JNS SPINE

CLINICAL ARTICLE

J Neurosurg Spine 37:80–91, 2022

Patient-reported outcome measure clustering after surgery for adult symptomatic lumbar scoliosis

James P. Wondra II, BS,[1] Michael P. Kelly, MD, MSc,[1] Elizabeth L. Yanik, PhD,[1] Jacob K. Greenberg, MD, MSCI,[1] Justin S. Smith, MD, PhD,[2] Shay Bess, MD,[3] Christopher I. Shaffrey, MD,[4] Lawrence G. Lenke, MD,[5] and Keith Bridwell, MD[1]

[1]Department of Orthopedic Surgery, Washington University School of Medicine, St. Louis, Missouri; [2]Department of Neurological Surgery, University of Virginia Health System, Charlottesville, Virginia; [3]Denver International Spine Center, Denver, Colorado; [4]Department of Neurological Surgery, Duke University, Durham, North Carolina; and [5]Och Spine Hospital, Columbia University College of Physicians and Surgeons, New York, New York

OBJECTIVE Adult symptomatic lumbar scoliosis (ASLS) is a widespread and debilitating subset of adult spinal deformity. Although many patients benefit from operative treatment, surgery entails substantial cost and risk for adverse events. Patient-reported outcome measures (PROMs) are patient-centered tools used to evaluate the appropriateness of surgery and to assist in the shared decision-making process. Framing realistic patient expectations should include the possible functional limitation to improvement inherent in surgical intervention, such as multilevel fusion to the sacrum. The authors' objective was to predict postoperative ASLS PROMs by using clustering analysis, generalized longitudinal regression models, percentile analysis, and clinical improvement analysis of preoperative health-related quality-of-life scores for use in surgical counseling.

METHODS Operative results from the combined ASLS cohorts were examined. PROM score clustering after surgery investigated limits of surgical improvement. Patients were categorized by baseline disability (mild, moderate, moderate to severe, or severe) according to preoperative Scoliosis Research Society (SRS)–22 and Oswestry Disability Index (ODI) scores. Responder analysis for patients achieving improvement meeting the minimum clinically important difference (MCID) and substantial clinical benefit (SCB) standards was performed using both fixed-threshold and patient-specific values (MCID = 30% of remaining scale, SCB = 50%). Best (top 5%), worst (bottom 5%), and median scores were calculated across disability categories.

RESULTS A total of 171/187 (91%) of patients with ASLS achieved 2-year follow-up. Patients rarely achieved a PROM ceiling for any measure, with 33%–43% of individuals clustering near 4.0 for SRS domains. Patients with severe baseline disability (< 2.0) SRS-pain and SRS-function scores were often left with moderate to severe disability (2.0–2.9), unlike patients with higher (≥ 3.0) initial PROM values. Patients with mild disability according to baseline SRS-function score were unlikely to improve. Crippling baseline ODI disability (> 60) commonly left patients with moderate disability (median ODI = 32). As baseline ODI disability increased, patients were more likely to achieve MCID and SCB ($p < 0.001$). Compared to fixed threshold values for MCID and SCB, patient-specific values were more sensitive to change for patients with minimal ODI baseline disability ($p = 0.008$) and less sensitive to change for patients with moderate to severe SRS subscore disability ($p = 0.01$).

CONCLUSIONS These findings suggest that ASLS surgeries have a limit to possible improvement, probably due to both baseline disability and the effects of surgery. The most disabled patients often had moderate to severe disability (SRS < 3, ODI > 30) at 2 years, emphasizing the importance of patient counseling and expectation management.

https://thejns.org/doi/abs/10.3171/2021.11.SPINE21949

KEYWORDS adult; lumbar; patient-reported outcomes; scoliosis; spine deformity; surgery

ABBREVIATIONS ASD = adult spinal deformity; ASLS = adult symptomatic lumbar scoliosis; HRQOL = health-related quality of life; MCID = minimum clinically important difference; ODI = Oswestry Disability Index; PROM = patient-reported outcome measure; SCB = substantial clinical benefit; SRS = Scoliosis Research Society.

SUBMITTED August 13, 2021. **ACCEPTED** November 9, 2021.

INCLUDE WHEN CITING Published online January 14, 2022; DOI: 10.3171/2021.11.SPINE21949.

©AANS 2022, except where prohibited by US copyright law

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

Adult symptomatic lumbar scoliosis (ASLS), a subset of adult spinal deformity (ASD), is a debilitating disease that is increasing in prevalence as the population ages.[1–3] ASD, including ASLS, causes disability similar to other chronic conditions such as diabetes mellitus and heart disease.[4,5] Although surgery is associated with greater improvements in health-related quality of life (HRQOL) than nonoperative care, outcomes vary.[6–9] Patient-centered counseling on realistic expectations for surgery is paramount, because treatment can be painful, expensive, and risks life-threatening complications.[10–14] Patient-reported outcome measures (PROMs) are patient-focused tools that are essential in both shared decision-making and the evaluation of success in a value-driven healthcare economy.[15–20]

To date, few studies have examined the relationship between preoperative and postoperative PROMs in ASD.[21,22] Specifically, no study has used the baseline PROM to provide expected outcomes or improvements to patients after surgery. PROM change is often evaluated relative to established minimum clinically important difference (MCID) and substantial clinical benefit (SCB) thresholds for these outcome instruments. These values are fixed and are frequently applied across populations of differing disability, possibly resulting in poor classification of success and failure.[23] This is particularly true for those patients starting with little disability in some domains affected by ASLS, such as function or pain. There are probably limits to what surgery can achieve, akin to a ceiling effect of the intervention, which may cause clustering of PROM scores after surgery.

The purpose of this study was to determine whether patient PROM clustering occurs after ASLS surgery. If such a clustering effect exists, it will assist with setting expectations for patients considering surgical intervention. Furthermore, we sought to determine the effect of baseline PROMs on 2-year outcomes. Finally, we examined the perceived success of ASLS surgery by using fixed and patient-specific thresholds for MCID and SCB. We hypothesize that patients with ASLS who have varying initial disability achieve similar final HRQOL scores, given that patients with severe initial disability have a greater quantitative improvement from baseline disability than patients with mild initial disability. These analyses will assist the shared decision-making process for patients with ASLS.

Methods

Study Population

This is a secondary analysis of data collected during the ASLS-1 study.[6] The study was conducted at 9 centers in North America and included randomized and observational patient cohorts. All sites obtained IRB approval and the study was registered with the ClinicalTrials.gov website (NCT00854828). Eligible patients were 40–80 years of age with ASLS, defined as a lumbar curve with a coronal Cobb measurement ≥ 30° and Oswestry Disability Index (ODI) score ≥ 20 or Scoliosis Research Society (SRS)–22 score ≤ 4.0 in pain, function, and/or self-image domains. Patients with prior spinal fusion or multilevel decompression surgery were excluded. Enrollment began in April 2010 and ended July 2014. Patients were offered enrollment in a randomized cohort. If randomization was declined, then they were offered enrollment in an observational cohort. This is an as-treated analysis of all patients (randomized and observational) who received surgery during the conduct of the study.

Outcome Measures

The primary outcome measures were the SRS-22 subscore and the ODI. The SRS-22 is a disease-specific instrument for spinal deformity, and the ODI is specific for lumbar spine disability. The SRS-22 comprises 5 domains: pain, function, mental health, self-image, and satisfaction. Scores range from 1 to 5 (the ceiling), with higher scores indicating a better quality of life. Scores for the ODI range from 0 to 100, with higher scores indicating more disability. Given that the ODI is a disability index, a score of 0 is the ceiling for improvement. The SRS-22 and ODI scores were obtained at baseline and at 3, 12, and 24 months after surgery.

HRQOL Clustering

The 2-year postoperative SRS-22 and ODI scores were used to evaluate the extent of score clustering into quantifiable score ranges. SRS-22 and ODI histogram and kernel density distributions were plotted. PROM distributions were examined to determine the highest-density score ranges.

Outcomes by Baseline PROM

Patients were stratified into groups according to baseline disability. For ODI, 4 groups were used as previously described: minimal (0–20 disability), moderate (21–40), severe (41–60), and crippled (61–100).[24,25] An analogous stratification was used for the SRS subscore, using phrasing from the SRS-22 survey: mild (4–4.9), moderate (3–3.9), moderate to severe (2–2.9), and severe (< 2.0). Waterfall plots displayed change from baseline PROM. Stacked bar plots of 2-year PROMs by baseline disability subgroups were created. To evaluate the impact of baseline SRS subscore and ODI on 2-year outcomes, generalized linear mixed models were created. The coefficient for the baseline PROM value describes the effect of the baseline value on the estimated 2-year PROM. The 2-year outcome percentiles were calculated for SRS and ODI scores for each baseline disability group. The best-case (top 5% of PROMs), worst-case (bottom 5%), and median scenarios were determined.

Clinically Relevant Improvements

To evaluate the impact of baseline HRQOL measures on a patient's likelihood of achieving or exceeding a clinically meaningful improvement, responder analyses were performed for MCID and SCB according to baseline disability group (ODI MCID = 11, SCB = 19; and SRS MCID = 0.4, SCB = 0.7).[26–29] In addition to a fixed MCID and SCB, we examined the effect of patient-specific MCID and SCB, which are calculated as a function of the baseline score and PROM scale.[30–33] Spratt has proposed 30%

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

TABLE 1. Baseline characteristics of 171 surgically treated patients with ASLS

Pt Characteristic	Value
Age in yrs, median (IQR)	60.2 (53.4, 66.6)
Female sex, no. (%)	154 (90.1)
Race, no. (%)	
White	165 (96.5)
Black	4 (2.3)
Other	2 (1.2)
Ethnicity, no. (%)	
Hispanic	3 (1.8)
Non-Hispanic	162 (94.7)
Did not report	6 (3.5)
Education, no. (%)	
Less than high school	7 (4.1)
High school diploma or GED	44 (25.7)
Technical or associate degree	30 (17.5)
Bachelor degree	38 (22.2)
Graduate degree	52 (30.4)
Income per year, no. (%)	
<$20,000	9 (5.3)
$20,000–$39,999	22 (12.9)
$40,000–$74,999	31 (18.1)
≥$75,000	84 (49.1)
Did not report	25 (14.6)
Tobacco use, no. (%)	
Current	10 (5.8)
Former	54 (31.6)
Never	107 (62.6)
Body mass index, median (IQR)	26.2 (23.3, 29.9)
Osteopenia/osteoporosis, no. (%)	
None/does not apply	66 (38.6)
T-score –1 to –1.5	48 (28.1)
T-score –1.6 to –2.4	44 (25.7)
T-score –2.5 or worse, or vertebral compression fracture	13 (7.6)
Hypertension—uncontrolled or requiring medications, no. (%)	
No	103 (60.2)
Yes, controlled w/ diet/exercise	8 (4.7)
Yes, controlled w/ medication	60 (35.1)
Yes, poorly controlled w/ medication	0
Diabetes—uncontrolled or requiring medications, no. (%)	
No	161 (94.2)
Yes, controlled w/ diet	1 (0.6)
Yes, controlled w/ oral hypoglycemics	7 (4.1)
Yes, insulin-dependent	2 (1.2)
Depression/anxiety/psychiatric disorder, no. (%)	55 (32.2)
Duration of back Sxs in mos, median (IQR)	36 (0, 102)
Duration of leg Sxs in mos, median (IQR)	0 (0, 24)
Lumbar Cobb angle in °, median (IQR)	54 (43, 67)
Lumbar lordosis, T12–sacrum, in °, median (IQR)	–37 (–48, –25)
Sagittal balance absolute value in mm, median (IQR)	29 (12, 59)
Coronal balance absolute value in mm, median (IQR)	20 (9, 34)
Pelvic incidence–lumbar lordosis mismatch in °, median (IQR)*	19 (4, 32)
No. of stenosis levels, no. (%)	
0	159 (93)
1	11 (6.4)
2	1 (0.6)
Listhesis, no. (%)	157 (91.8)
Baseline PROMs, median (IQR)	
ODI score	36 (26, 46)
SRS subscore	3.1 (2.7, 3.5)
SRS-pain	2.8 (2.4, 3.2)
SRS-function	3.2 (2.6, 3.8)
SRS–self-image	2.8 (2.3, 3.2)
SRS–mental health	3.8 (3, 4.4)
SRS-satisfaction*	3 (2.5, 3.5)
NRS back pain	7 (5, 8)
NRS leg pain	4 (1, 6)
Mental component score	51.5 (42.1, 60)
Physical component score	32.9 (26.4, 40.2)

CONTINUED IN NEXT COLUMN »

» CONTINUED FROM PREVIOUS COLUMN

GED = general equivalency development; NRS = numeric rating scale; Pt = patient; Sxs = symptoms.
* Pelvic incidence–lumbar lordosis mismatch: 8 missing; SRS-satisfaction: 2 missing.

of the remaining scale on a PROM as the MCID.[23] For example, for a patient with a baseline SRS subscore of 4.0, the MCID is equal to 0.3, which is less than anchor-based proposed values (0.4). We propose a 50% change of the remaining scale as an SCB (e.g., SRS subscore = 4.0; SCB = 0.5).

Satisfaction With Treatment

To investigate the relationship between baseline HRQOL and 2-year satisfaction, we performed a responder analysis defining "unsatisfied" as SRS-satisfaction < 4. The median and IQR for satisfaction scores were calculated according to baseline SRS subscore disability.

Statistical Analysis

Generalized linear mixed models displayed the mean PROM scores at each follow-up time point. Corresponding baseline PROMs were covariates to estimate the effect of baseline PROM on 2-year PROM. When calculating confidence limits of the best-case (top 5%), worst-case (bottom 5%), and median outcomes, a distribution-free method was used.[34] Confidence intervals were not calculated for patient groups lacking a sufficient sample size.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

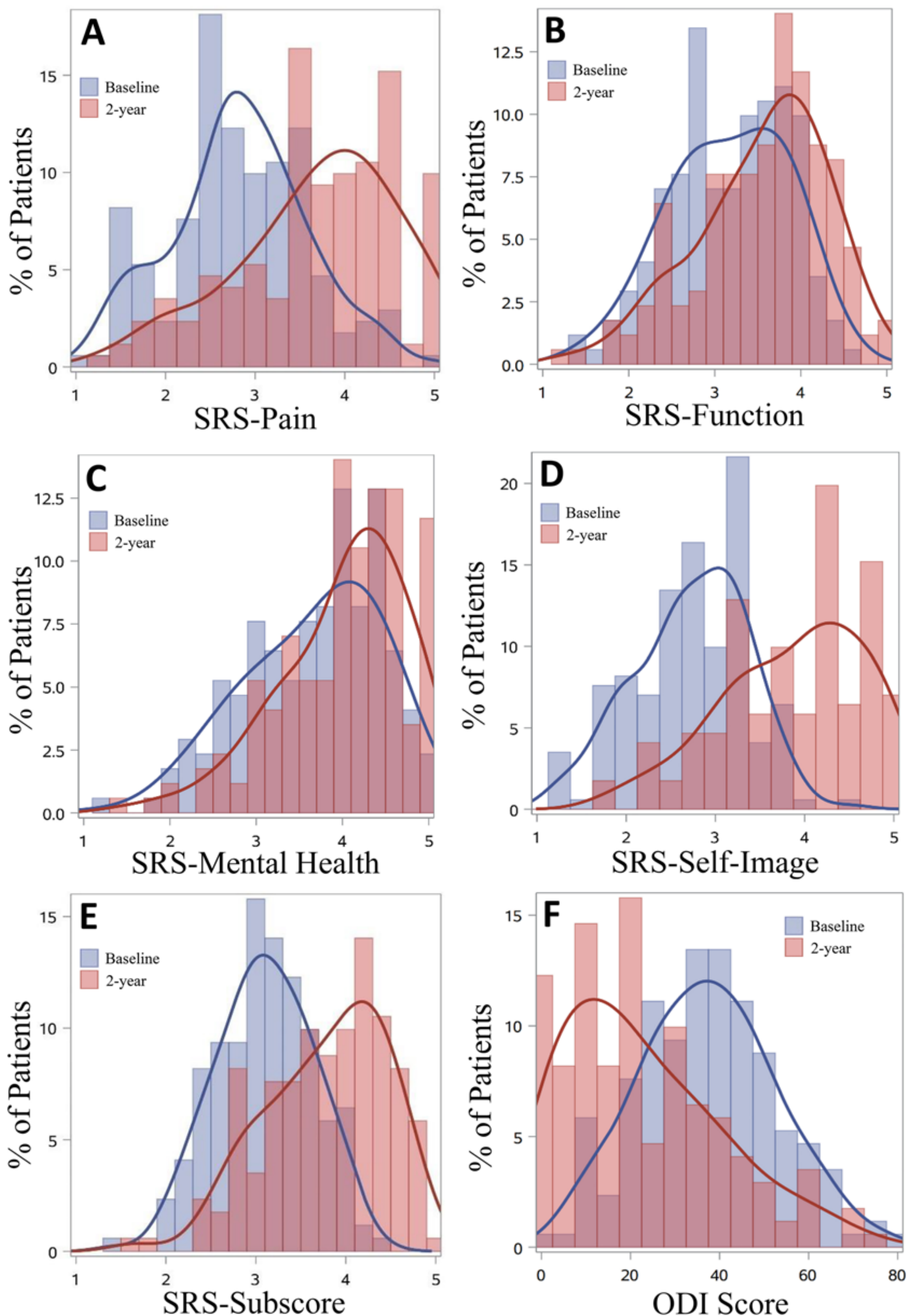

FIG. 1. Histograms of baseline and 2-year patient-reported outcomes. *Solid lines* represent kernel density estimates. **A:** Pain. **B:** Function. **C:** Mental health. **D:** Self-image. **E:** SRS subscore. **F:** ODI score.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

Classification rates for patients achieving fixed and patient-specific MCID/SCB thresholds according to baseline disability were calculated. The Cochran-Armitage test for trend examined MCID/SCB achievement across baseline disability categories. Within-category differences between fixed and patient-specific MCID/SCB thresholds were examined with McNemar's test. As this is an exploratory study, no corrections for multiple comparisons were performed. Statistical analyses were performed using SAS (version 9.4).

TABLE 2. Two-year PROM clustering ranges

Domain	Range	No.	% of Pts
Pain	3.8–4.2	51	30%
Function	3.6–4.2	74	43%
Mental health	4.0–4.6	86	50%
Self-image	4.0–4.5	55	32%
Subscore	4.0–4.5	56	33%
ODI	8–18	54	32%

Results

A total of 286 patients with ASLS were enrolled, and 187 patients received operative treatment. A total of 171 (91%) operatively treated patients achieved complete 2-year follow-up and were included in our analyses (Table 1).

Clustering

Figure 1 presents baseline and 2-year follow-up SRS-22 domain and ODI score distributions. The 2-year SRS-function (43% of patients with scores from 3.6 to 4.2) and SRS–mental health (50% of patients with scores from 4.0 to 4.6) demonstrated the highest-density clustering. SRS-pain, SRS–self-image, and SRS subscore also show clustering, with approximately one-third of patients falling between 4.0 and 4.5. Thirty-two percent of ODI scores clustered between 8 and 18. Twelve percent of patients achieved the SRS–mental health ceiling and 10% the SRS-pain ceiling. Clustering data are summarized in Table 2.

Generalized Linear Mixed Models

Better baseline PROMs were associated with higher 2-year scores for both the ODI and all SRS-22 domains. For each point improvement in ODI at baseline (decreasing score), a 0.6-point decrease is seen at 2 years. Similarly, for each point improvement in baseline SRS subscore (increasing score), a 0.6-point increase is seen at 2 years. The changes across SRS-22 domains were not substantially different (pain = 0.4; function = 0.5; self-image = 0.3; and mental health = 0.5).

PROM Change by Baseline Disability

Figure 2A and B shows waterfall plots of 2-year change in SRS subscore and ODI, respectively. Patients with the largest quantitative improvement tended to have a greater disability at baseline. General trends were similar for SRS subscore and ODI, with 61% and 59% of patients improving from baseline group at 2 years, 34% and 30% of patients remaining in their baseline group, and 5% and 11% of patients worsening from baseline group, respectively. For both SRS subscore and ODI, patients with the worst baseline disability were the most likely to see improvement (Fig. 3A and B).

Best-Case and Worst-Case Outcomes

Best-case (top 5% PROM), worst-case (bottom 5%), and median percentile outcomes are summarized in Table 3. More baseline disability is associated with more 2-year disability, despite larger quantitative improvement. Individual patient results vary greatly, even when comparing patients with similar baseline PROMs.

MCID and SCB Responder Analysis

Responder analyses for MCID and SCB are presented in Table 4. For SRS subscore, the percentage of patients achieving fixed MCID increased from 33% to 100% as baseline disability increased ($p = 0.002$). The percentage of patients achieving fixed SCB increased from 11% to 75% as baseline disability increased ($p = 0.005$). The percentages of patients achieving patient-specific SCB significantly increased with baseline disability ($p = 0.019$). A greater percentage of moderate to severely disabled patients achieved MCID with static thresholds versus patient-specific thresholds. Significant differences were not seen for other baseline disability groups. Table 5 presents satisfaction with care data. Most patients were satisfied with their outcomes, with 17% of moderately disabled and 25% of moderate to severely disabled patients expressing dissatisfaction.

For ODI, the percentage of patients achieving fixed MCID increased from 19% to 75% as baseline disability increased ($p < 0.001$). The percentage of patients achieving fixed SCB increased from 0% to 75% as baseline disability increased ($p < 0.001$). More patients with minimal baseline disability achieved patient-specific MCID (46%) than fixed MCID (19%, $p = 0.008$) and achieved patient-specific SCB (38%) than fixed SCB (0%). Patients with moderate baseline disability achieved patient-specific SCB (51%) at a higher rate than fixed SCB (41%, $p = 0.005$). For all other baseline disability groups, the number of patients achieving fixed and patient-specific MCID and SCB were not different from each other.

Discussion

PROMs are a critical component of a patient-centered shared decision-making process and are essential to a value-based healthcare system. Surgical treatment of ASLS is associated with greater improvement in PROMs when compared with nonoperative care, but comes with a substantial risk of treatment-associated adverse events.[6,8,35,36] Establishing patient preferences, values, and expected results is critical to achieving success from the patient's perspective.[37] Surgery expectations may be framed by improvement limitations inherent to the surgical intervention. That is, surgery may improve patient PROM

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

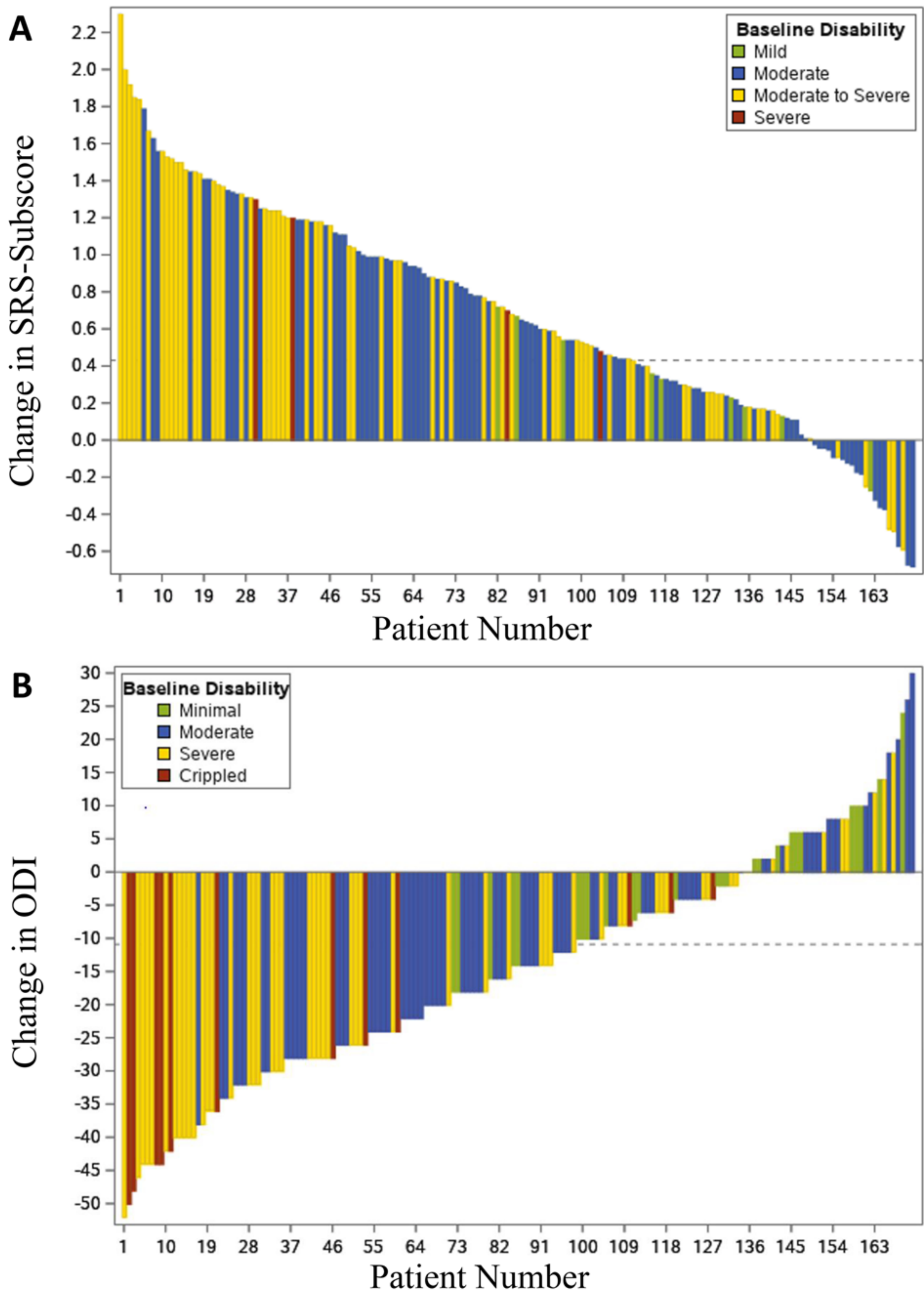

FIG. 2. Waterfall plots of individual outcomes. *Horizontal dashed lines* represent MCID. **A:** SRS subscore. **B:** ODI score.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

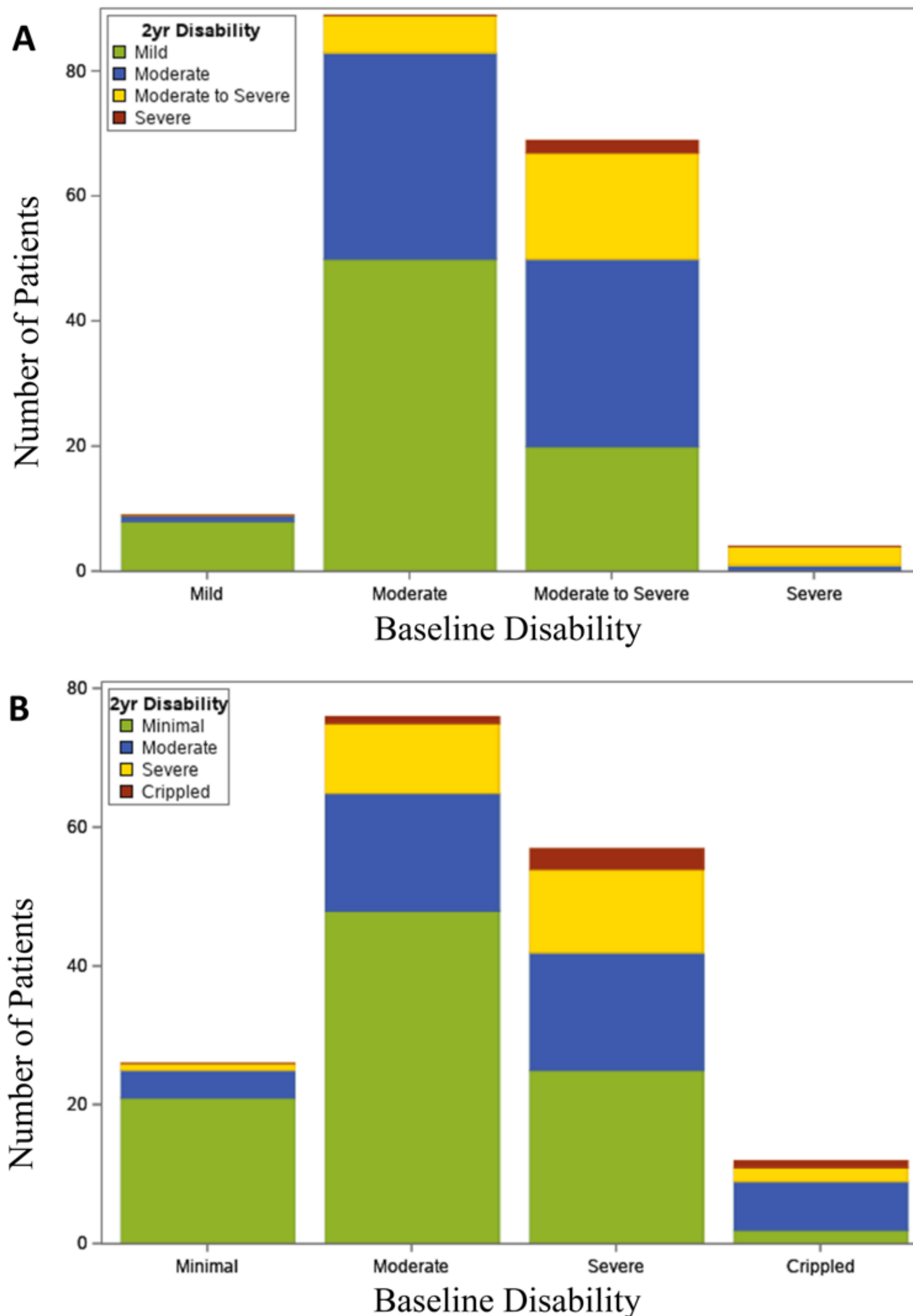

FIG. 3. Stacked bar plots of 2-year PROM scores by baseline disability groups. **A:** SRS subscore. **B:** ODI score.

scores to a limit below the true ceiling of the PROM scale due to constraints of the surgery itself. The purpose of this study was to describe PROM score clustering after surgery for ASLS and to examine the effect of initial PROM scores on 2-year outcomes.

A clustering effect of surgery was most prominent for the SRS-function domain, with 43% of 2-year scores falling between 3.6 and 4.2. The remainder of the SRS domains, subscore, and ODI had no more than one-third of patients falling within any cluster, except SRS–mental health, for which scores in half of the patients lay between 4.0 and 4.6. A clustering effect suggests that there may be limits to the function improvements possible with surgery. ASLS is frequently associated with multilevel spinal fusion including the sacrum. It is not unexpected that there are limits to what function a patient can reasonably expect after thoracolumbar fusion. The lack of other clear clusters suggests a more heterogeneous contribution of factors related to these domains, of which an extensive fusion is only one aspect.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

TABLE 3. Two-year best-case (top 5%), worst-case (bottom 5%), and most likely (median) outcomes by baseline disability

Baseline Disability	No.	Best-Case Outcome	Worst-Case Outcome	Most Likely Outcome
SRS subscore				
Mild disability (4–4.9)	9	4.8*	3.9*	4.5 [4.2–4.7]
Moderate disability (3–3.9)	89	4.8 [4.7–4.9]	2.9 [2.5–2.9]	4.1 [3.9–4.2]
Moderate to severe disability (2–2.9)	69	4.5 [4.3–4.7]	2.5 [1.6–2.5]	3.5 [3.3–3.8]
Severe disability (<2)	4	3.2*	2.4*	2.7 [2.4–3.2]
ODI				
Minimal disability (1–20)	26	0 [0–0]	34 [30–42]	9 [4–18]
Moderate disability (21–40)	76	0 [0–2]	44 [42–68]	14 [10–20]
Severely disabled (41–60)	57	0 [0–6]	62 [54–76]	26 [18–36]
Crippled (≥61)	12	18 [18–26]	68 [58–68]	32 [26–58]
Pain				
Mild disability (4–4.9)	13	5.0 [4.6–5.0]	3.0 [3.0–3.4]	4.2 [3.4–4.6]
Moderate disability (3–3.9)	64	5.0 [4.8–5.0]	2.6 [2.0–2.8]	3.9 [3.8–4.9]
Moderate to severe disability (2–2.9)	69	5.0 [5.0–5.0]	2.0 [1.8–2.2]	4.0 [3.6–4.2]
Severe disability (<2)	25	4.2 [3.8–5.0]	1.4 [1.2–1.4]	2.8 [2.5–3.4]
Function				
Mild disability (4–4.9)	27	4.8 [4.6–5.0]	3.0 [3.0–3.4]	4.2 [4.0–4.4]
Moderate disability (3–3.9)	78	4.6 [4.4–4.8]	2.4 [1.4–2.8]	3.8 [3.6–4.0]
Moderate to severe disability (2–2.9)	60	4.3 [4.0–5.0]	1.9 [1.2–2.2]	3.2 [3.0–3.6]
Severe disability (<2)	6	3.6*	1.8*	2.4 [1.8–3.6]
Mental health				
Mild disability (4–4.9)	80	5.0 [5.0–5.0]	3.4 [2.4–3.4]	4.4 [4.4–4.6]
Moderate disability (3–3.9)	60	4.8 [4.6–5.0]	2.5 [1.8–3.0]	4.0 [3.6–4.0]
Moderate to severe disability (2–2.9)	29	4.6 [4.4–5.0]	2.4 [1.4–2.6]	3.6 [3.2–4.0]
Severe disability (<2)	2	4.8*	2*	3.4 [2.0–4.8]
Self-image				
Mild disability (4–4.9)	2	4.8*	4.5*	4.7 [4.5–4.8]
Moderate disability (3–3.9)	72	5.0 [5.0–5.0]	2.8 [2.3–2.8]	4.2 [4.0–4.3]
Moderate to severe disability (2–2.9)	77	4.8 [4.8–5.0]	2.2 [1.8–2.7]	3.8 [3.3–4.2]
Severe disability (<2)	20	4.9 [4.5–5.0]	2.1 [1.8–2.7]	3.6 [3.2–4.3]

Brackets indicate 95% confidence intervals.
* Confidence interval omitted due to insufficient sample size.

Treatment success may be interpreted according to whether a change equal to or greater than the MCID is achieved, although MCID values are often fixed values with sometimes poor discriminative properties.[38,39] Considering the high cost and risk associated with surgeries for ASLS, achieving the MCID may not be an appropriate threshold for success, and an SCB has been proposed. In both cases, fixed values may both underestimate and overestimate the perceived benefit when the change relative to the PROM scale is not considered. In response, patient-specific values for MCID and SCB that consider baseline PROM and potential improvement are proposed.

Patients with the lowest baseline PROMs had, on average, the greatest improvements after surgery. Conversely, patients classified as mildly disabled were more likely to have a worsening of ODI or SRS subscore, as well as less frequent rates of MCID or SCB improvement, irrespective of fixed (ODI = 19%, SRS = 33%) or patient-specific (ODI = 46%, SRS = 56%) thresholds, underscoring the important effect of baseline PROM on expected results. With respect to SRS-pain scores, the median outcome of 3.9–4.2 suggests that many patients will improve to near mild disability. Conversely, SRS-function improvements are not nearly as uniform, with only modest improvements expected across each disability class. Not surprisingly, severely disabled patients were more likely to achieve MCID and SCB improvements when using fixed threshold values, because a small change relative to the amount of scale remaining is required. The potential value of patient-specific thresholds is seen with 30% more of the mildly disabled patients achieving SCB for both the SRS subscore and ODI, suggesting that patient-specific thresholds may be more sensitive to MCID/SCB categorization. In these less disabled patients, it may be impossible to achieve MCID/SCB thresholds due to the ceiling of the instruments used. Patient-specific threshold values also suggested slightly less success for surgery in the moderately to severely disabled patients.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

TABLE 4. Proportion of patients achieving fixed and patient-specific MCID and SCB by baseline disability

Baseline Disability	No.	MCID			SCB		
		Fixed	Pt-Specific	p Value*	Fixed	Pt-Specific	p Value*
SRS subscore							
Mild (4–4.9)	9	33%	56%	0.16	11%	44%	0.08
Moderate (3–3.9)	89	60%	58%	0.56	45%	42%	0.08
Moderate to severe (2–2.9)	69	74%	65%	0.01	58%	57%	0.56
Severe (<2)	4	100%	100%	—	75%	100%	—
Total	171	65%	62%	0.13	49%	49%	1.00
p value†		0.002	0.14		0.005	0.019	
ODI							
Minimal (1–20)	26	19%	46%	0.008	0%	38%	—
Moderate (21–40)	76	62%	66%	0.08	41%	51%	0.005
Severe (41–60)	57	65%	61%	0.16	54%	53%	0.32
Crippled (≥61)	12	75%	75%	1.00	75%	75%	—
Total	171	57%	62%	0.02	42%	51%	<0.001
p value†		<0.001	0.18		<0.001	0.068	

— = p value omitted due to insufficient sample size or sampling zero.
* According to McNemar's test.
† According to the Cochran-Armitage test for trend.

In general, worse baseline health scores predict more quantitative PROM improvement after surgery for ASD.[9,40] As such, patients with worse baseline PROMs are more likely to achieve MCID and SCB improvements. Fixed MCID values are limited by their relatively poor discrimination and by their poor sensitivity for patients with better baseline PROM scores.[23] Individualized MCID and SCB values may offer an improvement over more anchor-based or distribution-based fixed values. Our analysis demonstrated that patient-specific thresholds for MCID/SCB can improve sensitivity for deeming patients with minimal initial disability clinically improved, while reducing the potentially inflated clinical improvements seen in patients with moderate to severe initial disability.

In the ASLS-1 study there were significant PROM improvements on average.[6] The use of averages fails to account for patient-specific differences such as baseline PROM, which we investigate here, and the data presented may help set expectations for patients with ASLS. Clinical use of baseline PROM scores will continue to increase as machine learning methodologies are being used to develop models for predicting surgical outcomes for patients with ASD from baseline measurements.[41–44] Together, personalized MCID/SCB values and personalized machine learning predictions for achieving these values will further enhance patients' individualized surgical counseling and expectation management. Until personalized machine learning–based tools are available, the use of best-case/worst-case/most likely outcome may assist shared decision-making.

This study is limited by its retrospective design, although the data come from a prospective, dual-arm study with high rates of patient retention. An anchor question to ask whether patients feel improved or substantially improved would be most useful for examining classification of success by MCID and SCB. This is particularly true for the so-called 50% rule proposed for SCB. We have chosen this value because it is approximately 50% more than the MCID value. In general, extreme changes in PROMs (twice MCID) are uncommon and a value between MCID and twice MCID may be a fair target. Patient expectations for pain relief, functional improvement, and improved cosmesis will certainly vary, and determining patient expectations according to the various domains affected by ASLS is necessary. In this exploratory analysis we have relied on the primary outcome measures of the ASLS-1 study, which were the ODI and the SRS subscore. We have not analyzed the effect of complications on these

TABLE 5. Patient-reported satisfaction at 2 years by baseline disability

Baseline Disability	No.	Median Satisfaction	IQR	Surgery Success	Surgery Failure
SRS subscore					
Mild disability (4–4.9)	9	4.5	4.0–5.0	8 (89%)	1 (11%)
Moderate disability (3–3.9)	89	4.4	4.0–5.0	74 (83%)	15 (17%)
Moderate to severe disability (2–2.9)	69	4.2	3.8–5.0	52 (75%)	17 (25%)
Severe disability (<2)	4	4.5	4.1–4.9	4 (100%)	0 (0%)

Scores ≥ 4.0 were considered successful, and scores < 4.0 were considered failures.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

outcomes, given that we sought to use baseline PROMs to guide the informed–decision-making process. Analyses of complications within this data set suggest that single, recoverable complications have minimal effect on outcomes, as opposed to permanent deficits or repeated complications. Similarly, we have not controlled for the effect of postoperative alignment. Sagittal plane malalignment at baseline is not common in this data set and, more importantly, sagittal plane alignment after surgery may have only a small effect on postoperative PROMs.[41] Finally, this cohort is composed of only patients with ASLS, a distinct subset of ASD. Our results may not extrapolate to all ASD populations, given that there are quantifiable differences in postoperative PROM domain score improvements according to deformity type.[45]

This study has direct clinical implications regarding surgical outcomes based on disability at the time of surgery. Our finding that 2-year patient-reported outcomes cluster into definable ranges shows that many patients achieve similar final outcomes after operative treatment regardless of initial disability. With the significant cost, pain, and risk associated with operative treatment, it may be tempting for patients and surgeons to wait until the scoliosis becomes crippling before pursuing this option. However, this delay and extended suffering may not be prudent, given that many patients achieve similar 2-year PROMs regardless of starting point. Patients with great initial disability see the largest improvements from the starting point but remain more disabled than their counterparts who had less initial disability. Our best-case, worst-case, and most likely case analysis is useful in providing patients with accurate data on the heterogeneity of possible outcomes. We conclude that presurgical PROMs should be considered when predicting the expected benefit from surgical treatment of ASLS, and they should be directly applied to patient-centered, informed–decision-making tools. Future work will focus on applying PROM data to developing tools to facilitate discussions with patients regarding the expected benefits from surgical intervention.

Conclusions

PROMs are a critical part of describing patient disability before and after surgery. Clustering of outcomes after surgery for ASLS may help set expectations for patients as surgeons inform them of possible and most likely results. Patients with varying initial disability often achieve good outcomes, given that more severe initial disability corresponds to greater quantitative improvement from baseline. Although average patient-reported outcomes are essential, this study provides additional guidance to surgeons for the shared decision-making process. Best-case and worst-case outcome information highlights the heterogeneity of possible outcomes and helps frame patient expectations. Dynamic MCID and SCB values may offer more sensitive criteria to define surgical success or failure.

Acknowledgments

Funding was provided by the NIH (R01AR05517601A2) and the SRS.

References

1. Ames CP, Scheer JK, Lafage V, et al. Adult spinal deformity: epidemiology, health impact, evaluation, and management. *Spine Deform*. 2016;4(4):310-322.
2. Passias PG, Jalai CM, Worley N, et al. Adult spinal deformity: national trends in the presentation, treatment, and perioperative outcomes from 2003 to 2010. *Spine Deform*. 2017;5(5):342-350.
3. Kebaish KM, Neubauer PR, Voros GD, Khoshnevisan MA, Skolasky RL. Scoliosis in adults aged forty years and older: prevalence and relationship to age, race, and gender. *Spine (Phila Pa 1976)*. 2011;36(9):731-736.
4. Bess S, Line B, Fu KM, et al. The health impact of symptomatic adult spinal deformity: comparison of deformity types to united states population norms and chronic diseases. *Spine (Phila Pa 1976)*. 2016;41(3):224-233.
5. Pellisé F, Vila-Casademunt A, Ferrer M, et al. Impact on health related quality of life of adult spinal deformity (ASD) compared with other chronic conditions. *Eur Spine J*. 2015; 24(1):3-11.
6. Kelly MP, Lurie JD, Yanik EL, et al. Operative versus nonoperative treatment for adult symptomatic lumbar scoliosis. *J Bone Joint Surg Am*. 2019;101(4):338-352.
7. Liu S, Schwab F, Smith JS, et al. Likelihood of reaching minimal clinically important difference in adult spinal deformity: a comparison of operative and nonoperative treatment. *Ochsner J*. 2014;14(1):67-77.
8. Bridwell KH, Glassman S, Horton W, et al. Does treatment (nonoperative and operative) improve the two-year quality of life in patients with adult symptomatic lumbar scoliosis: a prospective multicenter evidence-based medicine study. *Spine (Phila Pa 1976)*. 2009;34(20):2171-2178.
9. Moal B, Lafage V, Smith JS, et al. Clinical improvement through surgery for adult spinal deformity: what can be expected and who is likely to benefit most? *Spine Deform*. 2015;3(6):566-574.
10. McCarthy IM, Hostin RA, O'Brien MF, et al. Analysis of the direct cost of surgery for four diagnostic categories of adult spinal deformity. *Spine J*. 2013;13(12):1843-1848.
11. McCarthy IM, Hostin RA, Ames CP, et al. Total hospital costs of surgical treatment for adult spinal deformity: an extended follow-up study. *Spine J*. 2014;14(10):2326-2333.
12. Smith HE, Rihn JA, Brodke DS, et al. Spine care: evaluation of the efficacy and cost of emerging technology. *Am J Med Qual*. 2009;24(6)(suppl):25S-31S.
13. Smith JS, Shaffrey CI, Kelly MP, et al. Effect of serious adverse events on health-related quality of life measures following surgery for adult symptomatic lumbar scoliosis. *Spine (Phila Pa 1976)*. 2019;44(17):1211-1219.
14. Carreon LY, Glassman SD, Lurie J, et al. Cost-effectiveness of operative versus nonoperative treatment of adult symptomatic lumbar scoliosis an intent-to-treat analysis at 5-year follow-up. *Spine (Phila Pa 1976)*. 2019;44(21):1499-1506.
15. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. Can preoperative patient-reported outcome measures be used to predict meaningful improvement in function after TKA? *Clin Orthop Relat Res*. 2017;475(1):149-157.
16. Berliner JL, Brodke DJ, Chan V, SooHoo NF, Bozic KJ. John Charnley Award: Preoperative patient-reported outcome measures predict clinically meaningful improvement in function after THA. *Clin Orthop Relat Res*. 2016;474(2):321-329.
17. Rihn JA, Currier BL, Phillips FM, Glassman SD, Albert TJ. Defining the value of spine care. *J Am Acad Orthop Surg*. 2013;21(7):419-426.
18. Porter ME. What is value in health care? *N Engl J Med*. 2010; 363(26):2477-2481.
19. Noonan VK, Lyddiatt A, Ware P, et al. Montreal Accord on Patient-Reported Outcomes (PROs) use series—Paper 3: Patient-reported outcomes can facilitate shared decision-

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

making and guide self-management. *J Clin Epidemiol*. 2017;89:125-135.
20. Barry MJ, Edgman-Levitan S. Shared decision making—pinnacle of patient-centered care. *N Engl J Med*. 2012;366(9):780-781.
21. Kang DG, Baldus C, Glassman SD, Shaffrey CI, Lurie JD, Bridwell KH. Neurologic deficits have a negative impact on patient-related outcomes in primary presentation adult symptomatic lumbar scoliosis surgical treatment at one-year follow-up. *Spine (Phila Pa 1976)*. 2017;42(7):479-489.
22. Chapman TM Jr, Baldus CR, Lurie JD, et al. Baseline patient-reported outcomes correlate weakly with radiographic parameters: a multicenter, prospective NIH adult symptomatic lumbar scoliosis study of 286 patients. *Spine (Phila Pa 1976)*. 2016;41(22):1701-1708.
23. Spratt KF. Patient-level minimal clinically important difference based on clinical judgment and minimally detectable measurement difference: a rationale for the SF-36 physical function scale in the SPORT intervertebral disc herniation cohort. *Spine (Phila Pa 1976)*. 2009;34(16):1722-1731.
24. Fairbank JCT, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)*. 2000;25(22):2940-2952.
25. Breitenseher MJ, Eyb RP, Matzner MP, Trattnig S, Kainberger FM, Imhof H. MRI of unfused lumbar segments after spondylodesis. *J Comput Assist Tomogr*. 1996;20(4):583-587.
26. Crawford CH III, Glassman SD, Bridwell KH, Berven SH, Carreon LY. The minimum clinically important difference in SRS-22R total score, appearance, activity and pain domains after surgical treatment of adult spinal deformity. *Spine (Phila Pa 1976)*. 2015;40(6):377-381.
27. Hägg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J*. 2003;12(1):12-20.
28. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. *J Bone Joint Surg Am*. 2008;90(9):1839-1847.
29. Crawford CH III, Glassman SD, Bridwell KH, Carreon LY. The substantial clinical benefit threshold for SRS-22R domains after surgical treatment of adult spinal deformity. *Spine Deform*. 2016;4(5):373-377.
30. Little DG, MacDonald D. The use of the percentage change in Oswestry Disability Index score as an outcome measure in lumbar spinal surgery. *Spine (Phila Pa 1976)*. 1994;19(19):2139-2143.
31. Austevoll IM, Gjestad R, Grotle M, et al. Follow-up score, change score or percentage change score for determining clinical important outcome following surgery? An observational study from the Norwegian registry for spine surgery evaluating patient reported outcome measures in lumbar spinal stenosis and lumbar degenerative spondylolisthesis. *BMC Musculoskelet Disord*. 2019;20(1):31.
32. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1-2):9-19.
33. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. 2008;33(1):90-94.
34. Hahn GJ, Meeker WQ. *Statistical Intervals: A Guide for Practitioners*. John Wiley & Sons; 1991.
35. Acaroglu E, Yavuz AC, Guler UO, et al. A decision analysis to identify the ideal treatment for adult spinal deformity: is surgery better than non-surgical treatment in improving health-related quality of life and decreasing the disease burden? *Eur Spine J*. 2016;25(8):2390-2400.
36. Smith JS, Lafage V, Shaffrey CI, et al. Outcomes of operative and nonoperative treatment for adult spinal deformity: a prospective, multicenter, propensity-matched cohort assessment with minimum 2-year follow-up. *Neurosurgery*. 2016;78(6):851-861.
37. Whitebird RR, Solberg LI, Norton CK, Ziegenfuss JY, Asche SE, Grossman ES. What outcomes matter to patients after joint or spine surgery? *J Patient Cent Res Rev*. 2020;7(2):157-164.
38. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7(5):541-546.
39. Chung AS, Copay AG, Olmscheid N, Campbell D, Walker JB, Chutkan N. Minimum clinically important difference: current trends in the spine literature. *Spine (Phila Pa 1976)*. 2017;42(14):1096-1105.
40. Hey HWD, Luo N, Chin SY, et al. The predictive value of preoperative health-related quality-of-life scores on postoperative patient-reported outcome scores in lumbar spine surgery. *Global Spine J*. 2018;8(2):156-163.
41. Ames CP, Smith JS, Pellisé F, et al. Development of deployable predictive models for minimal clinically important difference achievement across the commonly used health-related quality of life instruments in adult spinal deformity surgery. *Spine (Phila Pa 1976)*. 2019;44(16):1144-1153.
42. Khor S, Lavallee D, Cizik AM, et al. Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. *JAMA Surg*. 2018;153(7):634-642.
43. McGirt MJ, Sivaganesan A, Asher AL, Devin CJ. Prediction model for outcome after low-back surgery: individualized likelihood of complication, hospital readmission, return to work, and 12-month improvement in functional disability. *Neurosurg Focus*. 2015;39(6):E13.
44. Pellisé F, Serra-Burriel M, Smith JS, et al. Development and validation of risk stratification models for adult spinal deformity surgery. *J Neurosurg Spine*. 2019;31(4):587-599.
45. Line B, Bess S, Lafage V, et al. Counseling guidelines for anticipated postsurgical improvements in pain, function, mental health, and self-image for different types of adult spinal deformity. *Spine (Phila Pa 1976)*. 2020;45(16):1118-1127.

Disclosures

Dr. Smith is a consultant for Zimmer Biomet, NuVasive, Carlsmed, Cerapedics, DePuy Synthes, SeaSpine, and Stryker. He reports direct stock ownership in NuVasive and Alphatec. He received clinical or research support for the study described (includes equipment or material) from DePuy Synthes/ISSGF, and he received support of a non–study-related clinical or research effort that he oversaw from DePuy Synthes/ISSGF, NuVasive, and AO Spine. He receives royalties from Zimmer Biomet, NuVasive, and Thieme. Dr. Bess is a consultant for Stryker, Mirus, and Atec. He is a patent holder with Stryker and NuVasive. He received clinical or research support for the study described (includes equipment or material) from ISSFG, Medtronic, Globus, Stryker, Zimmer, and NuVasive. He received support of a non–study-related clinical or research effort that he oversaw from ISSGF, NuVasive, Medtronic, Stryker, Globus, SI Bone, and SeaSpine. He receives royalties from Stryker and NuVasive. Dr. Shaffrey is a consultant for Medtronic, NuVasive, and SI Bone. He reports direct stock ownership in NuVasive. He is a patent holder with Medtronic, NuVasive, Zimmer Biomet, and SI Bone. He receives royalties from Medtronic, NuVasive, and SI Bone. Dr. Lenke reports the following. 1) Medtronic: paid consultant (monies donated to a charitable foundation), royalties. 2) Broadwater: reimbursement for airfare/hotel. 3) Scoliosis Research Society: reimbursement for airfare/hotel, grant support (monies to institution). 4) Setting Scoliosis Straight Foundation: grant support (monies to institution). 5) Evans Family donation: philanthropic research funding from grateful patient/family. 6) Fox Family

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC

Foundation: philanthropic research funding from grateful patient. 7) AO Spine: reimbursement for airfare/hotel, grant support (monies to institution), and fellowship support to institution. 8) Acuity Surgical: paid consultant (monies donated to charitable foundation), royalties. 9) Abryx: paid consultant (monies donated to charitable foundation).

Author Contributions

Conception and design: Kelly, Yanik, Shaffrey. Acquisition of data: Kelly, Smith, Shaffrey, Lenke, Bridwell. Analysis and interpretation of data: Kelly, Wondra, Smith, Bess, Shaffrey, Lenke. Drafting the article: Kelly, Wondra, Smith, Bess. Critically revising the article: Kelly, Greenberg, Smith, Bess, Shaffrey, Lenke, Bridwell. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Kelly. Statistical analysis: Kelly, Yanik, Greenberg.

Supplemental Information

Previous Presentations

Podium presentation at the Scoliosis Research Society's 56th annual meeting, September 23, 2021, in St. Louis, MO.

Correspondence

Michael P. Kelly: Washington University School of Medicine, St. Louis, MO. kellymi@wustl.edu.

Brought to you by Duke University | Unauthenticated | Downloaded 06/16/23 03:19 PM UTC