Cost-effectiveness of Operative versus Nonoperative Treatment of Adult Symptomatic Lumbar Scoliosis an Intent-to-treat Analysis at 5-year Follow-up

Leah Y. Carreon, MD, MSc, Steven D. Glassman, MD,*, Jon Lurie, MD, MS, Christopher I. Shaffrey, MD,§ Michael P. Kelly, MD, MSc, Christine R. Baldus, RN, MHA, Kelly R. Bratcher, RN,§ Charles H. Crawford, III, Elizabeth L. Yanik, PhD, and Keith H. Bridwell, MD

Study Design. Secondary analysis using data from the NIH-sponsored study on adult symptomatic lumbar scoliosis (ASLS) that included randomized and observational arms.

Objective. The aim of this study was to perform an intent-to-treat cost-effectiveness study comparing operative (Op) versus nonoperative (NonOp) care for ASLS.

Summary of Background Data. The appropriate treatment approach for ASLS continues to be ill-defined. NonOp care has not been shown to improve outcomes. Surgical treatment has been shown to improve outcomes, but is costly with high revision rates.

Methods. Patients with at least 5-year follow-up data were included. Data collected every 3 months included use of NonOp modalities, medications, and employment status. Costs for index and revision surgeries and NonOp modalities were determined using Medicare Allowable rates. Medication costs were determined using the RedBook and indirect costs were calculated based on reported employment status and income. Quality-adjusted life year (QALY) was determined using the SF6D.

Results. There were 81 of 95 cases in the Op and 81 of 95 in the NonOp group with complete 5-year follow-up data. Not all patients were eligible 5-year follow-up at the time of the analysis. All patients in the Op and 24 (30%) in the NonOp group had surgery by 5 years. At 5 years, the cumulative cost for Op was $96,000 with a QALY gain of 2.44 and for NonOp the cumulative cost was $49,546 with a QALY gain of 0.75 with an incremental cost-effectiveness ratio (ICER) of $27,480 per QALY gain.

Conclusion. In an intent-to-treat analysis, neither treatment was dominant, as the greater gains in QALY in the surgery group come at a greater cost. The ICER for Op compared to NonOp treatment was above the threshold generally considered cost-effective in the first 3 years of the study but improved over time and was highly cost-effective at 4 and 5 years.

Key words: adult lumbar scoliosis, cost-effectiveness, decision analysis, nonoperative treatments, spine surgery.

Level of Evidence: 2

Spine 2019;44:1499–1506

There continues to be uncertainty regarding the appropriate treatment approach for patients with adult spinal deformity. Similar to patients with lumbar degenerative disorders, patients are initially managed with nonoperative (NonOp) treatments such as physical therapy, manipulation, injections, and medication. However, there is no firm evidence that these NonOp interventions are effective. A systematic review found only level II evidence for the use of injections and level IV evidence for physical therapy and bracing. Another study showed that NonOp care does not improve a patient’s quality of life as measured by the Short Form 12(SF-12), the Scoliosis Research Society 22R (SRS22R), or the Oswestry Disability Index (ODI).
Although NonOp treatment did not change health status, it is unknown whether these patients would have worsened without treatment.\textsuperscript{1,4}

Surgical treatment is usually considered when NonOp treatment has failed to improve pain and function. The surgeon, the patient, the patient’s caregiver, and the payer all contribute to the decision of whether the risk of surgery outweighs the potential benefits. Several studies have shown that surgical treatment improves a patient’s quality of life but is costly with high complication\textsuperscript{1} and revision rates.\textsuperscript{9}

The purpose of this study is to perform cost-effectiveness analysis comparing operative (Op) to NonOp treatment using data from subjects enrolled in the National Institutes of Health (NIH) sponsored study on adult symptomatic lumbar scoliosis (ASLS) using an intent-to-treat analysis.

**METHODS**

This is a secondary analysis of prospective data collected from subjects enrolled in nine centers in North America. The primary study evaluated Op and NonOp treatments in patients with ASLS and included randomized and observational arms. It is widely accepted that patients with ASD represent a heterogeneous population in terms of symptoms, functional limitations, and radiographic characteristics. To mitigate this heterogeneity, the present study included only patients between 40 and 80 years’ old with lumbar scoliosis with coronal Cobb $\geq 30^\circ$, and either an ODI score of $\geq 20$, or an Scoliosis Research Society 22R (SRS-22R) Pain, Function, or Appearance score $<4.0$, and no previous fusion surgery. Exclusion criteria were the presence of medical comorbidities that precluded surgery, high-grade ($\geq 3$) spondylolisthesis, previous thoracic or lumbar fusion, previous multilevel ($\geq 3$) thoracolumbar decompression, severe osteoporosis (femoral neck $t$ score $\geq -3.0$), neuromuscular scoliosis and presence of congenital lumbar spine anomalies. Subjects were enrolled from 2010 to 2014. Funding was provided by the NIH through an RO1 grant: A Multi-Center Prospective Study of Quality of Life in Adult Scoliosis (R01AR055176-01A2). Institutional review board approval was obtained at each participating center before subject enrollment. Institutional review board approval was also obtained for this secondary analysis.

The study database was queried in October 2017 for patients undergoing Op or NonOp treatment with at least 5-year follow-up data. Not all patients had reached the 5-year follow-up time point. Surgical approach, technique, and levels fused in the Op cohort were under the discretion of the treating surgeon. Data collected every 3 months included frequency of use of NonOp modalities (physical therapy, chiropractor, pain management visits and epidural steroid injections), medication use (nonsteroidal anti-inflammatory drug, opioids, muscle relaxants), and employment status. Health-related quality-of-life scores (HRQOLs) collected included the ODI,\textsuperscript{7,8} the SRS-22R,\textsuperscript{6} and the SF-12.\textsuperscript{5}

The cumulative incremental cost-effectiveness ratios (ICERs) at each year of follow-up were determined. ICER is the difference in cost between two possible interventions, in this case NonOp versus Op treatment for ASLS, divided by the difference in their effect. Total costs included all surgical and nonsurgical costs in both cohorts. Surgical costs for the index and revision surgeries within 5 years and direct costs for NonOp care were determined using Medicare Allowable rates.\textsuperscript{30} Medication costs were determined using the lowest price quoted\textsuperscript{11} and indirect costs were determined based on reported employment status and income. Treatment effectiveness in terms of quality-adjusted life years (QALY) was determined using the Short Form-6 Dimensions\textsuperscript{12} derived from the SF-12. Although determination of a specific dollar value threshold for appropriate medical treatment is controversial, interventions with a cost per QALY gained (cost/QALY) between $50,000 and $100,000, or less, are generally considered cost-effective.\textsuperscript{13–16} A treatment is dominant if it costs less and is more effective compared to the alternative treatment.\textsuperscript{17}

The intent-to-treat analysis was used to allow for the inclusion of costs and QALY gains or losses associated with NonOp care before and after a surgical intervention, as well as the cost and QALY gains of surgery following attempts at NonOp care. For a patient who crosses over from NonOp to Op at 2 years after enrollment the cost and benefit at 2 years is attributed to the nonsurgical treatment. But at the 5-year follow-up time-point, the benefit of surgery 3 years post-op (5 years from enrollment) and the expense of continued nonsurgical and surgical treatments are all attributed to the NonOp intervention.

Baseline characteristics of the Op and NonOp cohorts were compared using unpaired independent $t$ tests for continuous variables and Fisher exact test for categorical variables. As this is a secondary analysis with multiple concurrent comparisons, a conservative $P$ value threshold of 0.05 was used for the difference to be statistically significant.

TreeAge Pro was used to perform the cost-effectiveness analysis to take into account the incidence of index and revision surgeries in both groups, as patients were allowed to cross-over to either arm during the study. A decision tree was used instead of using mean costs and mean QALY gains for each cohort. Using group means does not take into account the incidence of index and revision surgeries in both groups, as patients were allowed to cross-over to either arm during the study. A decision tree was used instead of using mean costs and mean QALY gains for each cohort. Using group means does not take into account the incidence of index and revision surgeries in each cohort. Subjects undergoing revision surgery would incur higher costs with less QALY gain. This higher cost and less QALY gain cannot be applied equally to each subject in the cohort. The decision tree model allows for apportioning of these higher costs and lower QALY gains only to the subjects who have revision surgery.\textsuperscript{18,19}

**RESULTS**

A total of 286 were enrolled in the overall study; 190 were eligible for 5-year follow-up at the time of analysis. At 5 years after enrollment, 81 (85%) of the 95 cases in the
NonOp group eligible for 5-year follow-up had data available, and 81 (85%) cases of 95 cases in the Op group eligible for 5-year follow-up had data available. All other patients (96) enrolled had not yet reached that time point. The proportion of patients from the Observational arm was similar between the Op (68, 84%) and NonOp groups (69, 85%, \( P = 0.828 \)). The Op and NonOp groups were similar in age, sex distribution, and body mass index (Table 1). There were 5 (6%) smokers in the NonOp group and none in the Op group (\( P = 0.049 \)). The baseline HRQOL scores including the ODI, SF-6D, SF-12 PCS, and SF-12 MCS were similar between the two groups. Baseline SRS-22R for Pain and Mental Health were similar for both groups. Patients in the NonOp group (3.14) had statistically better baseline SRS-22R Self-image scores compared to the Op group (2.74, \( P < 0.000 \)); Patients in the NonOp group (3.44) also had statistically better SRS-22R function compared to the Op group (3.22, \( P = 0.034 \)). The major Cobb angle was significantly smaller in the NonOp group (49.05°), compared to the Op group (55.91°, \( P = 0.001 \)), but this 7° difference may not be clinically relevant.

In the Op group, all patients had their index surgery within 6 months of enrollment; 24 patients had one revision surgery and one patient had two revisions after the index surgery within five years. In the NonOp group, 24 patients had surgery by 5 years, with 5 patients having one revision surgery, one patient having two revisions, and one patient having three revisions after the index surgery.

In the first year after enrollment, 10 (12%) of the 81 patients on the NonOp group had surgery and seven (9%) of the patients in the Op group had a revision surgery. The cumulative cost in the first year for the NonOp group was \$17,301 with a QALY gain of 0.05. The cumulative cost for the Op group was \$65,011 with a QALY gain of 0.16. The ICER was \$440,931 per QALY (Figure 1).

By the second year, 16 (20%) of patients in the NonOp group had surgery, two (12%) of whom had revision surgery, whereas 14 (17%) in the Op group had revision surgery. The cumulative cost over 2 years for the NonOp group was \$30,236 with a QALY gain of 0.17. The cumulative cost over 2 years for the Op group was \$74,102 with a QALY gain of 0.47. The ICER was \$143,900 per QALY (Figure 2).

By the third year, 17 (21%) in the NonOp group had surgery, of whom 2 had one revision surgery and one had

| TABLE 1. Summary of Baseline Characteristics of Subjects in the Nonoperative and Operative Cohorts |
|----------------------------------|-----------------|-----------------|
| Nonoperative                    | Operative       |
| N                               | 81              | 81              |
| Arm, N (%)                      | 71 (88%)        | 77 (95%)        |
| Smoking status, N (%)           | 36.87 (9.67)    | 35.42 (10.14)   |
| Sciroxor Research Society 22R  |
| Pain                            | 2.96 (0.66)     | 2.94 (0.75)     |
| Function                        | 3.44 (0.61)     | 3.22 (0.73)     |
| Appearance                      | 3.14 (0.69)     | 2.74 (0.72)     |
| Mental health                   | 3.78 (0.37)     | 3.79 (0.37)     |
| Subscore                        | 3.31 (0.48)     | 3.17 (0.56)     |
| Coronal Cobb magnitude, °, mean (SD) | 49.05 (11.51)  | 55.91 (14.16)  |
| Coronal balance, mm, mean (SD)  | 20.38 (18.05)   | 21.83 (17.25)   |
| Sagittal balance, mm, mean (SD) | 23.84 (18.19)   | 22.74 (16.81)   |

SD indicates standard deviation.
two revisions, with a cumulative cost of $35,863 and QALY gain of 0.32. In the Op group, 19 (23%) had one revision by the third year, with a cumulative cost of $79,987 and QALY gain of 0.95. This results in an ICER of $70,210 per QALY (Figure 3).

By the fourth year, 21 (26%) of patients in the NonOp group have had surgery with four having had one revision and one patient having two revisions with a cumulative cost of $39,669 and QALY gain of 0.52. In the Op group, 22 patients had one revision and one patient had two revisions, with a cumulative cost of $84,233 and QALY gain of 1.62. This results in an ICER of $40,595 (Figure 4).

The cumulative cost including direct and indirect costs for the Op group at 5 years was $96,000 with a cumulative QALY gain of 2.44. The cumulative cost for the NonOp group was $49,546 with a cumulative QALY gain of 0.75. This results in an ICER of $27,480 (Figure 5). Op treatment was not dominant as the higher utility gains came at a higher cost; however, this ICER is well below the threshold to be considered highly cost-effective.

**DISCUSSION**

Most published studies directly comparing surgical and nonsurgical treatment of patients with spinal deformity have evaluated outcomes using widely accepted and validated measures of HRQOL.2,20–23 Only four studies have evaluated costs associated with spinal deformity surgery.24–27 Three of these studies presented cost per QALY (cost/ QALY) for surgical treatment only.24,26,27 Another presented cost/QALY of surgical treatment compared to a statistically modeled nonsurgical treatment arm, assuming that the surgical cohort had not undergone surgical treatment.25 Only one study has reported on measured costs of NonOp treatment for this patient population.4 All of these studies used data that were collected retrospectively over a 2-year time horizon. The present study is the first study comparing both QALY gain and costs for surgically and nonsurgically treated patients with adult lumbar spinal deformity using prospectively collected data over a 5-year period.

As can be expected in comparative cohorts that are not completely randomized, there were some baseline differences between the two groups. As most surgeons will not perform spine fusions on smokers because of the higher risk of non-union,28 there were no smokers in the Op group. In addition, the SRS-22R domain scores for Function and Appearance were worse in the Op compared to the NonOp cohort. However, the differences between the means are less

---

**Figure 2.** Flowchart of subjects in the second year of follow-up. In the NonOp group 16 (20%) subjects had surgery, 2 (12%) of whom had revision surgery, whereas 14 (17%) in the Op group had revision surgery. The cumulative cost over 2 years for the NonOp group was $30,236 with a QALY gain of 0.17. The cumulative cost over 2 years for the Op group was $74,102 with a QALY gain of 0.47. The ICER was $143,900 per QALY. QALY indicates quality-adjusted life years.
than a half point, which may make these differences clinically irrelevant.

The ODI scores, a measure of low back disability, were similar in the Op and NonOp cohorts in the present study. This is in contrast to previously published studies in adult spine deformity that showed worse baseline ODI scores for patients in the Op group compared to the NonOp group. In addition, the radiographic parameters were also similar, with the coronal Cobb magnitude in the Op group being only larger than in the NonOp group. The similarities between the Op and NonOp cohorts allow for a more straightforward comparison of these two interventions compared to previous studies.

Similar to previous adult scoliosis studies, the present study showed that utility gains over all time points were greater in the Op compared to the NonOp group. QALY gains at 2 years in the Op group were similar to previous studies, but QALY gains at 2 years in the NonOp group were substantially better than in previously reported nonsurgical cohorts. The documentation of modest QALY gains in the 57 patients who did not have surgery at any point in this prospective study, is in marked contrast to what has been predicted using statistical models as a maintenance or loss of QALYs in previous studies.

This is the first study to include direct costs owing to continued use of nonsurgical treatment modalities after the index surgery and indirect costs owing to lost wages. Previous studies that reported on the cost of surgical treatment for adult spine deformity considered only direct costs associated with the index surgery and any revisions within the study-specified time horizon. Thus, the previously predicted cost per QALY gained in these previous studies was lower than that reported in the present study.

Previous studies have used modeling and not actual data to predict cost and QALY gains at longer follow-up time points, beyond the 2-year time horizon. However, as shown in the present study, assumptions made during the modeling process may not hold true in actuality. Some studies predicted a deterioration in HRQOLs without surgery and an incremental improvement or maintenance of improvement with surgery. The present study showed small improvements in HRQOLs without surgery and deterioration in HRQOLs with multiple revision surgeries.
The results of the present study showed that in the first 3 years of the study Op treatment produced more QALYs but at an increased cost that did not reach the widely accepted Willingness-to-Pay threshold of less than $50,000. This reverses on the fourth year of the study when Op treatment becomes cost-effective relative to NonOp treatment with an ICER of $40,496 and even more cost-effective at the 5 years after enrollment with an ICER of $27,480. The improving cost-effectiveness over time is not surprising, as even single-level fusion for spondylolisthesis, the most widely accepted and least invasiveness indication for fusion surgery, only becomes cost-effective at 5 years after surgery. This is because of the high up-front costs for Op treatment combined with higher QALY gains that is maintained over the 5-year period. However, Op treatment is not a dominant intervention, as the higher QALY gained came at a higher cost.

There are limitations to the present study. The two groups may not be directly comparable, as the Op group had no smokers, slightly larger coronal Cobb angles, and worse SRS-22R Appearance and Function domain scores. Indirect costs owing to lost wages by the caregiver were not considered. Costs were assigned to healthcare interventions used instead of using actual individual costs per patient. Although the number of patients available for follow-up was high, there still may be some selection bias.

Using an intent-to-treat analysis, Op treatment of ASLS becomes cost-effective relative to NonOp treatment with an ICER of $40,595 at 4 years after enrollment that is maintained at 5 years with an ICER of $27,480, despite the high rate (28%) of revisions. Much larger gains in QALY are seen in the much more expensive Op treatment, whereas very modest gains in QALY are seen in the less expensive NonOp treatment. Thus, neither treatment arm is dominant. Longer follow-up is needed to determine whether this ICER will improve or not based on the need for revision surgeries, continued use of nonsurgical treatment modalities, and gains or losses in QALY in both the Op and NonOp cohorts.
**Key Points**

- An intent-to-treat cost-effectiveness study comparing Op versus NonOp care using patients with 5-year follow-up data from the NIH sponsored study on ASLS that included randomized and observational arms was performed.
- There were 81 of 95 cases in the Op and 81 of 95 in the Non-Op group with complete 5-year follow-up data. All patients in the Op group and 24 (30%) in the NonOp group had surgery by 5 years.
- In an intent-to-treat analysis, neither treatment was dominant, as the greater gains in QALY in the surgery group come at a greater cost.
- At 5 years, the cumulative cost for Op was $96,000 with a cumulative QALY gain of 2.44 and for NonOp the cumulative cost was $49,546 with a QALY gain of 0.75 with an ICER of $27,480 per QALY gain.
- The ICER for Op compared to NonOp treatment was above the threshold generally considered cost-effective in the first three years of the study but improved over time and was highly cost-effective at 4 and 5 years.

**Acknowledgments**

The authors acknowledge Oheneba Boachie-Adjei, MD, Charles C. Edwards II, MD, Tyler R. Koski, MD, Stephen H. Lewis, MD, Stefan Parent, MD, PhD, Frank J. Schwab, MD, Jacob M. Buchowski, MD, MD, Lawrence G. Lenke, MD, Han Jo Kim, MD, and Justin D. Smith, MD, for their participation in this study.

**References**


**Figure 5.** Flowchart of subjects in the fifth year of follow-up. The cumulative cost, including direct and indirect costs for the Op group at 5 years was $96,000 with a cumulative QALY gain of 2.44. The cumulative cost for the NonOp group was $49,546 with a cumulative QALY gain of 0.75. This results in an incremental cost-effectiveness ratio of $27,480. QALY indicates quality-adjusted life year.