.7)
p-value		0.65	0.067	0.17
No Physical Therapy	187	−19.7 (1.3)	−9.8 (1.5)	−9.9 (−13, −6.7)
Had Physical Therapy	447	−19.3 (0.8)	−8 (1)	−11.3 (−13.4, −9.3)
p-value		0.83	0.32	0.43
Worker's Compensation	48	−11.4 (2.7)	0.3 (3.1)	−11.7 (−18.4, −5)

SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
No Worker's Compensation	583	−20.1 (0.7)	−9.5 (0.8)	−10.6 (−12.4, −8.9)
p-value		0.001	0.002	0.77
No Litigation	608	−19.9 (0.7)	−9.1 (0.8)	−10.8 (−12.5, −9)
Litigation Pending/Resolved	24	−8.9 (3.6)	3.3 (4)	−12.2 (−20.5, −4)
p-value		0.003	0.002	0.74
Dissatisfied with Symptoms	573	−19.3 (0.7)	−8.2 (0.9)	−11 (−12.9, −9.2)
Satisfied with Symptoms or Neutral	61	−19.6 (2.8)	−11.4 (2.2)	−8.2 (−13.7, −2.7)
p-value		0.91	0.18	0.33
Predominant Back Pain	159	−16.1 (1.4)	−8.5 (1.6)	−7.6 (−11.2, −4.1)
Predominant Leg Pain	456	−20.7 (0.8)	−8.8 (1)	−11.8 (−13.8, −9.8)
p-value		0.005	0.84	0.039
Problem Getting Better	46	−33 (3.4)	−16.6 (2.5)	−16.4 (−23.3, −9.5)
Problem Getting Worse	378	−17.5 (0.9)	−6.9 (1.1)	−10.6 (−12.8, −8.3)
Problem Staying about the Same	203	−21.8 (1.3)	−9.4 (1.4)	−12.4 (−15.3, −9.5)
p-value		<0.001	0.002	0.22
Treatment Preference Not Sure	121	−17.5 (1.7)	−6.3 (1.7)	−11.2 (−14.7, −7.6)
Prefer Nonoperative Treatment	229	−21.8 (1.4)	−10.3 (1.1)	−11.5 (−14.1, −8.8)
Prefer Surgery	283	−18.6 (1)	−8.7 (2.2)	−9.9 (−14.1, −5.7)
p-value		0.087	0.15	0.82
No Missed Work	148	−21.6 (1.4)	−9.5 (1.8)	−12.1 (−15.5, −8.8)
Missed Work	139	−19.2 (1.5)	−8.3 (1.9)	−10.9 (−14.7, −7.1)
p-value		0.26	0.66	0.63
No Lifting at Work	324	−21.4 (1)	−8.6 (1.1)	−12.8 (−15.2, −10.5)
Lifting at Work	307	−17.1 (1)	−8.4 (1.2)	−8.7 (−11.2, −6.3)
p-value		0.002	0.92	0.015
Small Chance of Being Pain Free with Surgery	48	−16.9 (3.1)	−8.5 (2.5)	−8.4 (−14.6, −2.1)
Moderate Chance of Being Pain Free with Surgery	207	−16.8 (1.3)	−7.6 (1.2)	−9.2 (−11.9, −6.5)
Big Chance of Being Pain Free with Surgery	366	−20.9 (0.9)	−9.4 (1.2)	−11.5 (−13.9, −9)
p-value		0.022	0.59	0.38
Small Chance of Being Pain Free with Nonoperative Treatment	344	−18.7 (0.9)	−9.7 (1.3)	−9 (−11.6, −6.4)
Moderate Chance of Being Pain Free with Nonoperative Treatment	187	−18.4 (1.4)	−6.3 (1.3)	−12.1 (−14.9, −9.2)

SpS	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI)[†]
Big Chance of Being Pain Free with Nonoperative Treatment	80	−27.7 (2.4)	−14.3 (2)	−13.4 (−18.4, −8.4)
p-value		0.002	0.003	0.16

[†]Treatment effect is the difference between the surgical and non-operative mean change from baseline. Analysis is done using a mixed effects longitudinal regression model including a random individual effect to account for correlation between repeated measurements within individuals. Treatment is a time-varying covariate where a patients' experience prior to surgery is attributed to the non-operative arm and time is measured from enrollment and his/her post-surgery outcomes are attributed to the surgical arm and time is measured from time of surgery.

[*]For baseline ODI subgroup analysis, the categorical baseline ODI (not the continuous ODI) was in the model.

[**]Other comorbidities include: stroke, cancer, fibromyalgia, chronic fatigue syndrome (CFS), post-traumatic stress (PTSD), alcohol, drug dependency, lung, liver, kidney, blood vessel, nervous system, migraine or anxiety.

[***]Insurance coverage was not included because only 8 out of 634 patients did not have insurance coverage.

Table 2

Area under curve subgroup results from the adjusted[*] as-treated Oswestry Disability Index (ODI) outcome analysis for the six significant treatment effect predictors (time weighted average 4 year treatment effects).

SPS Predictors	N	Surgical (n = 413)	Non-operative (n = 221)	Treatment Effect (95% CI) [†]
Baseline ODI 56	482	−18.9 (0.8)	−3.9 (0.9)	−15 (−17, −13)
Baseline ODI > 56	150	−28.3 (1.5)	−23.9 (2.2)	−4.4 (−8.7, −0.1)
p-value		<0.001	<0.001	<0.001
Smoker	62	−8.9 (2.2)	−7.4 (2.4)	−1.6 (−6.6, 3.5)
Non-Smoker	566	−20.7 (0.7)	−9 (0.9)	−11.7 (−13.6, −9.9)
p-value		<0.001	0.53	<0.001
No Neuroforaminal Stenosis	427	−19.6 (0.8)	−10.9 (1)	−8.7 (−10.8, −6.5)
Neuroforaminal Stenosis	207	−20 (1.2)	−5.8 (1.3)	−14.2 (−17, −11.3)
p-value		0.80	0.003	0.002
Predominant Back Pain	159	−16.3 (1.4)	−9.1 (1.6)	−7.3 (−10.8, −3.7)
Predominant Leg Pain	456	−20.6 (0.8)	−9.1 (0.9)	−11.5 (−13.5, −9.5)
p-value		0.008	0.97	0.035
No Lifting at Work	324	−20.9 (0.9)	−8.3 (1.1)	−12.5 (−14.9, −10.1)
Lifting at Work	307	−18.1 (1)	−9.5 (1.1)	−8.5 (−11, −6)
p-value		0.039	0.46	0.017
Any Neurological Deficit	349	−21 (0.9)	−7.7 (1.1)	−13.3 (−15.5, −11)
No Neurological Deficit	285	−18 (1)	−10.8 (1.2)	−7.2 (−9.8, −4.6)
p-value		0.037	0.058	<0.001

[*] Adjusted for center, age, gender, baseline ODI score, income, treatment preference, duration of symptoms, compensation, smoking status, BMI, baseline Stenosis Bothersomeness, joint, stomach and bowel problems.

[†] Treatment effect is the difference between the surgical and non-operative mean change from baseline. Analysis is done using a mixed model with a random subject intercept term. Treatment is a time-varying covariate where a patients' experience prior to surgery is attributed to the non-operative arm and time is measured from enrollment and his/her post-surgery outcomes are attributed to the surgical arm and time is measured from time of surgery.

Table 3

Area under curve subgroup results from the adjusted[*] as-treated Oswestry Disability Index (ODI) outcome analysis for combinations of the six significant treatment effect predictors (time weighted average 4 year treatment effects).

	N	Surgical	Non-operative	Treatment Effect[†] (95% CI)
Non-smokers with ODI 56, foraminal stenosis, neurological deficit, predominant leg pain, and no lifting at work	35	−23.8 (2.7)	0.3 (3.1)	−24.1 (−30.2, −18)
Smokers with ODI > 56	23	−0.9 (3.8)	−4.5 (5.2)	3.6 (−5.6, 12.8)
Others	576	−20.2 (0.7)	−9.9 (0.8)	−10.2 (−12.1, −8.4)
p-value		<0.001	0.003	<0.001

[*] Adjusted for center, age, gender, baseline ODI score, income, treatment preference, duration of symptoms, compensation, smoking status, BMI, baseline Stenosis Bothersomeness, joint, stomach and bowel problems.

[†] Treatment effect is the difference between the surgical and non-operative mean change from baseline. Analysis is done using a mixed model with a random subject intercept term. Treatment is a time-varying covariate where a patients' experience prior to surgery is attributed to the non-operative arm and time is measured from enrollment and his/her post-surgery outcomes are attributed to the surgical arm and time is measured from time of surgery.