Published in final edited form as:

Med Care. 2014 December ; 52(12): 1055–1063. doi:10.1097/MLR.0000000000000250.

The SPORT Value Compass: Do the Extra Costs of Undergoing Spine Surgery Produce Better Health Benefits?

James N. Weinstein^{1,2,3}, Anna N. A. Tosteson^{2,3}, Tor D. Tosteson^{2,3}, Jon Lurie^{1,2,3}, William A. Abdu^{1,2,3}, Sohail K. Mirza^{1,2,3}, Wenyan Zhao^{2,3}, Tamara S. Morgan^{2,3}, and Eugene C. Nelson^{2,3}

¹The Dartmouth-Hitchcock Medical Center, One Medical Center Dr., Lebanon, NH 03756

²The Dartmouth Institute for Health Policy and Clinical Practice, 35 Centerra Pkwy, Lebanon, NH 03756

³Geisel School of Medicine at Dartmouth, 11 Rope Ferry Rd, Hanover, NH 03755

Abstract

Background—The Spine Patient Outcomes Research Trial (SPORT) aimed to determine the comparative effectiveness of surgical care vs. non-operative care by measuring longitudinal value: outcomes, satisfaction and costs.

Methods—This paper aims to summarize available evidence from the Spine Patient Outcomes Research Trial (SPORT) by addressing two important questions about outcomes and costs for three types of spine problem: 1) How do outcomes and costs of spine patients differ depending on whether they are treated surgically compared to non-operative care? 2) What is the incremental cost per quality adjusted life year for surgical care over non-operative care?

Results—After 4 years of follow up, patients with 3 spine conditions that may be treated surgically or non-operatively, have systematic differences in value endpoints. The average surgical patient enjoys better health outcomes and higher treatment satisfaction but incurs higher costs.

Conclusion—Spine care is preference-sensitive and because outcomes, satisfaction and costs vary over time and between patients, data on value can help patients make better-informed decisions and help payers know what their dollars are buying.

Keywords

low back pain; surgery; non-surgical treatment; preference-sensitive; costs

Introduction

Disorders of the lumbar spine – including intervertebral disk herniation (IDH), spinal stenosis (SPS), and degenerative spondylolithesis (DS) – are common, disabling and costly. Different stakeholders in the healthcare system have important questions that demand

Corresponding Author: Eugene C. Nelson, PhD, The Dartmouth Institute, 35 Centerra Pkwy, Lebanon, NH 03756, 603-650-6670, FAX: 603-653-0896, Eugene.c.nelson@gmail.com, tamara.s.morgan@dartmouth.edu.

Trial Registration: clinicaltrials.gov NCT00000409-DS; NCT00000410-IDH; NCT00000411-SpS

answers. Patients suffering from these problems want to know if the anxiety, risks, recovery time, and costs associated with surgery will result in better and sustained health outcomes. Physicians caring for spine patients want to give them the best information about their chances for pain relief and for regaining full functioning with or without surgery. Payers covering the costs of healthcare for beneficiaries with spine conditions need to know if their investment in surgical care for spine patients will be rewarded by superior health results in the short and long run.

Although the challenges of answering these questions are manifold, they are important questions that require the best answers based on available data. This paper aims to summarize available evidence from the Spine Patient Outcomes Research Trial (SPORT) by addressing two important questions about outcomes and costs for three types of spine problems. Both questions will be answered for each of the three spine conditions based on patient results after two and four years of follow-up.(1-11)

- Outcomes and Costs: How do outcomes and costs of spine patients differ depending on whether they are treated surgically versus non-operatively?
- Cost-Effectiveness: What is the incremental cost per quality adjusted life year (QALY) for surgery versus non-operative care?

As we answer these questions, we hope to demonstrate how the graphical presentation of data, using the clinical value compass approach, provides a parsimonious way to summarize diverse study endpoints associated with the value of healthcare. We show how this tool can be used to answer questions about outcomes and costs for different types of spine problems.

The data used to answer these questions come from SPORT, the large prospective, multicenter, comprehensive cohort study (i.e. a concurrent, prospective, preference-based observational cohort alongside a randomized trial). Most of the literature that has heretofore been published focuses on outcomes at two years (4, 5, 9, 10) or at four years (6-8) for the three spine conditions. Some of the papers report health economic results using QALYs and estimates for the incremental costs per QALY conferred by surgery.(8-10) Publications from SPORT, taken as a whole, provide extensive information on a comprehensive set of data measuring case-mix factors, patient demographics, condition-specific outcomes, general health status outcomes, patient reports on their satisfaction with treatment, and the comparative cost-effectiveness of surgical versus non-surgical care. However, it may be difficult to get a clear overall picture of the results from the vast amount of data in these numerous individual publications. The hallmark of this paper is that it uses a small set of outcome and cost results and a value compass graphical framework to efficiently communicate clinical, patient satisfaction, and economic endpoints of interest to patients, providers and policy makers.

Methods

Overview of Research Design

Extensive information on SPORT's study design and methods has been published previously. (4, 5, 9-11) The SPORT study followed patients from 13 medical centers in 11

states; enrollment started in March 2000 and finished in March 2005. Eligible participants were age 18 and older with well-defined symptoms, physical findings and imaging confirmed diagnoses of spinal stenosis either alone (SPS) or associated with degenerative spondylolisthesis (DS), or of intervertebral disc herniation (IDH). After determining that prospective spine patients were study eligible, they were provided detailed information on the known risks and benefits of treatment alternatives that explained what was known about the risks and benefits of surgical care and non-operative care of patients with their index condition. Patients who agreed to participate in the randomized study were assigned by chance to receive one of two treatment options – surgical care vs. non-operative care. Patients who agreed to participate but were unwilling to be randomized were entered into the concurrent observational cohort based on their treatment preference. All subsets of patients were followed to determine whether or not they actually received spine surgery, regardless of the group that they originally entered, and to measure a comprehensive set of outcomes. Data were collected at 6 weeks, 3 months, 6 months, 1 year, and continues to be collected annually. The randomized component of the study experienced high degrees of cross-over, with substantial proportions of those randomized to non-operative treatment eventually getting surgery and those randomized to surgery never receiving it. This limits the interpretation of an intention-to-treat analysis. Interestingly, analysis revealed a high degree of similarity in both the baseline characteristics and the outcomes of patients in the randomized and observational components of SPORT. As a result, this report is based on an “as treated” analysis of the combined randomized and observational patients and their results at baseline and at 2 and 4 years.

Data Collection Methods

At baseline the following types of data were collected from different sources: diagnosis, spine disease status, spine disease severity, functional health status, general health status, patient demographics, descriptive variables and comorbidities. Then, at the intervals described above, follow-up data were collected to measure changes in health status, patient reports on their care experiences including satisfaction with treatment results, utilization and costs of care during the preceding time interval, and other pertinent data.

Main Outcome Measures

We selected six measures to answer the two questions that are featured in this report.

Oswestry Disability Index (ODI)—The AAOS/Modems version of the ODI uses 9 items to measure patient self-reported, disease-specific burden of illness. (12) It was selected because it has been widely accepted as a valid measure, and changes in the level of the Index are considered to be clinically relevant. A change of 15 points on the ODI is considered a substantial and clinically relevant change.

Physical Function Scale (PF)—This is one of the 8 scales of the SF-36. (13-16) The PF scale uses 10 items to estimate physical function based on patient self-report. It was selected because it is valid, reliable, sensitive to changes in underlying disease states, and among the most widely used measures of functional health. A change of 10 points on the PF is considered substantial and clinically relevant.

Patient Treatment Satisfaction (PTS)—This single-item measure reflects the individual patient's assessment of the degree to which he/she is “very satisfied” or “somewhat satisfied” with the care received; it is an indicator of the patient's perception of treatment benefit.(17-19)

Quality Adjusted Life Years (QALYs)—The EuroQoL EQ-5D with U.S. scoring was used to estimate QALYs. (20, 21) The EQ-5D uses 5 questions (mobility, self-care, pain, usual activities, anxiety) based on patient-reported function at each of 3-levels to define 245 distinct health states. By taking time-weighted estimates of EQ-5D health state values we estimated patient changes in health related quality of life (denoted by QALYs) for surgical treatment or non-operative care.

Costs of Spine Care (CSC)—CSC includes two types of costs – direct costs associated with treatment and indirect social costs associated with having the spine condition. The direct portion of CSC was based on an estimate of the patient's direct costs of receiving spine-related care. Indirect costs were estimated by assessing lost productivity due to spine-related problems. Direct costs were based on patients' reports of utilization of major categories of health care services (e.g., physician visits, emergency department visits, hospital stays, etc.); these were monetized based on applying standard national average costs associated with using these services. Indirect costs were based on patients' reports of the number of days of paid work that they either missed or had to reduce, or caregiver and homemaker expenses imposed by role limitations; these were monetized based on applying each patient's salary level to the amount of time that work was missed or limited due to their spine condition.

Incremental Cost-Effectiveness Ratio (ICER)—This metric uses the data on direct and indirect costs to provide an estimate of the additional expenses associated with gains in QALYs from surgery compared to non-operative treatment. In general, if the cost per QALY is less than \$100,000, then it is often considered by health economists to be sufficiently cost-effective to merit its use in clinical practice in the United States. (22)

In summary, two measures, ODI and PF, reflect the changes in functional health status from baseline to follow-up at either 24 or 48 months. PTS reflects the patient's assessment at 24 or 48- month follow-up of how much they were helped by their treatment. CSC reflects the direct and indirect costs associated with spine problems during the interval starting at baseline and extending until the point of follow-up, a time span of either 24 or 48 months. QALYs and ICERs reflect the improvement associated with treatment observed at 24 or 48 months using a single numeric value to reflect each result.

Data Analysis Methods

Data are reported based on “as treated” patients in each of the three respective disease groups. This means that patients are classified by the actual treatment received (surgical vs. non-operative care at the time of analysis and not on their original study assignment.) The treatment indicator in this analysis is treated as a time-varying covariate, with all outcomes occurring prior to surgery being assigned to the non-operative group and outcomes

occurring after surgery being assigned to the surgery group. The timing of the outcome is assessed relative to the date of surgery (“re-set clock”) and the baseline data for calculating change scores is drawn from the most recent data-point prior to surgery (“re-set baseline”). (6) The outcome and cost measures described above were calculated after 2 and 4 years of follow-up observations.

Detailed descriptions of the analysis methods have been published elsewhere. (4-10) In brief, all longitudinal result measures were adjusted for important case-mix differences that existed at baseline. Primary analyses compared surgical and non-operative treatments using changes from baseline at 2 and 4 years of follow-up using a mixed effects longitudinal regression model which included a random individual effect to correct for associations between repeated measurements within individual patients.

Data are displayed using conventional tables as well as with graphical value compass figures.(23, 24) The basic idea behind the value compass is that the value of healthcare, for an individual patient or a clinical population, is a function of several factors. The compass display provides a compact way to summarize a large amount of data in a single view. Thus, the value compass has four major domains of value-related outcomes and each “point” on the compass displays a particular category of results. In general, the west compass point shows clinical outcomes, the north point shows functional outcomes, the east point shows patient perceptions of care, and the south point shows costs incurred by the patient: direct expenses for medical care (such as out-of-pocket payments and insurance payments) plus indirect societal costs (such as time lost from work).

The value compass displays in this report include specific measures for each domain: ODI (disease outcomes), (PF) functional outcomes, (PTS) patient satisfaction with care, and CSC (direct and indirect costs associated with treatment and role limitations). In addition, cost-effectiveness data on QALYs and ICERs are shown at the center of the compass display.

Results

Characteristics of Patients at Baseline

The results in Table 1 show the similarities and differences across the three disease groups and between the surgical and non-operative treatment groups within disease groups at baseline using the “as treated” classification for patients.

In general, the data suggest that IDH patients tend to be substantially younger, employed, and more likely to smoke, while the SPS and DS patients tend to be older, retired, and less likely to smoke. IDH and SPS patients are more likely to be males whereas DS patients are more likely to be females. The average body mass indices (BMIs) for all cohorts range from 27 to 29, indicating that the average patient is overweight but not obese. Surgery patients tended to have worse functional health status at baseline across all disease groups.

The baseline measures that compare surgical and non-operative IDH patients show several pre-treatment differences that were significant ($p<0.05$). IDH patients in the surgery group were younger, less likely to be working full or part time, had higher BMIs, were less likely

to have joint problems or other co-morbidities, tended to have worse health status (Pain, ODI, PF, Symptoms), were more likely to report that they were getting worse at time of enrollment; and had higher rates of L5-S1 herniation.

Similarly, there were significant differences between surgical and non-operative SPS and DS patients at baseline. SPS and DS surgical patients were younger, scored worse on health measures (Pain, ODI, PF, Symptoms), and were more likely to report that they were getting worse at enrollment. There were, however, no significant differences between surgical and non-operative patients on the levels of stenosis.

The patterns revealed by Table 1 show that, because of the as-treated nature of the analysis, the treatment groups differed in several important respects. These differences underscore the importance of careful case-mix adjustment for baseline differences in the longitudinal regression models to provide a fair comparison between the treatment groups.

Differences in Outcomes and Costs at 2 and 4 Years: Surgical and Non-Operative Patients

The general pattern of results on health outcomes and costs that are contained in Table 2 suggests that at both 2 and 4 years, after adjusting for case-mix differences at baseline, patients in all three disease groups that were treated surgically had better health outcomes, higher levels of treatment satisfaction, and higher healthcare costs. Table 3 shows the unadjusted results for comparison.

With respect to the overall cost-effectiveness of care, surgical patients enjoyed greater gains in QALYs, though the differences between these gains in absolute terms were not large. As for the “bottom line” measure of cost-effectiveness, the ICERs varied greatly. At 2 years, the ICERs for IDH, SPS and DS surgical patients were \$34,355, \$77,600 and \$115,600 respectively, whereas the cost-effectiveness values at 4 years were \$20,600, \$59,400 and \$64,300.

Figures 1, 2, and 3 display the value compass results at 2 years and 4 years for IDH, SpS, and DS patients. The compass layout shows the results for surgical and non-operative patients. The absolute change from baseline on ODI is on the west or “clinical” point of the compass, the change from baseline on PF is on the north or “functional” point of the compass, the PTS is on the east “satisfaction” point of the compass, and the CSC is on the south “cost” point of the compass. Higher absolute values on the ODI, PF and PTS reflect better results for patients whereas greater costs reflect higher expenses incurred by patients. The values in the center of the compass show the QALY results as well as the ICER results. Values shown in blue are for surgical patients and values shown in yellow are for non-operative patients.

The compasses shown in Figure 1 suggest that at 2 years (left display), IDH patients treated surgically as well as non-operatively had large improvements in ODI and PF; however, the surgical patients had greater gains on these metrics, were more satisfied with their treatments, and made greater gains in QALYs. Surgical patients incurred greater expenses for spine care over the 2-year period and the ICER was \$34,355. The same general pattern for health outcomes and treatment satisfaction was sustained at 4 years (right display), but

there is a shift on the cost metrics whereby the gap in costs between surgical and non-operative care diminished and the incremental cost per QALY of surgery over non-operative care decreased (from \$34,355 at 2 years to \$20,600 at 4 years).

The compasses shown in Figure 2 for SPS patients have a pattern similar to that seen for IDH patients, but in this case the gains for both surgical and non-operative patients on health outcomes are muted. After 2 years, both groups show improvement on the ODI and PF but the absolute value of the improvement is much less than for IDH patients. Once again surgical patients are markedly more satisfied with their results and have greater expenses. Surgery is associated with greater gains in QALYs at an incremental cost for surgery of \$77,600 per QALY. After 4 years, health outcome and satisfaction results are for the most part sustained, but total costs continue to increase in both groups while the incremental cost for surgery per QALY decreases to \$59,400.

Finally, the compasses shown in Figure 3 for DS patients provide yet another variation on the patterns seen before. These findings are not muted when contrasted with the IDH patterns, but rather in some respects amplified. At 2 years, the surgical patients show much greater improvement on the ODI and PF and are much more satisfied but they have incurred much higher total costs. The difference in the gain in QALYs by surgical patients compared to non-operative patients is not large whereas the incremental cost-effectiveness ratio for surgery is quite high (\$115,600). The same general pattern holds after 4 years, but the non-operative patients have doubled their total costs whereas the surgical patients have increased but at a much lower rate. All of this produces a more favorable ICER of \$64,300 down from \$115,600.

Discussion

Results Summary

SPORT findings indicate that, for the average patient, surgical treatment produced better health outcomes and higher satisfaction but was more costly. The cost-effectiveness results suggest that these better outcomes for surgery range from \$34,355 (IDH) to \$115,600 (DS) incremental costs per QALY at 2 years, but that these incremental costs decrease at 4 years, ranging from \$20,600 (IDH) to \$64,300 (DS). If we were to adopt a convention that cost-effectiveness ratios of less than \$100,000 per year are reasonable, then we might conclude that surgery is sufficiently cost-effective to warrant use for all three patient populations, if it matches the preferences and values of the individual patient, and if our time horizon is 4 years.

Research Methods: Merits and Limitations

SPORT has methodological limitations that have been discussed in prior reports. (4-10) This paper used an “as-treated” analysis combining both the randomized and observational groups as the basis for the findings. While this is a reasonable choice, because it enables inclusion of all patients with follow-up data, it also dilutes the benefits of focusing solely on patients who were part of the RCT. Other limitations include: (a) patients who were lost to follow-up over time; (b) heterogeneity in treatments within and across centers; and (c)

reliance on patient reports of utilization of care and role limitations in order to develop estimates of the costs of care because claims data were not available. Nevertheless, the pattern of the results between treatment groups was consistent and provides important insights on longitudinal outcomes, comparative effectiveness and the tradeoffs involved in choosing between surgical and non-surgical care.

Relevance of SPORT Trial to Critical Issues

In closing, we briefly touch on the relevance of SPORT to important policy issues.

Patient-Centered Care Supported by Patient-Reported Outcome Measures—

The need for patient-centered care delivery and research is clear, yet the use of patient-reported outcome measures as key endpoints in research has been limited. SPORT contributes to the movement towards acceptance of patient-reported outcome measures as being credible, valid and reliable endpoints for certain diagnoses. The SPORT team is working to provide individual patients with better estimates of value compass treatment outcomes (based on using results from patients with similar characteristics) and thereby refine the data used for shared decision-making.

Health Information Technology (HIT) Infrastructure—While enormous resources are being invested in and electronic health records (EHRs) for use in clinical practice, and “meaningful use” has become the target of HIT development, deployment, and financial incentives,(25) the need for HIT to support the routine collection and analysis of patient-reported outcome data has lagged behind. SPORT underscores the importance of building standard measures of patient-reported outcomes into clinical trials as well as into EHRs, patient-controlled health records, and clinical/research registries for measuring longitudinal outcomes and costs. This practice would allow every patient encounter to become a data-point, thus bridging the gap between clinical research and clinical practice.

Value Measurement for Care Improvement and Healthcare Reform—SPORT offers a model that can be adopted for other clinical populations to measure outcomes in relationship to costs for specific clinical populations over time, i.e. value. Value measurement should include health outcomes that are most relevant to patients, clinicians and payers as well as measures of patient experience and costs of care. (23, 26, 27) We believe the methods we have used to summarize the results from SPORT for spine patients could be applied to many other clinical populations. Measures of value are needed to improve both clinical effectiveness as well as healthcare payment reform that rewards the value of care experienced by patients as opposed to the volume of services delivered to patients. The value compass display provides a way to present complex data to summarize results published separately in many different papers. In the future, graphical displays can be used by patients to help them make more informed decisions based on customized estimates of results for “patients like me.”

Acknowledgments

Funding: 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, and the National Institute of Occupational Safety

and Health, the Centers for Disease Control and Prevention. The Multidisciplinary Clinical Research Center in Musculoskeletal Diseases is funded by NIAMS (P60-AR048094 and P60-AR062799).

References

1. Tosteson AN, Lurie JD, Tosteson TD, Skinner JS, Herkowitz H, Albert T, et al. Surgical treatment of spinal stenosis with and without degenerative spondylolisthesis: cost-effectiveness after 2 years. *Ann Intern Med.* 2008 Dec 16; 149(12):845–53. [PubMed: 19075203]
2. Tosteson AN, Skinner JS, Tosteson TD, Lurie JD, Andersson GB, Berven S, et al. The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). *Spine.* 2008 Sep 1; 33(19): 2108–15. [PubMed: 18777603]
3. Tosteson ANA, Lurie J, Tosteson TD, Weinstein J, Blood EA. 4 yr CEA of IDH/SpS/DS. *Spine.* 2011 In Press.
4. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med.* 2007 May 31; 356(22):2257–70. [PubMed: 17538085]
5. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA.* 2006 Nov 22; 296(20):2451–9. [PubMed: 17119141]
6. Weinstein JN, Lurie JD, Tosteson TD, Tosteson AN, Blood EA, Abdu WA, et al. Surgical versus nonoperative treatment for lumbar disc herniation: four-year results for the Spine Patient Outcomes Research Trial (SPORT). *Spine.* 2008 Dec 1; 33(25):2789–800. [PubMed: 19018250]
7. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, 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 Jun; 91(6):1295–304. [PubMed: 19487505]
8. Weinstein JN, Tosteson TD, Lurie JD, Tosteson A, Blood E, Herkowitz H, 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 Jun 15; 35(14):1329–38. [PubMed: 20453723]
9. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, et al. Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med.* 2008 Feb 21; 358(8):794–810. [PubMed: 18287602]
10. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA.* 2006 Nov 22; 296(20):2441–50. [PubMed: 17119140]
11. Birkmeyer NJ, Weinstein JN, Tosteson AN, Tosteson TD, Skinner JS, Lurie JD, et al. Design of the Spine Patient Outcomes Research Trial (SPORT). *Spine.* 2002 Jun 15; 27(12):1361–72. [PubMed: 12065987]
12. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine.* 2000 Nov 15; 25(22):2940–52. Comparative Study Review. discussion 52. [PubMed: 11074683]
13. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care.* 1994 Jan; 32(1):40–66. [PubMed: 8277801]
14. Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, et al. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. *Jama.* 1989 Aug 18; 262(7):907–13. [PubMed: 2754790]
15. Ware J, Sherbourne D. The MOS 36-item short-form health survey. *Medical Care.* 1992; 30:473–83. [PubMed: 1593914]
16. Ware, JJ. SF-36 Health Survey: Manual and Interpretation Guide. Boston, MA: Nimrod Press; 1993.
17. Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, et al. The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. *Spine.* 1996 Aug 1; 21(15):1787–94. discussion 94-5. [PubMed: 8855463]

18. Deyo RA, Diehl AK. Patient satisfaction with medical care for low-back pain. *Spine*. 1986 Jan-Feb;11(1):28–30. [PubMed: 2939566]
19. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995 Sep 1; 20(17):1899–908. discussion 909. [PubMed: 8560339]
20. The EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16(3):199–208. comment. [PubMed: 10109801]
21. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005 Mar; 43(3):203–20. [PubMed: 15725977]
22. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How Attractive Does a New Technology Have to Be to Warrant Adoption and Utilization - Tentative Guidelines for Using Clinical and Economic Evaluations. *Can Med Assoc J*. 1992 Feb 15; 146(4):473–81. [PubMed: 1306034]
23. Nelson EC, Greenfield S, Hays RD, Larson C, Leopold B, Batalden PB. Comparing outcomes and charges for patients with acute myocardial infarction in three community hospitals: an approach for assessing “value”. *Int J Qual Health Care*. 1995 Jun; 7(2):95–108. Comparative Study Multicenter Study. [PubMed: 7655815]
24. Nelson EC, Mohr JJ, Batalden PB, Plume SK. Improving health care, Part 1: The clinical value compass. *Jt Comm J Qual Improv*. 1996 Apr; 22(4):243–58. [PubMed: 8743061]
25. Jain SH, Seidman J, Blumenthal D. How health plans, health systems, and others in the private sector can stimulate ‘meaningful use’. *Health Aff (Millwood)*. 2010 Sep; 29(9):1667–70. [PubMed: 20820024]
26. Porter ME. What is value in health care? *The New England journal of medicine*. 2010 Dec 23; 363(26):2477–81. [PubMed: 21142528]
27. Porter, ME.; Teisberg, EO. *Redefining health care: creating value-based competition on results*. Boston: Harvard Business School Press; 2006.

Herniated Disk Outcomes

Figure 1. Herniated Disk: Outcomes and Costs at 2 and 4 Years for Surgical and Non-Surgical Patients

Stenosis Outcomes

Figure 2. Spinal Stenosis: Outcomes and Costs at 2 and 4 Years for Surgical and Non-Surgical Patients

Degenerative Spondylolisthesis Outcomes

Figure 3. Degenerative Spondylolisthesis: Outcomes and Costs at 2 and 4 Years for Surgical and Non-Surgical Patients

Table 1
Patient Baseline Demographic Characteristics, Comorbid Conditions, Disease Level and Health Status Measures

	IDH			SPS			DS		
	Surgery (n=789)	Non-operative (n=403)	p-value	Surgery (n=414)	Non-operative (n=220)	p-value	Surgery (n=391)	Non-operative (n=210)	p-value
Mean Age (SD)	40.6 (10.8)	43.9 (12.2)	<0.001	63.8 (12.2)	66.1 (10.5)	0.02	64.7 (10.1)	68.6 (10.2)	<0.001
Female - no. (%)	340 (43%)	167 (41%)	0.63	159 (38%)	90 (41%)	0.60	269 (69%)	143 (68%)	0.93
White - no. (%)	695 (88%)	338 (84%)	0.053	349 (84%)	184 (84%)	0.92	334 (85%)	172 (82%)	0.31
At least some college - no. (%)	572 (72%)	312 (77%)	0.077	260 (63%)	141 (64%)	0.82	264 (68%)	136 (65%)	0.55
Marital Status - Married - no. (%)	549 (70%)	286 (71%)	0.67	301 (73%)	145 (66%)	0.091	263 (67%)	133 (63%)	0.38
Work Status - no. (%)			<0.001			0.36			0.92
Full or part time	470 (60%)	276 (68%)		147 (36%)	69 (31%)		142 (36%)	76 (36%)	
Disabled	121 (15%)	37 (9%)		40 (10%)	20 (9%)		35 (9%)	16 (8%)	
Retired	21 (3%)	24 (6%)		183 (44%)	113 (51%)		167 (43%)	90 (43%)	
Other	177 (22%)	66 (16%)		44 (11%)	18 (8%)		47 (12%)	28 (13%)	
Mean Body Mass Index (BMI) (SD)	28.2 (5.7)	27.5 (5.3)	0.043	29.4 (5.3)	29.8 (6.1)	0.43	29.3 (6.4)	28.9 (5.8)	0.46
Smoker - no. (%)	197 (25%)	85 (21%)	0.16	37 (9%)	25 (11%)	0.40	35 (9%)	16 (8%)	0.69
Comorbidities - no. (%)									
Depression	93 (12%)	48 (12%)	0.97	46 (11%)	24 (11%)	0.96	70 (18%)	28 (13%)	0.18
Joint Problem	127 (16%)	94 (23%)	0.003	222 (54%)	124 (56%)	0.56	218 (56%)	126 (60%)	0.36
Other (IDH)	328 (42%)	198 (49%)	0.015	314 (76%)	184 (84%)	0.03	312 (80%)	164 (78%)	0.70
Health Status Measures, Mean (SD)									
Bodily Pain (BP)	23.2 (17.9)	34.9 (21.9)	<0.001	30.6 (18.5)	38.8 (20.8)	<0.001	31 (18.7)	37.6 (19.3)	<0.001
Physical Functioning (PF)	32.4 (23.4)	48.2 (26.4)	<0.001	31.8 (21.9)	40.4 (24.8)	<0.001	31.4 (20.9)	39.7 (24)	<0.001
Mental Component Summary (MCS)	44.6 (11.4)	46.2 (11.8)	0.026	48.6 (12)	50.7 (11.6)	0.031	49.5 (11.7)	51.1 (11.1)	0.11
Oswestry (ODI)	54.9 (19.6)	38.8 (20.4)	<0.001	45.6 (17.9)	36.3 (18.2)	<0.001	44.5 (16.7)	36.1 (18.7)	<0.001
Very dissatisfied with symptoms - no. (%)	697 (88%)	257 (64%)	<0.001	320 (77%)	113 (51%)	<0.001	303 (77%)	113 (54%)	<0.001

	IDH			SPS			DS		
	Surgery (n=789)	Non-operative (n=403)	p-value	Surgery (n=414)	Non-operative (n=220)	p-value	Surgery (n=391)	Non-operative (n=210)	p-value
Patient self assessed health trend - no. (%)			<0.001			<0.001			<0.001
Getting better	65 (8%)	114 (28%)		14 (3%)	32 (15%)		19 (5%)	19 (9%)	
Staying about the same	339 (43%)	195 (48%)		116 (28%)	87 (40%)		98 (25%)	96 (46%)	
Getting worse	379 (48%)	93 (23%)		277 (67%)	101 (46%)		270 (69%)	91 (43%)	
Herniation Level - no. (%)			<0.001						
L2-L3/L3-L4	42 (5%)	46 (11%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	
L4-L5	305 (39%)	151 (37%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	
L5-S1	442 (56%)	205 (51%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	
Stenosis Levels - no. (%)									
L2-L3	0 (0%)	0 (0%)		124 (30%)	55 (25%)	0.22	34 (9%)	19 (9%)	1
L3-L4	0 (0%)	0 (0%)		279 (67%)	141 (64%)	0.45	156 (40%)	80 (38%)	0.73
L4-L5	0 (0%)	0 (0%)		381 (92%)	198 (90%)	0.47	378 (97%)	202 (96%)	0.94
L5-S1	0 (0%)	0 (0%)		106 (26%)	67 (30%)	0.23	33 (8%)	24 (11%)	0.30

Results at Two and Four Years for Spine Patients Treated Surgically and Non-Surgically**Table 2**

	2 Years: Surgery	2 Years: No Surgery	4 Years: Surgery	4 Years: No Surgery
Herniated Disk				
Oswestry Disability Index (ODI) (SE) [*]	(n = 775) -37.2 (0.68)	(n = 416) -24.6 (0.91)	(n = 789) -38.1 (0.74)	(n = 403) -24.9 (1)
Physical Function (PF) (SE) [†]	43.6 (0.82)	30.2 (1.1)	44.6 (0.89)	29.7 (1.2)
Patient Satisfaction w/Care (%) [‡]	91.7	79.3	91.7	77.3
Total Costs of Spine Care (CSC) (95%CI)	\$27,341 (\$25,882, \$28,799)	\$13,108 (\$11,244, \$14,902)	\$29,061 (\$26,758, \$31,384)	\$22,067 (\$18,979, \$25,825)
Quality Adjusted Life Years (QALYs) (95%CI)	1.64 (1.62, 1.67)	1.44 (1.41, 1.47)	3.24 (3.22, 3.30)	2.90 (2.81, 2.92)
Incremental cost-effectiveness ratio (ICER) (95%CI) [§]	\$34,355 (\$20,419, \$52,512)		\$20,600 (\$4,539, \$33,088)	
Spinal Stenosis				
Oswestry Disability Index (ODI) (SE) [*]	(n = 394) -20.5 (1)	(n = 240) -9.3 (1.2)	(n = 414) -18.7 (1.1)	(n = 220) -9.3 (1.3)
Physical Function (PF) (SE) [†]	23 (1.3)	11.8 (1.4)	20.3 (1.3)	11.6 (1.7)
Patient Satisfaction w/Care (%) [‡]	82.5	66.2	77.8	63.6
Total Costs of Spine Care (CSC) (95%CI)	\$26,222 (\$24,308, \$28,129)	\$13,519 (\$10,921, \$15,796)	\$37,758 (\$35,773, \$43,748)	\$24,611 (\$20,814, \$27,838)
Quality Adjusted Life Years (QALYs) (95%CI)	1.54 (1.51, 1.56)	1.37 (1.33, 1.40)	2.94 (2.87, 2.98)	2.72 (2.63, 2.76)
Incremental cost-effectiveness ratio (ICER) (95%CI) [§]	\$77,600 (\$49,600, \$120,000)		\$59,400 (\$37,059, \$125,162)	
Degenerative Spondylolisthesis				
Oswestry Disability Index (ODI) (SE) [*]	(n = 368) -24.2 (0.98)	(n = 233) -7.5 (1.2)	(n = 391) -23 (1)	(n = 210) -8.6 (1.3)
Physical Function (PF) (SE) [†]	26.6 (1.3)	8.3 (1.5)	26.6 (1.3)	7.7 (1.7)
Patient Satisfaction w/Care (%) [‡]	89.2	67.3	86.6	67.8
Total Costs of Spine Care (CSC) (95%CI)	\$42,081 (\$39,800, \$44,373)	\$16,046 (\$13,862, \$18,234)	\$53,222 (\$49,393, \$57,368)	\$31,095 (\$26,802, \$37,959)
Quality Adjusted Life Years (QALYs) (95%CI)	1.55 (1.53, 1.58)	1.33 (1.30, 1.36)	3.00 (2.96, 3.07)	2.66 (2.56, 2.68)
Incremental cost-effectiveness ratio (ICER) (95%CI) [§]	\$115,600 (\$90,800, \$144,900)		\$64,300 (\$32,864, \$83,117)	

^{*} The Oswestry Disability Index ranges from 0 to 100, with lower scores indicating less severe symptoms

[†] SF-36 scores range from 0 to 100, with higher score indicating less severe symptoms.

[‡] Combines and averages the "Very" and "Somewhat" Satisfied with Care responses (%)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Incremental cost-effectiveness ratio (ICER) is the cost per QALY (EQ-5D) gained for surgical treatment relative to non-operative care with Medicare surgery cost.

Table 3
Unadjusted Mean Outcome Scores and Unadjusted Mean Change from Baseline at One, Two and Four Years for Spine Patients Treated Surgically and Non-Surgically

	Baseline		1-Year		2-Year		4-Year	
	Mean (SD)		Mean (SD) or percent		Mean (SD) or percent		Mean (SD) or percent	
	Surgery	Non-Operative*	Surgery	Non-Operative*	Surgery	Non-Operative*	Surgery	Non-Operative*
Herminated Disk								
Mean Score or percent	(n=789)	(n=391)	(n=659)	(n=340)	(n=663)	(n=322)	(n=526)	(n=280)
EQ-5D US Score (SD)	0.49 (0.22)	0.65 (0.19)	0.83 (0.17)	0.81 (0.15)	0.81 (0.18)	0.82 (0.16)	0.84 (0.17)	0.84 (0.15)
Oswestry Disability Index (ODI) (SD) [§]	55.3 (20)	38.6 (20.5)	15.8 (19)	17.8 (18.4)	15.7 (20.2)	16.2 (18.7)	13.7 (18.8)	14.1 (17.7)
Physical Function (PF) (SD) [‡]	32.4 (23.6)	48.7 (26.3)	78.1 (23.3)	74.8 (23.7)	77.9 (25)	76.8 (24.2)	80.4 (24.1)	78.4 (24.6)
Patient Satisfaction w/Care (%) [‡]			91	84	90	82	90	80
Mean change from baseline								
Oswestry Disability Index (ODI) (SD) [§]	55.3 (20)	38.6 (20.5)	-40.1 (24.5)	-20.3 (21.9)	-39.4 (25.8)	-21.6 (21.9)	-40.4 (24.8)	-22.6 (20.6)
Physical Function (PF) (SD) [‡]	32.4 (23.6)	48.7 (26.3)	46.9 (28.7)	25.3 (24.7)	45.8 (30.2)	26.9 (25.9)	47.6 (29.9)	26.5 (24.9)
Spinal Stenosis								
Mean Score or percent	(n=414)	(n=214)	(n=316)	(n=188)	(n=351)	(n=177)	(n=278)	(n=139)
EQ-5D US Score (SD)	0.58 (0.2)	0.66 (0.2)	0.77 (0.19)	0.74 (0.17)	0.76 (0.19)	0.74 (0.16)	0.75 (0.19)	0.74 (0.18)
Oswestry Disability Index (ODI) (SD) [§]	47 (17.9)	36.4 (18.2)	23.5 (21.7)	27.8 (19.8)	24.7 (21.3)	27 (18.7)	24.7 (20.9)	26 (19.9)
Physical Function (PF) (SD) [‡]	31.3 (21.3)	40.4 (24.7)	58.9 (28.8)	50.6 (28.1)	55.3 (30)	51.9 (28.3)	54.8 (30.4)	52.5 (28.3)
Patient Satisfaction w/Care (%) [‡]			84	73	79	71	76	68
Mean change from baseline								
Oswestry Disability Index (ODI) (SD) [§]	47 (17.9)	36.4 (18.2)	-23 (22.6)	-8.4 (19.9)	-21.3 (22.2)	-8.3 (18.1)	-19.5 (21.9)	-8.9 (20.6)
Physical Function (PF) (SD) [‡]	31.3 (21.3)	40.4 (24.7)	27.7 (29.3)	10.2 (24.4)	23 (30.4)	10.6 (25.1)	22.4 (29.5)	10.2 (25)
Degenerative Spondylolisthesis								
Mean Score or percent	(n=391)	(n=196)	(n=302)	(n=172)	(n=353)	(n=159)	(n=292)	(n=132)
EQ-5D US Score (SD)	0.57 (0.2)	0.65 (0.19)	0.78 (0.18)	0.74 (0.15)	0.77 (0.17)	0.73 (0.18)	0.77 (0.18)	0.75 (0.16)
Oswestry Disability Index (ODI) (SD) [§]	46 (17.2)	36.3 (18.8)	20.1 (20.4)	28.2 (19.7)	21.2 (19.7)	27.9 (21.1)	22.1 (19.9)	27.5 (20.5)

	Baseline			1-Year			2-Year			4-Year		
	Mean (SD)			Mean (SD) or percent			Mean (SD) or percent			Mean (SD) or percent		
	Surgery	Non-Operative*		Surgery	Non-Operative*		Surgery	Non-Operative*		Surgery	Non-Operative*	
Physical Function (PF) (SD) [‡]	29.6 (20.6)	39.1 (23.8)		58.7 (28.2)	51.5 (26.4)		57.9 (28.6)	50.6 (28.7)		57.1 (28.6)	48.1 (28.7)	
Patient Satisfaction w/Care (%) [‡]				89	69		88	69		83	67	
Mean change from baseline												
Oswestry Disability Index (ODI) (SD) [§]	46 (17.2)	36.3 (18.8)		-25.8 (20.8)	-8.4 (17.2)		-24.6 (20.2)	-8.2 (17.8)		-24.1 (21)	-8.4 (19.2)	
Physical Function (PF) (SD) [‡]	29.6 (20.6)	39.1 (23.8)		29 (27)	11.7 (23.4)		27.9 (27.3)	10.5 (23.3)		27.1 (27.6)	8.5 (23.8)	

* Non-Operative is the group of patients who never had surgery and never crossed over to surgery.

[‡]The SF-36 Physical Function score 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.

[†]Combines and averages the “Very” and “Somewhat” Satisfied with Care responses (%).