



Published in final edited form as:

Spine (Phila Pa 1976). 2013 April 15; 38(8): 678–691. doi:10.1097/BRS.0b013e318278e571.

Outcome Variation Across Centers After Surgery for Lumbar Stenosis and Degenerative Spondylolisthesis:

The SPORT Experience

Atman Desai, MD^{*}, Kimon Bekelis, MD^{*}, Perry A. Ball, MD^{*}, Jon Lurie, MD^{†,‡,§}, Sohail K. Mirza, MD[‡], Tor D. Tosteson, DSc[†], Wenyan Zhao, MS[§], and James N. Weinstein, DO^{‡,§}

^{*}Section of Neurosurgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH

[†]Department of Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH

[‡]Department of Orthopaedics, Dartmouth-Hitchcock Medical Center, Lebanon, NH

[§]The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, NH

Abstract

Study Design—Retrospective review of a prospectively collected database.

Objective—To examine whether short- and long-term outcomes after surgery for lumbar stenosis (SPS) and degenerative spondylolisthesis (DS) vary across centers.

Summary of Background Data—Surgery has been shown to be of benefit for both SPS and DS. For both conditions, surgery often consists of laminectomy with or without fusion. Potential differences in outcomes of these overlapping procedures across various surgical centers have not yet been investigated.

Methods—Spine patient outcomes research trial cohort participants with a confirmed diagnosis of SPS or DS undergoing surgery were followed from baseline at 6 weeks, 3, 6, and 12 months, and yearly thereafter, at 13 spine clinics in 11 US states. Baseline characteristics and short- and long-term outcomes were analyzed.

Results—A total of 793 patients underwent surgery. Significant differences were found between centers with regard to patient race, body mass index, treatment preference, neurological deficit, stenosis location, severity, and number of stenotic levels. Significant differences were also found in operative duration and blood loss, the incidence of durotomy, the length of hospital stay, and wound infection. When baseline differences were adjusted for, significant differences were still seen between centers in changes in patient functional outcome (SF-36 bodily pain and physical function, and Oswestry Disability Index) at 1 year after surgery. In addition, the cumulative

Address correspondence and reprint requests to Atman Desai, M.D., Dartmouth-Hitchcock Medical Center, Section of Neurosurgery, One Medical Center Dr., Lebanon NH 03756; atman.desai@hitchcock.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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

adjusted change in the Oswestry Disability Index Score at 4 years significantly differed among centers, with SF-36 scores trending toward significance.

Conclusion—There is a broad and statistically significant variation in short- and long-term outcomes after surgery for SPS and DS across various academic centers, when statistically significant baseline differences are adjusted for. The findings suggest that the choice of center affects outcome after these procedures, although further studies are required to investigate which center characteristics are most important.

In this retrospective study, outcome variation after surgery for spinal stenosis and degenerative spondylolisthesis among centers participating in the spine patient outcomes research trial (SPORT) were analyzed. Significant variation, including the presence of outlier centers, was found. This suggests that the choice of center influences the outcome after these procedures, although further studies are required to ascertain which center characteristics may be most important.

Keywords

clinical outcomes; lumbar spine; spondylolisthesis; lumbar stenosis; laminectomy; lumbar fusion

Surgery for lumbar spinal stenosis (SPS) or degenerative spondylolisthesis (DS) is one of the most common procedures performed by spine surgeons in the United States.^{1–3} The rates of this procedure have been increasing in the US Medicare population, although there remains broad variation in rates of surgery across the country.^{4–6} The goal of the procedure is usually to decompress the thecal sac from either degenerative bony overgrowth or spinal canal compromise due to instability from listhesis of vertebral bodies.⁷ Typical presentation from SPS and DS includes back, buttock, and leg pain, often seen during walking and severely limiting walking tolerance.⁷ Symptoms are classically relieved by lumbar flexion. Most patients without related cauda equina syndrome or progressive weakness are initially managed conservatively.⁸ If symptoms persist or worsen, then a significant number will undergo surgical intervention with overall good postoperative outcomes.^{8,9} A certain proportion of these patients will, however, remain symptomatic and then may require further operative management. Poor outcomes after surgery have been attributed to several factors, including misdiagnosis, surgical technique failure, new pathology, untreated instability, facet syndrome, and adjacent level disease.^{10,11} The literature examining these factors is largely based on single-institution studies, therefore, the effect on outcomes of different centers has not been studied.

The spine patient outcomes research trial (SPORT),^{8,9,12,13} a large multicenter trial including both randomized and observational cohorts initiated in March 2000, provides standardized outcome measures and long-term follow-up and a valuable opportunity to examine short-term and long-term outcomes across various study centers.^{12,13} Furthermore, all laminectomies and fusions were performed using the same approach (open, posterior) and all were performed in patients without a previous history of lumbar spine surgery.

This study was undertaken to investigate whether short- and long-term outcomes after surgery for SPS and DS varied significantly among the study centers participating in the SPORT.

MATERIALS AND METHODS

Study Design

The SPORT was conducted at 13 medical centers with multidisciplinary spine practices in 11 states in the United States. Institutional review board approval was obtained at each center. The trial was registered with ClinicalTrials.gov (NCT00000409, NCT00000411). Additional background information is available in literature.^{12,13}

Patient Population

As part of the original SPORT protocol, all patients had neurogenic claudication or radicular leg pain with associated neurological signs, spinal stenosis, or degenerative spondylolisthesis seen on cross-sectional imaging, symptoms that had persisted for at least 6 weeks, and physician confirmation that they were surgical candidates. Pre-enrollment nonoperative care included physical therapy, anti-inflammatory medications, opioid analgesics, chiropractic care, and epidural injections. Enrollment began in March 2000 and ended in February 2005.

Surgery Performed

The protocol surgery consisted of a standard open posterior lumbar laminectomy at the affected level or levels with or without fusion. The use of a microscope was at the discretion of the surgeon but was not recorded as part of the SPORT database.

Study Measures

The short-term outcome measures were operative duration, operative blood loss, requirement for blood transfusion, perioperative nerve root injury, wound complications (e.g., infection), inpatient length of stay, and postoperative mortality.

The long-term outcome measures were the need for repeat surgery at 1, 2, 3, and 4 years, the Short Form-36 (SF-36) bodily pain and physical function component scores and the American Academy of Orthopaedic Surgeons MODEMS (Musculoskeletal Outcomes Data Evaluation and Management System) version of the Oswestry Disability Index (ODI) measured at 6 weeks, 3 months, 6 months, and yearly up to 4 years. SF-36 scores range from 0 to 100 points, with higher scores indicating less severe symptoms; the ODI ranges from 0 to 100 points, with lower scores indicating less severe symptoms. The effect of the center variability on long-term outcome was defined as the difference in the mean changes, compared with baseline, between the various groups (the difference in the difference).

Statistical Methods

The baseline characteristics of patients of each center were compared. The analyses consisted of comparisons between all the groups. Computations were performed with the use of the PROC MIXED procedure for continuous data and the PROC GENMOD procedure

for binary and non-normal secondary outcomes from the SAS software package (version 9.1; SAS Institute, Cary, NC). Significance was defined as $P < 0.05$ on the basis of a 2-sided hypothesis test with no adjustments made for multiple comparisons. The data for these analyses were collected through August 29, 2009. The outcome analyses for SF-36 and ODI scores were adjusted for age, sex, diagnosis, baseline scores for SF-36 and ODI, and those baseline variables that were significantly different among centers and also those that had significant effect on outcomes.

RESULTS

Baseline Characteristics

A total of 793 patients underwent standard open lumbar laminectomy and / or posterior fusion (Table 1). The enrollment in various centers was in the range of 29 to 125 patients. There were no significant differences among the study centers in some patient characteristics (age, sex, the prevalence of smoking, diabetes), baseline levels of disability (preoperative SF-36 bodily pain and physical functioning, ODI), or listhesis level. Patients demonstrated significant differences in race, body mass index (BMI), incidence of hypertension, joint problems, treatment preferences, neurological deficit, and stenosis level, location, and severity. Certain baseline characteristics did not reach overall statistical significance in terms of difference, but did seem to be in a wide range among centers. These included compensation (3%–19%), smoking (3%–12%), and certain comorbidities.

Follow-up

As of August 2009, the mean (SD) follow-up among all analyzed SPS/DS patients was 45.2 (13.8) months. The median (range) follow-up time among all analyzed SPS/DS patients was 47.4 (1.2–95.6) months.

Operative Events

There was a significant difference among the study centers in the type of procedure performed (laminectomy-only, laminectomy with instrumented or noninstrumented fusion), the stenotic level decompressed, and the number of stenotic levels decompressed (Table 2). The various centers demonstrated statistically significant differences with broad ranges in the operative duration (74–235.9 min), blood loss (174.5–892.5 mL), the need for intraoperative (4%–36%) and postoperative blood transfusions (3%–32%), and the incidence of dural tears (0%–18%). There were no differences in the incidence of intraoperative nerve root injury.

Short-term Outcomes

There were statistically significant differences in the length of hospital stay (1.3–4.1 d) and in the rate of wound infection (0%–10%) among various study centers (Table 2). There were no statistically significant overall differences in the incidence of postoperative nerve root injury, or wound hematoma. There were no occurrences of wound dehiscence, bone graft complication, cerebrospinal fluid fistula formation, paralysis or cauda equina injury, or other complications attributable to surgery. There was no mortality within the first 3

months of surgery, except at one center where 2 patients died within 6 weeks of lumbar laminectomy.

Adjusting Covariables

The set of adjusting covariates for the analysis of SF-36 scores and ODI scores were selected by putting baseline variables that were significantly different among centers into the outcome models. These were sex, diagnosis, baseline scores for (SF-36 and ODI), BMI, race, education, pseudoclaudication, straight-leg raising or femoral tension, pain radiation, any neurological deficit, reflex deficit, sensory deficit, motor deficit, hypertension, stomach problem, joint problem, and other comorbidities, L2–L3 level stenosis, L3–L4 level stenosis, L5–S1 level stenosis, central stenosis, lateral recess stenosis, neuroforamen stenosis, stenosis severity, number of moderate/severe stenotic levels, and treatment preference at baseline. Patient age and those baseline variables that had a significant effect on outcomes were selected as adjusting covariates. The set of adjusting covariates were age, sex, race, BMI, diagnosis, education, any neurological deficit, stomach problem, joint problem, other comorbidities, baseline treatment preference, and baseline scores for SF-36 and ODI.

Long-term Outcomes

One-, 2-, 3- and 4-year postsurgical reoperation rates were calculated from Kaplan-Meier plots (Tables 2–4; Figures 1, 2). There was no difference in the incidence of reoperation among the study centers at 1, 2, 3, and 4 years. (Figure 2, Table 2). Despite not reaching statistical significance, however, the variation in the rates of reoperation at 4 years was relatively broad (5%–21%).

There were significant differences across the centers in all long-term functional outcomes (SF-36 and ODI) at 1-year of follow-up, with a broad range (SF-36 BP: 22–37.7; SF-36 physical function: 19.3–36.3; ODI: 16.9–27.9) and in the ODI scores at 4 years (12.3–28.6) (Table 3, Figure 1). When the “area under the curve” was analyzed for each adjusted functional outcome measure, the cumulative difference in ODI scores showed broad and significant variation across centers (15–28.1), whereas there was also a trend toward significant differences in the SF-36 scores across centers (Table 4).

DISCUSSION

SPS refers to the compression of the lumbar neural elements by the intervertebral disc, hypertrophied facet joints, and ligamentum flavum. Symptoms not relieved through nonoperative management may be treated by the removal of offending bone and ligament (decompressive laminectomy). When there is instability and listhesis of one vertebral body on another, laminectomy and fusion (with or without instrumentation) is a common surgical treatment.⁷ Symptomatic patients have a significant degree of disability and several studies have demonstrated improvement in pain and function with operative interventions.^{8,9,14–17}

In a meta-analysis of 19 trials, surgery for SPS and DS has demonstrated a favorable risk-benefit profile.¹⁸ Amundsen et al¹⁹ demonstrated good results in 91% of patients undergoing surgery versus 71% in patients undergoing nonoperative management. A Cochrane database review²⁰ in 2005 noted that most published articles consisted of

uncontrolled case series or prospective cohort studies, and the few randomized trials that have been reported until then were performed on a small number of patients and reported mixed surgical indications.²¹ More recently, the effectiveness of operative intervention has also been supported by SPORT,^{8,9} a large multicenter randomized study that demonstrated equivalence of surgery to conservative management in the intention-to-treat analysis and superiority of surgery in the as-treated analysis. Although there is wide variation in the reported outcomes of surgery for SPS and DS, most studies reveal good or excellent outcomes in approximately 50% to 70% of patients,^{14,15,17} despite there being a need for further well-designed trials with a more systematic conservative care arm.

The broader surgical literature has investigated the potential correlation of various surgical outcomes to different centers' characteristics.^{22,23} Several studies have suggested that center characteristics, in particular volume, are responsible for varied outcomes among them.^{24–28} The strongest correlation has been observed in treatment of abdominal aortic aneurysm, congenital heart disease, pediatric renal transplantation, pancreatic and esophageal cancer, and heart and liver transplant.^{22,29–32} On the contrary, a recent study³³ from the Veterans Affairs National Surgical Quality Improvement Database did not demonstrate any center variation in outcomes for 8 common surgical procedures, including carotid endarterectomies, hip replacements, and cholecystectomies, but no spine operations. It, therefore, seems likely that the correlation of surgical outcomes with center characteristics is dependent on the procedure type and is not universal. Such research has, however, not been performed in spine surgery. Furthermore, in these studies, the investigators have tended to analyze 30-day mortality data and not functional objective outcome indices, which are more pertinent to spine surgery.

In this study, we observed statistically significant differences with broad variation across centers in several short-term and long-term outcomes after surgery in patients with SPS or DS. Rates of surgical site infection and dural tear varied across centers, as did operative blood loss and inpatient length of stay. In addition, all of the adjusted changes in functional indices (SF-36 bodily pain and physical function, ODI) were significantly different across centers at 1 year from surgery, and the ODI at 4 years. The cumulative change in ODI during 4 years (area under the curve) also showed a significant difference between centers, with the other 2 functional outcome indices (SF-36 scores) trending toward significance. These results are in accordance with this literature, which has suggested that there is considerable variation in the proportion of individuals reporting long-term benefit from surgical treatment of SPS and DS.^{11,17,19,34–36} Of note, a meta-analysis demonstrated considerable variation in reported success rates among several single-institution studies.³⁷

Several studies have investigated the outcomes after surgery for SPS and DS, and they have attributed its variable success to multiple factors.^{10,11,15,16,35} Factors such as age, sex, smoking, low socioeconomic status, working or disability status, regular exercise, and disease level and type have been investigated.^{10,11,15,35} For most of these factors, the data remain inconclusive, with one meta-analysis unable to identify any predictive factors of good outcome across the 74 selected studies.³⁷ In our study, the patient population showed significant differences with respect to baseline characteristics across centers, including sex, race, BMI, diagnosis, the presence of pseudoclaudication, straight-leg raise or femoral

tension, pain radiation, neurological deficits, certain comorbidities, treatment preference, and stenosis location, severity, and levels. The factors that had a significant effect on outcome were then selected as adjusting covariates to account for potential confounders and attempt to isolate the effect of center alone.

These results demonstrate that the first-time open surgery for SPS or DS, performed by neurosurgeons or orthopedic surgeons, provides the patients with variable short- and certain long-term functional outcomes across different centers, when baseline differences in patient characteristics are adjusted for. The etiology of this observation is uncertain and cannot be answered by these data but may be multifactorial. Individual surgeon experience and volume, center volume and facilities, and nuances in operative technique (e.g., microscope use) may all play a role in affecting outcome. Furthermore, the variability seen in this study in the type of surgery performed across centers (Table 2) may, in addition to being a product of variability in baseline diagnosis (Table 1), reflect individual surgeon preference for certain procedures, which may itself be a factor in outcome. Several other factors potentially affecting outcomes in spine surgery have been investigated. These have often had conflicting findings, possibly because much of this literature is based on retrospective single-center studies. These factors are, however, not available in the data investigated by the SPORT trial, therefore cannot be further analyzed using data from this multicenter prospectively collected cohort. It is also possible that some yet unrecognized pathophysiological features of SPS and DS (e.g., genetic) may influence surgical outcome in a way that cannot currently be appreciated, an interesting topic for further investigation. Variations in treatment effects related to socioeconomic factors may also be at play. Although workers compensation and secondary gain are not generally considered in these populations (SPS and DS), neurobiological responses to degenerative conditions such as these remain to be studied. There is clear evidence that pain response can be associated with psychosocial factors, economics, and education. SPORT data may allow for further explanation of the variation in treatments as we continue to unravel the mystery of why one patient does better than another when all else seems equal.

This study has several limitations that should be considered. It includes subgroup analysis of 13 different centers; therefore, although the study population of 793 is large, each center has a significantly smaller population, making the detection of statistically significant differences among groups difficult. To maximize the size of the study population, we analyzed patients with 2 overlapping conditions. It is retrospective and vulnerable to certain confounding variables. Although the patients were selected from major academic centers, given that the initial SPORT guidelines were relatively broad, there were potentially major differences in patient selection criteria and surgical techniques, and these could possibly confound outcomes. It should be noted, however, that all surgical cases in the SPORT were performed using an “open” technique, that is, without the use of a tubular retractor. A further potential limitation is that, the statistical analysis only detects group differences across the multiple centers, but does not necessarily detect “outliers.” In addition, baseline characteristics not having a significant effect on outcome despite a wide variability were not adjusted for in the analysis of long-term center effect. Compensation in particular has previously been reported to have a significant effect on outcome but did not reach statistical significance in this study. We think, however, that adjusting only for those variables with a

statistically significant effect on outcome in our own data is the most valid analysis to perform. As discussed earlier, although this study suggests that certain centers have better outcomes, it cannot answer which center characteristics, such as volume, facilities, surgeon experience and training, and operative technique, may play an important role in outcome. Further studies are clearly required to address this issue. Finally, patient population sizes at certain centers were relatively small, limiting statistical power in detecting significant differences across centers in certain outcomes, for example, functional outcome indices during the 4-year period.

The data presented in this study demonstrate clear heterogeneity and variation in outcomes across centers. This has not been demonstrated previously for spinal surgery and is not necessarily an intuitive finding, given that these surgical procedures are frequently performed in a range of centers across the United States, and that several surgical procedures,³³ for example, hip arthroplasty, have been shown to have little variation among centers. For this reason, we think that the data are of value to the reader, although we agree that it cannot answer which intrinsic characteristics of the center are responsible for this variation. We hope that our study may provide the impetus for further investigation in this area.

CONCLUSION

The effect of individual center on short- and long-term outcomes after surgery for SPS and degenerative spondylolisthesis has not been previously investigated. In this study, the data of the SPORT trial demonstrated statistically significant, broad variation in several short-term and long-term outcomes across centers, even when significant differences in baseline characteristics were adjusted for. The presence of “outlier centers” was also seen for several baseline characteristics and outcome measures. Therefore, choice of surgical center seems to affect outcome after surgery for SPS and DS, although further studies are required to investigate which specific center characteristics may be of particular importance.

Acknowledgments

The National Institute of Arthritis and Musculoskeletal and Skin Diseases (U01-AR45444) and the Office of Research on Women’s Health, the National Institutes of Health, the National Institute of Occupational Safety and Health, and the Centers for Disease Control and Prevention grant funds were received in support of this work.

Relevant financial activities outside the submitted work: consultancy, grant.

References

1. Deyo RA, Weinstein JN. Low back pain. *N Engl J Med*. 2001; 344:363–370. [PubMed: 11172169]
2. Atlas SJ, Keller RB, Wu YA, et al. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. *Spine (Phila Pa 1976)*. 2005; 30:936–943. [PubMed: 15834339]
3. Ekman P, Möller H, Hedlund R. Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis. *Spine (Phila Pa 1976)*. 2009; 34:1204–1210. [PubMed: 19444068]
4. Allen RT, Rihn JA, Glassman SD, et al. An evidence-based approach to spine surgery. *Am J Med Qual*. 2009; 24:15S–24S. [PubMed: 19890181]

5. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. *J Bone Joint Surg Am*. 2009; 91:1295–1304. [PubMed: 19487505]
6. Schurman SJ, Stablein DM, Perlman SA, et al. Center volume effects in pediatric renal transplantation. A report of the North American Pediatric Renal Transplant Cooperative Study. *Pediatr Nephrol*. 1999; 13:373–378. [PubMed: 10412856]
7. Atlas SJ, Chang Y, Kammann E, et al. Long-term disability and return to work among patients who have a herniated lumbar disc: the effect of disability compensation. *J Bone Joint Surg Am*. 2000; 82:4–15. [PubMed: 10653079]
8. Turner JA, Ersek M, Herron L, et al. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. *Spine (Phila Pa 1976)*. 1992; 17:1–8. [PubMed: 1531550]
9. Hurri H, Slätis P, Soini J, et al. Lumbar spinal stenosis: assessment of long-term outcome 12 years after operative and conservative treatment. *J Spinal Disord*. 1998; 11:110–115. [PubMed: 9588466]
10. Birkmeyer NJ, Weinstein JN. Medical versus surgical treatment for low back pain: evidence and clinical practice. *Eff Clin Pract*. 1999; 2:218–227. [PubMed: 10623054]
11. Javid MJ, Hadar EJ. Long-term follow-up review of patients who underwent laminectomy for lumbar stenosis: a prospective study. *J Neurosurg*. 1998; 89:1–7. [PubMed: 9647165]
12. Iguchi T, Kurihara A, Nakayama J, et al. Minimum 10-year outcome of decompressive laminectomy for degenerative lumbar spinal stenosis. *Spine (Phila Pa 1976)*. 2000; 25:1754–1759. [PubMed: 10888941]
13. Rompe JD, Eysel P, Zöllner J, et al. Degenerative lumbar spinal stenosis. Long-term results after undercutting decompression compared with decompressive laminectomy alone or with instrumented fusion. *Neurosurg Rev*. 1999; 22:102–106. [PubMed: 10547007]
14. Djurasovic M, Glassman SD, Carreon LY, et al. Contemporary management of symptomatic lumbar spinal stenosis. *Orthop Clin North Am*. 2010; 41:183–191. [PubMed: 20399357]
15. Mandal AK, Drew N, Lapidus JA. The effect of center volume on pancreas transplant outcomes. *Surgery*. 2004; 136:225–231. [PubMed: 15300184]
16. Sosa JA, Bowman HM, Gordon TA, et al. Importance of hospital volume in the overall management of pancreatic cancer. *Ann Surg*. 1998; 228:429–433. [PubMed: 9742926]
17. Hannan EL, Racz M, Kavey RE, et al. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. *Pediatrics*. 1998; 101:963–966. [PubMed: 9606220]
18. 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; 35:1329–1338. [PubMed: 20453723]
19. Dewing CB, Provencher MT, Riffenburgh RH, et al. The outcomes of lumbar microdiscectomy in a young, active population: correlation by herniation type and level. *Spine (Phila Pa 1976)*. 2007; 33:33–38. [PubMed: 18165746]
20. Jönsson B, Annertz M, Sjöberg C, et al. A prospective and consecutive study of surgically treated lumbar spinal stenosis. Part II: Five-year follow-up by an independent observer. *Spine (Phila Pa 1976)*. 1997; 22:2938–2944. [PubMed: 9431630]
21. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA*. 2006; 296:2451–2459. [PubMed: 17119141]
22. Kizer KW. The volume-outcome conundrum. *N Engl J Med*. 2003; 349:2159–2161. [PubMed: 14645645]
23. Edwards EB, Roberts JP, McBride MA, et al. The effect of the volume of procedures at transplantation centers on mortality after liver transplantation. *N Engl J Med*. 1999; 341:2049–2053. [PubMed: 10615076]
24. Ciol MA, Deyo RA, Howell E, et al. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc*. 1996; 44:285–290. [PubMed: 8600197]
25. Taylor VM, Deyo RA, Cherkin DC, et al. Low back pain hospitalization. Recent United States trends and regional variations. *Spine (Phila Pa 1976)*. 1994; 19:1207–1212. [PubMed: 8073311]

26. Katz JN, Lipson SJ, Chang LC, et al. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine (Phila Pa 1976)*. 1996; 21:92–98. [PubMed: 9122770]
27. Begg CB, Cramer LD, Hoskins WJ, et al. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA*. 1998; 280:1747–1751. [PubMed: 9842949]
28. Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. *N Engl J Med*. 2003; 349:2117–2127. [PubMed: 14645640]
29. Weinstein JN, Lurie JD, Tosteson TD, et al. 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. [PubMed: 19018250]
30. Scholz M, Firsching R, Lanksch WR. Long-term follow up in lumbar spinal stenosis. *Spinal Cord*. 1998; 36:200–204. [PubMed: 9554022]
31. Chou R, Baisden J, Carragee EJ, et al. Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine (Phila Pa 1976)*. 2009; 34:1094–1109. [PubMed: 19363455]
32. Hosenpud JD, Breen TJ, Edwards EB, et al. The effect of transplant center volume on cardiac transplant outcome. A report of the United Network for Organ Sharing Scientific Registry. *JAMA*. 1994; 271:1844–1849. [PubMed: 8196141]
33. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. 1979. *Clin Orthop Relat Res*. 2007; 457:3–9. [PubMed: 17415058]
34. Amundsen T, Weber H, Nordal HJ, et al. Lumbar spinal stenosis: conservative or surgical management?: A prospective 10-year study. *Spine (Phila Pa 1976)*. 2000; 25:1424–1435. [PubMed: 10828926]
35. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev*. 2005:CD001352.
36. Goodney PP, Stukel TA, Lucas FL, et al. Hospital volume, length of stay, and readmission rates in high-risk surgery. *Ann Surg*. 2003; 238:161–167. [PubMed: 12894006]
37. Khuri SF, Daley J, Henderson W, et al. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. *Ann Surg*. 1999; 230:414–429. [PubMed: 10493488]

Key Points

There is statistically significant variation in many short- and long-term outcomes after surgery for spinal stenosis and degenerative spondylolisthesis in centers participating in the SPORT trial when significant differences in baseline characteristics are controlled for.

In addition, there is broad variation and the presence of “outlier centers” for some short-term and long-term outcomes.

The findings suggest that choice of center affects outcomes after these procedures, although further studies are required to investigate which center characteristics are most important.

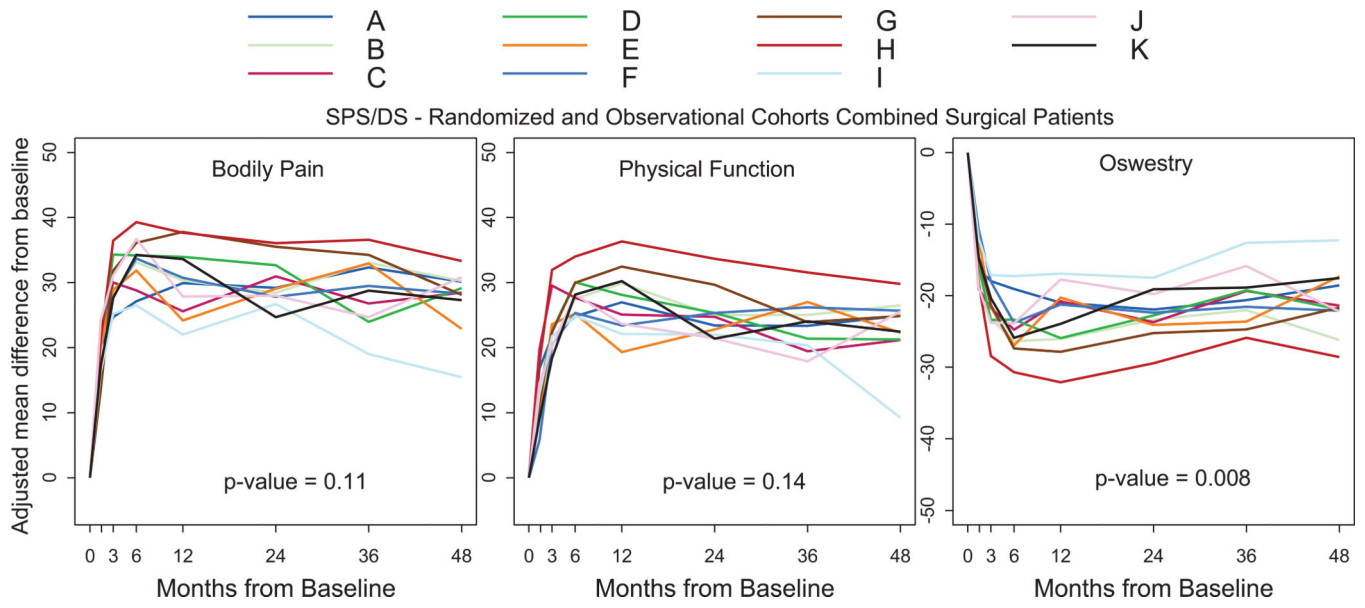


Figure 1.
SPORT SPS+DS surgical patients primary outcome results over time by center

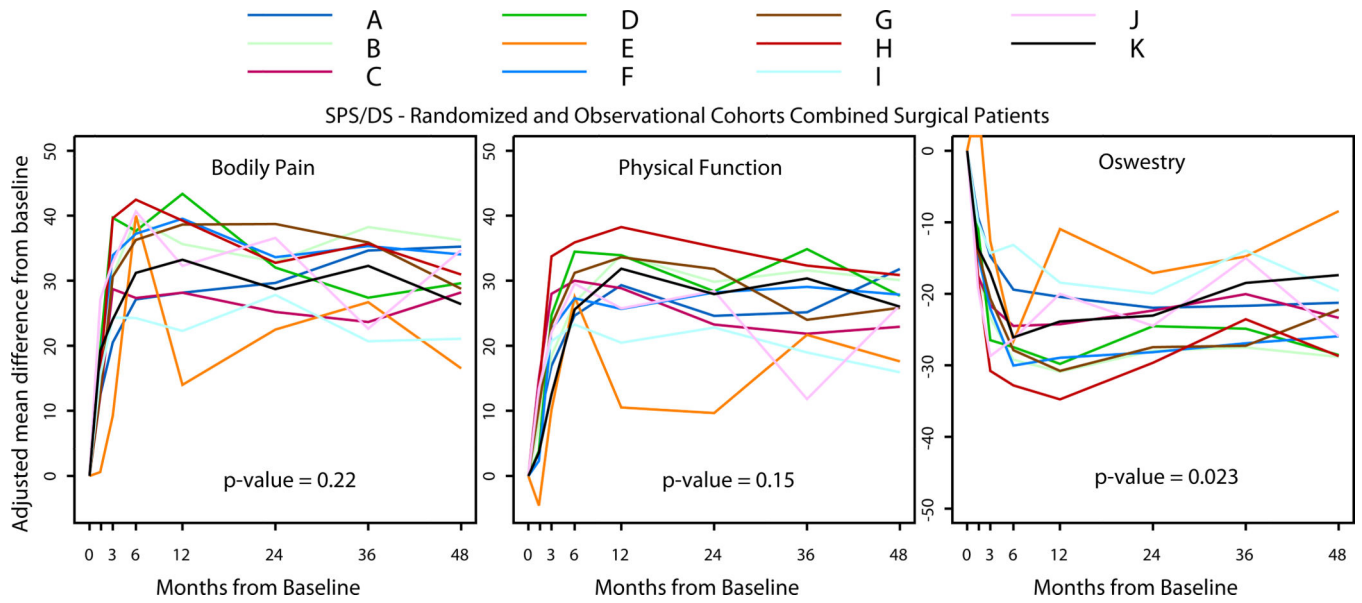


Figure 2. SPORT SPS+DS surgical patients percent of repeated surgery at 4 years follow-up by center

Table 1

Baseline characteristics for SPORT SPS/DS surgical patients by center*

Characteristics (SPS+DS)	A (n=125)	B (n=88)	C (n=90)	D (n=72)	E (n=32)	F (n=77)	G (n=80)	H (n=79)	I (n=29)	J (n=47)	K (n=74)	p-value
Mean Age (SD)	64.3 (10)	65.1 (9.8)	64.9 (12.7)	65.8 (9.7)	64.3 (12.9)	63.9 (12.7)	63.8 (9.8)	62.1 (13.2)	65.9 (9.7)	62.1 (13.1)	64.5 (10.8)	0.66
Female - no.(%)	69 (55%)	51 (58%)	49 (54%)	36 (50%)	16 (50%)	37 (48%)	53 (66%)	34 (43%)	17 (59%)	22 (47%)	38 (51%)	0.27
Ethnicity: Not Hispanic	123 (98%)	87 (99%)	90 (100%)	69 (96%)	31 (97%)	75 (97%)	78 (98%)	79 (100%)	27 (93%)	38 (81%)	72 (97%)	<0.001
Race - White	113 (90%)	73 (83%)	85 (94%)	61 (85%)	23 (72%)	72 (94%)	70 (88%)	69 (87%)	19 (66%)	28 (60%)	60 (81%)	<0.001
Education - At least some college	87 (70%)	49 (56%)	53 (59%)	48 (67%)	22 (69%)	57 (74%)	40 (50%)	54 (68%)	18 (62%)	34 (72%)	52 (70%)	0.035
Marital Status - Married	89 (71%)	60 (68%)	64 (71%)	59 (82%)	21 (66%)	52 (68%)	54 (68%)	59 (75%)	17 (59%)	29 (62%)	53 (72%)	0.42
Work Status												0.065
Full or part time	41 (33%)	32 (36%)	28 (31%)	24 (33%)	10 (31%)	27 (35%)	34 (42%)	35 (44%)	14 (48%)	17 (36%)	24 (32%)	
Disabled	9 (7%)	6 (7%)	11 (12%)	5 (7%)	3 (9%)	4 (5%)	6 (8%)	4 (5%)	3 (10%)	12 (26%)	11 (15%)	
Retired	58 (46%)	39 (44%)	43 (48%)	33 (46%)	14 (44%)	39 (51%)	33 (41%)	32 (41%)	11 (38%)	9 (19%)	32 (43%)	
Other	17 (14%)	11 (12%)	8 (9%)	10 (14%)	5 (16%)	7 (9%)	7 (9%)	8 (10%)	1 (3%)	9 (19%)	7 (9%)	
Compensation - Any	7 (6%)	5 (6%)	10 (11%)	2 (3%)	5 (16%)	6 (8%)	6 (8%)	5 (6%)	1 (3%)	9 (19%)	7 (9%)	0.066
Mean Body Mass Index (BMI), (SD)	29.4 (6)	29.7 (5.1)	30.7 (8.6)	28.1 (5.1)	29.6 (5.8)	27.6 (5.3)	31 (5.8)	28.4 (4.7)	29.7 (4)	28.4 (4.8)	29.4 (5.9)	0.005
Smoker	11 (9%)	9 (10%)	6 (7%)	5 (7%)	3 (9%)	7 (9%)	10 (12%)	8 (10%)	1 (3%)	3 (6%)	7 (9%)	0.95
Comorbidities - no.(%)												
Hypertension	70 (56%)	34 (39%)	29 (32%)	30 (42%)	10 (31%)	35 (45%)	35 (44%)	31 (39%)	12 (41%)	18 (38%)	39 (53%)	0.036
Diabetes	23 (18%)	12 (14%)	12 (13%)	6 (8%)	4 (12%)	10 (13%)	11 (14%)	8 (10%)	4 (14%)	1 (2%)	16 (22%)	0.15
Osteoporosis	14 (11%)	14 (16%)	7 (8%)	7 (10%)	4 (12%)	9 (12%)	3 (4%)	2 (3%)	2 (7%)	3 (6%)	7 (9%)	0.16
Heart Problem	34 (27%)	19 (22%)	16 (18%)	14 (19%)	7 (22%)	21 (27%)	11 (14%)	20 (25%)	6 (21%)	7 (15%)	16 (22%)	0.48
Stomach Problem	30 (24%)	24 (27%)	13 (14%)	14 (19%)	1 (3%)	11 (14%)	20 (25%)	20 (25%)	5 (17%)	7 (15%)	22 (30%)	0.031
Bowel or Intestinal Problem	17 (14%)	10 (11%)	10 (11%)	5 (7%)	3 (9%)	7 (9%)	7 (9%)	2 (3%)	5 (17%)	3 (6%)	11 (15%)	0.27
Depression	10 (8%)	12 (14%)	19 (21%)	8 (11%)	2 (6%)	11 (14%)	13 (16%)	9 (11%)	3 (10%)	8 (17%)	17 (23%)	0.10
Joint Problem	67 (54%)	66 (75%)	45 (50%)	33 (46%)	14 (44%)	28 (36%)	57 (71%)	38 (48%)	15 (52%)	22 (47%)	49 (66%)	<0.001
Other	50 (40%)	23 (26%)	35 (39%)	27 (38%)	10 (31%)	29 (38%)	31 (39%)	27 (34%)	7 (24%)	15 (32%)	41 (55%)	0.038
Time since most recent episode > 6 months	74 (59%)	57 (65%)	59 (66%)	47 (65%)	13 (41%)	47 (61%)	45 (56%)	44 (56%)	16 (55%)	31 (66%)	45 (61%)	0.44

Characteristics (SPS+DS)	A (n=125)	B (n=88)	C (n=90)	D (n=72)	E (n=32)	F (n=77)	G (n=80)	H (n=79)	I (n=29)	J (n=47)	K (n=74)	P-value
SF-36 scores, mean(SD)												
Bodily Pain (BP)	32.1 (18.2)	33.5 (17.3)	30.9 (19.5)	28 (17)	27.8 (16.8)	35.3 (21.5)	30.4 (17.7)	29.9 (18.4)	26.3 (21.4)	27.7 (18.2)	28.9 (18)	0.19
Physical Functioning (PF)	32.7 (21.4)	30.6 (22.4)	31.9 (19.7)	27.3 (21)	31.6 (18.2)	34.7 (20.3)	32.6 (22.1)	35.5 (20.4)	29 (23.7)	31.5 (25.7)	27.3 (21.1)	0.33
Oswestry Disability Index (ODI) (SD)	43 (17.2)	43.6 (17.6)	43.4 (17.1)	47.3 (15.4)	48.2 (17.6)	44.4 (19.6)	44.7 (17.8)	42.9 (15.4)	45.9 (19.2)	52.1 (18.1)	47.4 (16.9)	0.10
Satisfaction with symptoms - very dissatisfied	97 (78%)	64 (73%)	60 (67%)	58 (81%)	26 (81%)	64 (83%)	62 (78%)	62 (78%)	22 (76%)	37 (79%)	61 (82%)	0.41
Patient self-assessed health trend - no.(%)												0.71
Getting better	5 (4%)	3 (3%)	4 (4%)	3 (4%)	1 (3%)	3 (4%)	5 (6%)	2 (3%)	1 (3%)	2 (4%)	4 (5%)	
Staying about the same	36 (29%)	16 (18%)	23 (26%)	15 (21%)	7 (22%)	29 (38%)	23 (29%)	28 (35%)	6 (21%)	10 (21%)	18 (24%)	
Getting worse	84 (67%)	69 (78%)	58 (64%)	54 (75%)	24 (75%)	45 (58%)	52 (65%)	49 (62%)	22 (76%)	30 (64%)	52 (70%)	
Treatment preference at baseline - no.(%)												<0.001
Preference for non-surg	30 (24%)	18 (20%)	21 (23%)	7 (10%)	9 (28%)	5 (6%)	18 (22%)	31 (39%)	10 (34%)	6 (13%)	11 (15%)	
Not sure	19 (15%)	18 (20%)	16 (18%)	19 (26%)	5 (16%)	12 (16%)	19 (24%)	23 (29%)	7 (24%)	6 (13%)	9 (12%)	
Preference for surgery	76 (61%)	52 (59%)	53 (59%)	46 (64%)	18 (56%)	60 (78%)	43 (54%)	25 (32%)	12 (41%)	35 (74%)	54 (73%)	
Pseudoclaudication - Any	73 (58%)	83 (94%)	77 (86%)	66 (92%)	27 (84%)	70 (91%)	71 (89%)	67 (85%)	21 (72%)	44 (94%)	61 (82%)	<0.001
SLR or Femoral Tension	12 (10%)	22 (25%)	18 (20%)	9 (12%)	3 (9%)	23 (30%)	22 (28%)	12 (15%)	1 (3%)	3 (6%)	14 (19%)	<0.001
Pain radiation - any	68 (54%)	64 (73%)	72 (80%)	62 (86%)	24 (75%)	72 (94%)	54 (68%)	68 (86%)	21 (72%)	46 (98%)	66 (89%)	<0.001
Any Neurological Deficit	28 (22%)	52 (59%)	58 (64%)	43 (60%)	17 (53%)	65 (84%)	40 (50%)	48 (61%)	7 (24%)	25 (53%)	47 (64%)	<0.001
Reflexes - Asymmetric Depressed	9 (7%)	10 (11%)	39 (43%)	22 (31%)	6 (19%)	48 (62%)	22 (28%)	26 (33%)	3 (10%)	8 (17%)	14 (19%)	<0.001
Sensory - Asymmetric Decrease	17 (14%)	35 (40%)	31 (34%)	15 (21%)	7 (22%)	32 (42%)	27 (34%)	22 (28%)	3 (10%)	16 (34%)	24 (32%)	<0.001
Motor - Asymmetric Weakness	12 (10%)	19 (22%)	17 (19%)	23 (32%)	11 (34%)	47 (61%)	9 (11%)	20 (25%)	2 (7%)	8 (17%)	30 (41%)	<0.001
Listhesis Level												0.29
L3-L4	3 (2%)	4 (5%)	8 (9%)	2 (3%)	0 (0%)	4 (5%)	1 (1%)	5 (6%)	1 (3%)	3 (6%)	5 (7%)	
L4-L5	39 (31%)	49 (56%)	37 (41%)	27 (38%)	17 (53%)	31 (40%)	50 (62%)	37 (47%)	16 (55%)	18 (38%)	29 (39%)	
Stenosis Levels												
L2-L3	40 (32%)	12 (14%)	15 (17%)	11 (15%)	6 (19%)	17 (22%)	7 (9%)	21 (27%)	4 (14%)	11 (23%)	13 (18%)	0.005
L3-L4	75 (60%)	44 (50%)	49 (54%)	37 (51%)	14 (44%)	47 (61%)	20 (25%)	53 (67%)	15 (52%)	37 (79%)	39 (53%)	<0.001

Characteristics (SPS+DS)	A (n=125)	B (n=88)	C (n=90)	D (n=72)	E (n=32)	F (n=77)	G (n=80)	H (n=79)	I (n=29)	J (n=47)	K (n=74)	P-value
L4-L5	121 (97%)	85 (97%)	82 (91%)	64 (89%)	30 (94%)	74 (96%)	79 (99%)	75 (95%)	28 (97%)	44 (94%)	67 (91%)	0.17
L5-S1	13 (10%)	26 (30%)	13 (14%)	20 (28%)	3 (9%)	12 (16%)	15 (19%)	10 (13%)	3 (10%)	11 (23%)	8 (11%)	0.003
Number of Moderate/Severe Stenotic Levels												<0.001
None	3 (2%)	0 (0%)	0 (0%)	2 (3%)	1 (3%)	1 (1%)	4 (5%)	0 (0%)	3 (10%)	1 (2%)	2 (3%)	
One	61 (49%)	43 (49%)	42 (47%)	41 (57%)	17 (53%)	30 (39%)	51 (64%)	39 (49%)	11 (38%)	9 (19%)	40 (54%)	
Two	36 (29%)	29 (33%)	35 (39%)	21 (29%)	9 (28%)	28 (36%)	21 (26%)	27 (34%)	12 (41%)	28 (60%)	24 (32%)	
Three+	25 (20%)	16 (18%)	13 (14%)	8 (11%)	5 (16%)	18 (23%)	4 (5%)	13 (16%)	3 (10%)	9 (19%)	8 (11%)	
Stenosis Locations												
Central	116 (93%)	80 (91%)	82 (91%)	56 (78%)	27 (84%)	73 (95%)	67 (84%)	74 (94%)	21 (72%)	44 (94%)	66 (89%)	0.002
Lateral Recess	102 (82%)	87 (99%)	72 (80%)	57 (79%)	30 (94%)	71 (92%)	72 (90%)	74 (94%)	26 (90%)	37 (79%)	56 (76%)	<0.001
Neuroforamen	13 (10%)	51 (58%)	37 (41%)	28 (39%)	9 (28%)	17 (22%)	50 (62%)	26 (33%)	15 (52%)	16 (34%)	19 (26%)	<0.001
Stenosis Severity												<0.001
Mild	3 (2%)	0 (0%)	0 (0%)	2 (3%)	1 (3%)	1 (1%)	4 (5%)	0 (0%)	3 (10%)	1 (2%)	2 (3%)	
Moderate	77 (62%)	22 (25%)	37 (41%)	25 (35%)	4 (12%)	18 (23%)	29 (36%)	33 (42%)	8 (28%)	17 (36%)	31 (42%)	
Severe	45 (36%)	66 (75%)	53 (59%)	45 (62%)	27 (84%)	58 (75%)	47 (59%)	46 (58%)	18 (62%)	29 (62%)	41 (55%)	
Instability	5 (4%)	2 (2%)	1 (1%)	2 (3%)	1 (3%)	1 (1%)	1 (1%)	8 (10%)	1 (3%)	9 (19%)	5 (7%)	<0.001

* Patients who received surgery were classified according to whether they received surgical treatment during the first 4 years of enrollment. Centers KAIS (n = 12) and MSR (n = 0) were not included.

† Race or ethnic group was self-assessed. Whites and blacks could be either Hispanic or non-Hispanic.

‡ This category includes patients who were receiving or had applications pending for workers compensation, Social Security compensation, or other compensation.

§ The body-mass index is the weight in kilograms divided by the square of the height in meters.

¶ Other = problems related to stroke, cancer, fibromyalgia, CFS, PTSD, alcohol, drug dependency, lung, liver, kidney, blood vessel, nervous system, migraine or anxiety.

‡‡ The SF-36 scores range from 0 to 100, with higher score indicating less severe symptoms.

‡‡‡ The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms.

§§ The Sciatica Bothersomeness index range from 0 to 24, with lower scores indicating less severe symptoms.

¶¶ The Low Back Pain Bothersomeness Scale ranges from 0 to 6, with lower scores indicating less severe symptoms.

Table 2

Operative treatments, complications and events by center

(SPS+DS)	A (n=126)	B (n=87)	C (n=91)	D (n=69)	E (n=32)	F (n=76)	G (n=81)	H (n=78)	I (n=30)	J (n=47)	K (n=72)	p-value
Procedure, no. (%)												<0.001
Decompression only	80 (65%)	31 (37%)	52 (57%)	39 (57%)	12 (40%)	40 (53%)	22 (27%)	35 (46%)	11 (38%)	23 (49%)	31 (43%)	
Non-instrumented fusion	11 (9%)	10 (12%)	36 (40%)	5 (7%)	4 (13%)	17 (23%)	4 (5%)	8 (11%)	4 (14%)	3 (6%)	1 (1%)	
Instrumented fusion	33 (27%)	43 (51%)	3 (3%)	24 (35%)	14 (47%)	18 (24%)	55 (68%)	33 (43%)	14 (48%)	21 (45%)	40 (56%)	
Multi-level fusion, no. (%)	10 (8%)	12 (14%)	2 (2%)	1 (1%)	3 (9%)	19 (25%)	23 (28%)	15 (19%)	4 (13%)	8 (17%)	7 (10%)	<0.001
Laminectomy Level, no (%)												
L2-L3	38 (30%)	18 (21%)	16 (18%)	10 (15%)	11 (35%)	33 (44%)	8 (11%)	23 (29%)	9 (31%)	12 (26%)	14 (20%)	<0.001
L3-L4	83 (66%)	39 (45%)	46 (51%)	37 (54%)	20 (65%)	61 (81%)	25 (34%)	56 (72%)	18 (62%)	40 (85%)	38 (55%)	<0.001
L4-L5	124 (99%)	84 (97%)	82 (90%)	62 (91%)	29 (94%)	73 (97%)	74 (97%)	75 (96%)	26 (90%)	45 (96%)	64 (90%)	0.045
L5-S1	28 (22%)	49 (56%)	19 (21%)	12 (18%)	6 (19%)	54 (72%)	24 (32%)	40 (51%)	7 (24%)	12 (26%)	12 (17%)	<0.001
Levels decompressed, no. (%)												
0	1 (1%)	0 (0%)	0 (0%)	1 (1%)	1 (3%)	1 (1%)	5 (6%)	0 (0%)	1 (3%)	0 (0%)	1 (1%)	<0.001
1	35 (28%)	27 (31%)	45 (49%)	29 (42%)	10 (31%)	10 (13%)	36 (44%)	13 (17%)	11 (37%)	3 (6%)	30 (42%)	
2	44 (35%)	31 (36%)	23 (25%)	28 (41%)	8 (25%)	12 (16%)	27 (33%)	28 (36%)	7 (23%)	28 (60%)	26 (36%)	
3+	46 (37%)	29 (33%)	23 (25%)	11 (16%)	13 (41%)	53 (70%)	13 (16%)	37 (47%)	11 (37%)	16 (34%)	15 (21%)	
Operation time, minutes (SD)	74 (41.4)	180.8 (69.4)	148.8 (63.4)	186.8 (82.3)	192.3 (73.3)	172.9 (67.1)	235.9 (106.1)	175.9 (69.4)	177.5 (79)	196 (74.8)	181.8 (68.1)	<0.001
Blood loss, cc (SD)	174.5 (144.1)	686.5 (587.5)	441.4 (425.8)	322.1 (228.1)	397.3 (293.9)	686.4 (506.8)	425.5 (355.1)	405.1 (336.1)	512.9 (398)	821.3 (892.5)	296.5 (223.5)	<0.001
Blood Replacement, no. (%)												
Intraoperative replacement	5 (4%)	31 (36%)	10 (11%)	17 (25%)	2 (7%)	24 (32%)	24 (30%)	23 (29%)	5 (17%)	12 (26%)	15 (21%)	<0.001
Post-operative transfusion	7 (6%)	27 (31%)	3 (3%)	2 (3%)	3 (10%)	9 (12%)	12 (15%)	8 (10%)	0 (0%)	6 (13%)	23 (32%)	<0.001
Length of hospital stay, days (SD)**	2.3 (1.3)	3.7 (1.8)	3.1 (3.5)	4.5 (3.1)	5.5 (3.8)	5.1 (2.7)	4.8 (3.5)	3.4 (2)	3.8 (1.4)	5.3 (4.1)	4.7 (2)	<0.001
Intraoperative complications [§]												
Dural tear/ spinal fluid leak	6 (5%)	11 (13%)	11 (12%)	5 (7%)	0 (0%)	13 (17%)	3 (4%)	14 (18%)	3 (10%)	8 (17%)	4 (6%)	0.003
Vascular injury	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.56
Other	0 (0%)	3 (3%)	5 (5%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	2 (3%)	0.047
None	120 (95%)	73 (84%)	76 (84%)	63 (91%)	31 (100%)	62 (83%)	77 (95%)	64 (82%)	27 (90%)	39 (83%)	66 (92%)	0.004

(SPS+DS)	A (n=126)	B (n=87)	C (n=91)	D (n=69)	E (n=32)	F (n=76)	G (n=81)	H (n=78)	I (n=30)	J (n=47)	K (n=72)	p-value	
Postoperative complications/events [¶]													
Nerve root injury	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.53
Wound dehiscence	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.56
Wound hematoma	1 (1%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	2 (2%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0.57
Wound Infection	3 (2%)	0 (0%)	5 (5%)	1 (1%)	3 (10%)	1 (1%)	2 (2%)	4 (5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.033
Other	18 (14%)	8 (9%)	2 (2%)	0 (0%)	5 (17%)	2 (3%)	4 (5%)	9 (12%)	1 (3%)	5 (11%)	6 (8%)	6 (8%)	0.002
None	98 (78%)	55 (64%)	81 (89%)	64 (96%)	21 (70%)	63 (85%)	60 (75%)	61 (78%)	28 (97%)	36 (77%)	45 (62%)	45 (62%)	<0.001
Post-operative mortality (death within 6 weeks of surgery)	0 (0%)	0 (0%)	2 (2.2%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.12
Post-operative mortality (death within 3 months of surgery)	0 (0%)	0 (0%)	2 (2.2%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.27
Additional surgeries (1-year rate) [§]	11 (9%)	1 (1%)	5 (5%)	2 (3%)	3 (9%)	5 (6%)	9 (11%)	5 (6%)	1 (3%)	3 (6%)	2 (3%)	2 (3%)	0.22
Additional surgeries (2-year rate) [§]	17 (13%)	6 (7%)	11 (12%)	3 (4%)	4 (12%)	9 (11%)	12 (15%)	7 (9%)	3 (10%)	5 (11%)	3 (4%)	3 (4%)	0.33
Additional surgeries (3-year rate) [§]	24 (19%)	7 (8%)	13 (14%)	6 (8%)	5 (15%)	10 (12%)	14 (17%)	9 (11%)	3 (10%)	7 (15%)	3 (4%)	3 (4%)	0.13
Additional surgeries (4-year rate) [§]	26 (21%)	11 (13%)	16 (17%)	6 (8%)	5 (15%)	10 (12%)	14 (17%)	10 (13%)	3 (10%)	9 (19%)	4 (5%)	4 (5%)	0.17
Recurrent stenosis / progressive listhesis	15 (12%)	3 (4%)	4 (4%)	3 (4%)	3 (9%)	4 (5%)	1 (1%)	3 (4%)	1 (3%)	6 (13%)	1 (1%)	1 (1%)	
Pseudarthrosis / fusion exploration	1 (0.8%)	1	0	0	0	0	1 (1.3%)	0	0	1	0	0	
Complication or Other	7 (5.6%)	4 (4.7%)	5 (5.5%)	1	1	6 (7.7%)	9 (11.4%)	3 (4.1%)	1	4 (8.5%)	2 (2.8%)	2 (2.8%)	
New condition	4 (3.2%)	2 (2.4%)	6 (6.6%)	0	0	1	3 (3.8%)	1	0	0	0	0	

[§]None of the following were reported: aspiration, nerve root injury, operation at wrong level.

[¶]Any reported complications up to 8 weeks post operation. None of the following were reported: bone graft complication, cerebrospinal fluid leak, paralysis, cauda equina injury, pseudarthrosis.

^{||}One-, two-, three- and four-year post-surgical re-operation rates are Kaplan Meier estimates and p-values are based on the log-rank test. Numbers and percentages are based on the first additional surgery if more than one additional surgery. Surgeries include any additional spine surgery not just re-operation at the same level.

^{***}One of the HSS patients had a length of hospital stay of 372 days—not counting that case the average length of hospital stay for the HSS site would be 5.1 (2.7).

Table 3
Adjusted* Change Scores for Primary Outcomes in the Randomized and Observational Cohorts Combined SPS+DS Surgical Patients According to Center.

Center (SPS+DS)	1Y		2Y		3Y		4Y	
	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)
A	(n = 92)	29.9 (2.5)	(n = 108)	29.2 (2.5)	(n = 98)	32.3 (2.6)	(n = 87)	30.1 (2.8)
B	(n = 73)	30.4 (2.9)	(n = 81)	28.4 (2.9)	(n = 74)	33 (3)	(n = 59)	30.3 (3.2)
C	(n = 73)	25.6 (2.9)	(n = 83)	30.9 (2.9)	(n = 79)	26.8 (2.9)	(n = 71)	28.5 (3)
D	(n = 51)	34 (3.4)	(n = 63)	32.7 (3.4)	(n = 55)	24 (3.6)	(n = 50)	29.2 (3.6)
E	(n = 29)	24.2 (4.6)	(n = 28)	29.1 (4.7)	(n = 24)	33 (5.1)	(n = 23)	22.9 (5.1)
F	(n = 54)	30.7 (3.3)	(n = 69)	27.8 (3.3)	(n = 66)	29.5 (3.4)	(n = 54)	28.3 (3.5)
G	(n = 61)	37.8 (3.1)	(n = 69)	35.5 (3)	(n = 66)	34.3 (3)	(n = 57)	28.1 (3.2)
H	(n = 56)	37.7 (3.1)	(n = 66)	36.1 (3.6)	(n = 66)	36.6 (3.4)	(n = 64)	33.3 (3.2)
I	(n = 24)	22 (5.1)	(n = 26)	26.7 (5.1)	(n = 24)	19 (5)	(n = 21)	15.4 (5.2)
J	(n = 36)	27.9 (4.2)	(n = 36)	28 (4.6)	(n = 33)	24.7 (4.5)	(n = 25)	30.8 (4.7)
K	(n = 62)	33.6 (3.1)	(n = 69)	24.7 (3.2)	(n = 67)	28.8 (3.1)	(n = 58)	27.3 (3.3)
p-value	0.024		0.34		0.05		0.36	

Center (SPS+DS)	1Y		2Y		3Y		4Y	
	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)
A	(n = 92)	26.9 (2.4)	(n = 108)	23.4 (2.5)	(n = 98)	23.4 (2.5)	(n = 87)	24.9 (2.7)
B	(n = 73)	29.9 (2.8)	(n = 81)	25 (2.9)	(n = 74)	25 (2.9)	(n = 59)	26.5 (3.1)
C	(n = 73)	25.1 (2.8)	(n = 83)	24.8 (2.8)	(n = 79)	19.4 (2.9)	(n = 71)	21.2 (3)
D	(n = 51)	28.1 (3.3)	(n = 63)	25.3 (3.3)	(n = 55)	21.4 (3.5)	(n = 50)	21.2 (3.5)
E	(n = 29)	19.3 (4.5)	(n = 28)	22.8 (4.6)	(n = 24)	27 (5)	(n = 23)	22.3 (4.9)
F	(n = 54)	23.4 (3.3)	(n = 69)	25.3 (3.2)	(n = 66)	26.2 (3.2)	(n = 54)	25.7 (3.4)
G	(n = 61)	32.5 (3)	(n = 69)	29.7 (2.9)	(n = 66)	23.9 (3)	(n = 57)	24.9 (3.1)
H	(n = 56)	36.3 (3.1)	(n = 66)	33.7 (3.5)	(n = 66)	31.5 (3.3)	(n = 64)	29.8 (3.2)
I	(n = 24)	22.1 (5)	(n = 26)	22 (5)	(n = 24)	20.4 (4.9)	(n = 21)	9.2 (5.1)

PF	1Y		2Y		3Y		4Y	
	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)
J	(n = 36)	23.7 (4.1)	(n = 36)	21.4 (4.5)	(n = 33)	17.9 (4.4)	(n = 25)	25.6 (4.6)
K	(n = 62)	30.2 (3)	(n = 69)	21.4 (3.1)	(n = 67)	24.1 (3.1)	(n = 58)	22.4 (3.2)
p-value	0.031		0.31		0.30		0.13	

ODI	1Y		2Y		3Y		4Y	
	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)	No. of patients	Mean Change (SE)
A	(n = 92)	-21 (1.9)	(n = 108)	-22 (2)	(n = 98)	-20.6 (2)	(n = 87)	-18.6 (2.1)
B	(n = 73)	-26.1 (2.3)	(n = 81)	-23.4 (2.3)	(n = 74)	-22 (2.3)	(n = 59)	-26.2 (2.4)
C	(n = 73)	-20.8 (2.2)	(n = 83)	-23.7 (2.2)	(n = 79)	-19.3 (2.3)	(n = 71)	-21.4 (2.4)
D	(n = 51)	-25.9 (2.6)	(n = 63)	-22.8 (2.6)	(n = 55)	-19.3 (2.7)	(n = 50)	-22 (2.7)
E	(n = 29)	-20.3 (3.5)	(n = 28)	-24.1 (3.6)	(n = 24)	-23.6 (3.9)	(n = 23)	-17.3 (4.1)
F	(n = 54)	-21.2 (2.6)	(n = 69)	-22.4 (2.5)	(n = 66)	-21.6 (2.5)	(n = 54)	-22.1 (2.7)
G	(n = 61)	-27.9 (2.3)	(n = 69)	-25.2 (2.3)	(n = 66)	-24.7 (2.3)	(n = 57)	-21.7 (2.5)
H	(n = 56)	-32.1 (2.4)	(n = 66)	-29.5 (2.7)	(n = 66)	-25.9 (2.6)	(n = 64)	-28.6 (2.5)
I	(n = 24)	-16.9 (3.9)	(n = 26)	-17.5 (3.9)	(n = 24)	-12.6 (3.8)	(n = 21)	-12.3 (4)
J	(n = 36)	-17.7 (3.2)	(n = 36)	-19.8 (3.5)	(n = 33)	-15.9 (3.5)	(n = 25)	-22.2 (3.6)
K	(n = 62)	-23.9 (2.4)	(n = 69)	-19.1 (2.4)	(n = 67)	-18.9 (2.4)	(n = 58)	-17.5 (2.5)
p-value	0.001		0.25		0.11		0.011	

* Adjusted for age, gender, race, body mass index, diagnosis, education, any neurological deficit, stomach problem, joint problem, other comorbidities**, baseline treatment preference and baseline scores (for SF36 and ODI).

** 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.

Table 4

Adjusted* SPS+DS surgical patients average four years area under curve (AUC) estimates according to center

(SPS+DS)	A	B	C	D	E	F	G	H	I	J	K	P-value
BP	29 (1.9)	29.8 (2.1)	27.5 (2.1)	29.5 (2.4)	27.8 (3.5)	28.5 (2.4)	33.3 (2.2)	35.3 (2.4)	21.6 (3.7)	27.8 (3.2)	28.1 (2.3)	0.11
PF	23.7 (1.9)	25.2 (2.2)	22.9 (2.2)	23.8 (2.4)	22.6 (3.5)	24 (2.4)	26.8 (2.2)	32 (2.4)	19.6 (3.8)	21.4 (3.2)	23.6 (2.3)	0.14
ODI	-20 (1.5)	-23.5 (1.7)	-21.2 (1.7)	-21.8 (1.9)	-21.8 (2.8)	-21.1 (1.9)	-24.4 (1.8)	-28.1 (1.9)	-15 (2.9)	-18.9 (2.5)	-20 (1.8)	0.008

* Adjusted for age, gender, race, body mass index, diagnosis, education, any neurological deficit, stomach problem, joint problem, other comorbidities**, baseline treatment preference and baseline scores for SF36 and ODI.

** 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.