



## Clinical Improvement Through Surgery for Adult Spinal Deformity: What Can Be Expected and Who Is Likely to Benefit Most ?

Bertrand Moal, Virginie Lafage, Justin.S Smith, Christopher B. Ames, Gregory Mundis, Jamie S. Terran, Eric Klineberg, Robert Hart, Oheneba Boachie-Adjei, Bess Shay, et al.

### ► To cite this version:

Bertrand Moal, Virginie Lafage, Justin.S Smith, Christopher B. Ames, Gregory Mundis, et al.. Clinical Improvement Through Surgery for Adult Spinal Deformity: What Can Be Expected and Who Is Likely to Benefit Most ?. Spine Deformity, 2015, 3 (6), pp.566-574. 10.1016/j.jspd.2015.04.004 . hal-02479172

**HAL Id: hal-02479172**

**<https://hal.science/hal-02479172>**

Submitted on 14 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Clinical Improvement Through Surgery for Adult Spinal Deformity: What Can Be Expected and Who Is Likely to Benefit Most?

Bertrand Moal, PhD<sup>a,i</sup>, Virginie Lafage, PhD<sup>a,\*</sup>, Justin S. Smith, MD, PhD<sup>b</sup>, Christopher P. Ames, MD<sup>c</sup>, Gregory Mundis, MD<sup>d</sup>, Jamie S. Terran, BS<sup>a</sup>, Eric Klineberg, MD<sup>e</sup>, Robert Hart, MD<sup>f</sup>, Oheneba Boachie-Adjei, MD<sup>g</sup>, Shay Bess, MD<sup>h</sup>, Wafa Skalli, PhD<sup>i</sup>, Frank Schwab, MD<sup>a</sup>, International Spine Study Group

<sup>a</sup>Department of Orthopedic Surgery, NYU Hospital for Joint Diseases, New York, NY, 10003, USA

<sup>b</sup>Department of Neurosurgery, University of Virginia Medical Center, Charlottesville, VA, 22903, USA

<sup>c</sup>Department of Neurosurgery, University of California San Francisco, San Francisco, CA, 94143, USA

<sup>d</sup>Department of Orthopedic Surgery, San Diego Center for Spinal Disorders, La Jolla, CA, 92037, USA

<sup>e</sup>Department of Orthopedic Surgery, University of California Davis, Sacramento, CA, 95817, USA

<sup>f</sup>Department of Orthopedic Surgery, Oregon Health Sciences University, Portland, OR, 97239, USA

<sup>g</sup>Department of Orthopedic Surgery, Hospital for Special Surgery, New York, NY, 10021, USA

<sup>h</sup>Department of Orthopedic Surgery, Rocky Mountain Hospital for Children, Denver, CO, 80205, USA

<sup>i</sup>Laboratory of Biomechanics, Arts et Metier ParisTech, Paris, France

## Abstract

**Study Design:** Multicenter, prospective, nonconsecutive, surgical case series from the International Spine Study Group.

**Objectives:** To evaluate the extent of clinical improvement after surgery for adult spinal deformity (ASD) based on minimal clinically important difference (MCID) and baseline measures.

**Summary of Background Data:** For ASD, evaluation of surgical treatment success using clinical scores should take into account baseline disability and pain and the improvement defined relative to the MCID.

**Methods:** Inclusion criteria included operative patients (age > 18 years) with baseline and 2-year SRS-22 scores. Normative values for the SRS scores were included and improvement for patients was expressed in number of MCIDs. At baseline, patients were classified by differences in activity and pain scores from normative values in four groups: “worst,” “severe,” “poor,” and “moderate.” At 2 years after surgery, patients were classified into four groups based on their change in SRS score as follows: “no improvement or deterioration,” “mediocre,” “satisfactory,” or “optimal.” Distinction among curve types was also performed based on the SRS-Schwab ASD classification.

**Results:** A total of 223 patients (age = 55 ± 15 years) were included. At baseline, for 77% of the patients, the worst scores were in Activity or Pain. At baseline, the distribution was 36% “worst,” 28% “severe,” 19% “poor,” and 17% “moderate.” Patients with sagittal malalignment only were more likely to be in the “worst” state (54%). The overall distribution of improvement was as follows: 24% no improvement or deterioration, 17% mediocre, 25% satisfactory, and 33% optimal. Forty-one percent of baseline “moderate” patients achieved no improvement. Of the baseline “worst” patients, 20% achieved no improvement, and 36% and 19% achieved “satisfactory” and “optimal” improvement, respectively.

**Conclusion:** Overall, 24% of patients did not experience improvement after surgery. Patients with baseline severe disability were more likely to perceive improvement than patients with less disability.

**Level of Evidence:** Level II.

**Keywords:** Adult spinal deformity; Clinical outcomes; Normative data; Sagittal alignment; MCID

## Introduction

To standardize the evaluation of adult spinal deformity (ASD), health-related quality of life (HRQOL) instruments are now widely used in clinical practice and the scientific community [1-3]. The Scoliosis Research Society instruments (SRS-22) were validated to provide a disease-specific health questionnaire [4].

Although its use is now common, the interpretation of HRQOL scores changes after treatment and involves considerations beyond simple numerical improvement [5]. Literature on HRQOLs advocates for the use of the minimal clinically important difference (MCID) concept in order to differentiate a statistical improvement from a clinical one, perceivable by the patient [5-9]. Although the use of MCID can enhance clinical relevance, it does not take into account the absolute values of HRQOL scores. Two patients experiencing the same improvement may not have the same outcome if they started at different baseline scores.

Moreover if part of those differences in baseline score is relative to the deficit due to the deformity, there is also an intrinsic clinical difference due to patient age. In a study on normative data, Baldus et al. [10] reported significant differences for SRS scores among sex and age groups. When evaluating the clinical improvement of ASD population, the age range is in general large (18 to >80 years old) [3,11-15], and therefore HRQOL evaluation should account for normative data instead of solely relying on generic scales.

The present study aimed to assess clinical outcomes for ASD treated surgically, with an emphasis on the baseline evaluation to include initial clinical state, age, body mass index (BMI), history of previous surgery, comorbidities, and type of spinal deformity. In contrast to previous studies, the outcomes assessment is based on comparisons with normative data, matched on the basis of age and gender, reported as multiples of MCID.

## Material and Methods

This is a retrospective analysis of a nonconsecutive series of ASD patients (age > 18 years) enrolled in a prospective multicenter study. Patients were drawn from the International Spine Study Group (ISSG) prospective database, derived from 10 clinical sites across the United States. Patients were enrolled through an institutional review board-approved protocol by each site. The radiographic inclusion criterion for the ISSG database was at least one of the following: Cobb angle  $\geq 20^\circ$ , sagittal vertical axis  $\geq 5$  cm, pelvic tilt  $\geq 25^\circ$ , or thoracic kyphosis  $\geq 60^\circ$ . Patients with inflammatory arthritis, tumor, or neuromuscular disease were excluded.

Specific inclusion criteria for the present study included operative treatment, availability of SRS-22 scores, and

availability of X-ray films at baseline and at the 2-year follow-up. Age, BMI, medical history, and comorbidities (Charlson score [16]) were collected. X-ray films were analyzed at baseline using validated software [17,18] (Spineview; Laboratory of Biomechanics Arts et Metiers ParisTech, Paris) to obtain the following parameters: Cobb angle and apex location, pelvic tilt, pelvic incidence (PI), L1-S1 lumbar lordosis, pelvic incidence—lumbar lordosis mismatch, and sagittal vertical axis. Each patient was

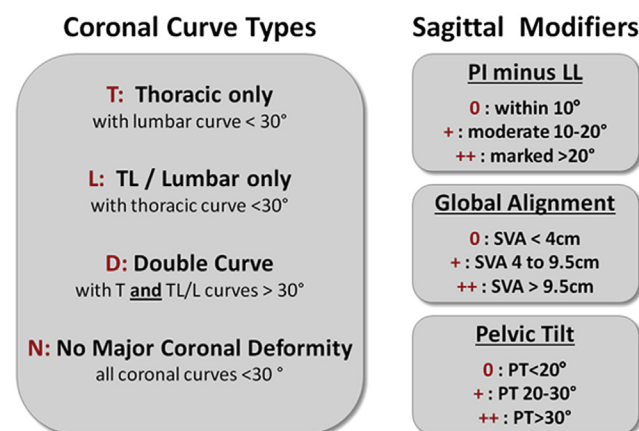


Fig. 1. Scoliosis Research Society (SRS)-Schwab classification.

Table 1

Description of the different types of curves derived from the SRS-Schwab classification for adult spinal deformity.

Type of curve groups	Acronym	Coronal criteria	Sagittal criteria
Thoracic	T	Type T	All modifiers at grade 0
Thoracic/sagittal	TS	Type T	At least one modifier at grade + or ++
(Thoraco)-lumbar	L	Type L	All modifiers at grade 0
(Thoraco)-lumbar/sagittal	LS	Type L	At least one modifier at grade + or ++
Double	D	Type D	All modifiers at grade 0
Double sagittal	DS	Type D	At least one modifier at grade + or ++
Sagittal only	S	Type N	At least one modifier at grade + or ++
Unclassified	U	Type N	All modifiers at grade 0

SRS, Scoliosis Research Society.

Table 2

Minimum clinically important difference (MCID) values used in this study.

SRS-22 domain	MCID
SRS Pain	+0.587 points/5
SRS Appearance	+0.8 points/5
SRS Activity	+0.375 points/5
SRS Mental	+0.42 points/5

SRS, Scoliosis Research Society.

Data from S. Berven, V. Deviren, D. Polly et al., presented at the International Meeting on Advanced Spine Techniques, Banff, Canada, July 7–9, 2005

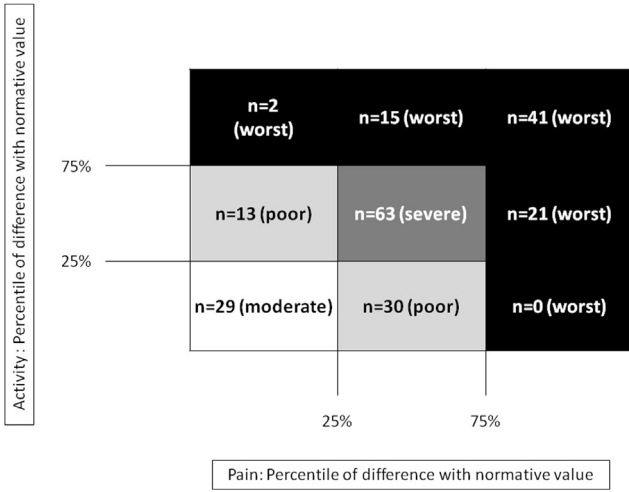


Fig. 2. Classification of baseline clinical state based on Pain and Activity SRS-22 domains.

classified according to the different type of curves derived from the SRS-Schwab classification (Fig. 1) [7,11,19,20], distinguishing eight groups based on deformity pattern (Table 1).

#### Clinical classifications

At baseline and 2-year follow-up, Activity, Pain, Appearance, and Mental domains of SRS-22 were

expressed as differences from normative data (age and gender matched) [3]. For each patient the differences were normalized using thresholds of MCIDs (Table 2, Equation 1). Because normative data and MCID are not available for the SRS Satisfaction domain and the Total SRS score, these scores were not included in the study.

$$\text{Diff}_{\text{norm domain}}^{\text{patient}} = \frac{\text{SRS}_{\text{domain}}^{\text{patients}} - \text{SRS}_{\text{domain}}^{\text{normative}}}{\text{MCID}_{\text{domain}}} \quad (\text{Equation 1})$$

For instance, a  $\text{Diff}_{\text{norm}}$  of  $-3$  MCID can be interpreted as a negative difference of 3 MCID compared with the SRS score of a normative population (age and gender matched).

For each of the four SRS domains each patient was classified into one of five groups reflecting clinical state at baseline and 2 years after surgery. Patients were then evaluated with the  $\text{Diff}_{\text{norm}}$  and change from baseline to the 2-year follow-up as follows:

Consistent:  $\text{Diff}_{\text{norm}}$  less than 1 MCID at baseline and at 2-year follow-up.

Deterioration: deterioration between baseline and 2-year follow-up greater than 1 MCID.

No improvement:  $\text{Diff}_{\text{norm}}$  at baseline greater than 1 MCID and improvement at 2-year follow-up less than 1 MCID.

		SRS Pain: surgical treatment effectiveness at 2 years				
		Consistent	Deterioration	No improvement	Improvement	Excellent
SRS Activity: surgical treatment effectiveness at 2 years	Consistent	9	0	4	3	13
	Deterioration	3	8	13	2	3
	No improvement	1	4	12	12	4
	Improvement	1	1	17	24	12
	Excellent	2	1	4	14	56

	Optimal		Satisfactory
	Mediocre		No improvement or deterioration

Fig. 3. Stratification of global clinical effectiveness of surgical treatment for adult spinal deformity at 2-year follow-up based on SRS Activity and SRS Pain domains.

Improvement: Diff<sub>norm</sub> at baseline greater than 1 MCID, improvement greater than 1 MCID, and Diff<sub>norm</sub> at 2-year follow-up greater than 1 MCID.

Excellent: Diff<sub>norm</sub> at baseline greater than 1 MCID, improvement greater than 1 MCID, and Diff<sub>norm</sub> at 2-year follow-up less than 1 MCID.

In addition, the results of the pain and activity domains were combined to classify the clinical condition of patients at baseline and the overall clinical treatment effectiveness.

At baseline, four groups were defined using the distribution of the Diff<sub>norm</sub> as illustrated in Fig. 2:

Worst: Diff<sub>norm</sub> above the 75th percentile (ie, greater difference) for Activity and Pain domains.

Severe: Diff<sub>norm</sub> between the 25th percentile and 75th percentile for both Activity and Pain domains.

Poor: Diff<sub>norm</sub> between the 25th percentile and 75th percentile for one of the domains and less than the 25th percentile for the other domain.

Moderate: Diff<sub>norm</sub> below the 25th percentile for both domains

To evaluate the overall treatment effectiveness, clinical scores in Pain and Activity were combined to establish the following four groups (Fig. 3):

Optimal: *excellent* in at least one domain + *consistent or excellent* in the other domain.

Satisfactory: *improvement* in at least one domain + *consistent or excellent* in the other.

Mediocre: *no improvement* in one domain + *improvement or excellent* in the other.

No improvement or deterioration: *deterioration* OR *no improvement* in at least one domain + *consistent or no improvement* in the other.

#### Statistical analysis

HRQOL scores were summarized by means and standard deviations (SDs). The distributions of patients according to the clinical treatment classification were expressed. Comparisons across the baseline clinical classification and overall clinical treatment effectiveness classification were performed using ANOVA t-test with Bonferroni post hoc or Games-Howell post hoc when equal variance was not assumed. Cross comparison between baseline clinical classification and overall clinical treatment effectiveness classification was studied using a chi-square test. The significance level was set at 0.05.

Table 3

Patients enrolled across participating sites.

Number of patients	Total	Site										
		A	B	C	D	E	F	G	H	J	K	
Enrolled between 2008 and 2011	327	6	15	34	54	25	14	18	62	59	40	
Enrolled with adequate baseline radiographs and SRS scores	310	6	15	34	47	25	14	17	54	58	40	
Enrolled with adequate baseline and 2-year follow-up radiographs and SRS scores	223	6	13	29	39	14	6	10	27	47	32	

SRS, Scoliosis Research Society.

## Results

### Enrollment

Between January 1, 2008, and December 31, 2011, 689 patients were eligible and 327 were enrolled. A total of 310 patients had adequate baseline radiographs and SRS scores. Out of these patients, 223 (35 men and 188 women) had adequate baseline and 2-year follow-up radiographs and SRS scores and thus were the only ones included in this study (Table 3). Mean age was 55 years (SD = 15) and mean BMI was 27.3 (SD = 5.9). Forty-two percent of the patients had a prior history of spine surgery.

The distribution of patients by curve type at baseline (Table 4) identified that 75% of patients had a coronal Cobb angle greater than 30° (T, TS, L, LS, D, and DS) and 73% had a sagittal deformity (TS, LS, DS, and TS). A total of 7 patients (3%) were unclassifiable because of a difference between the inclusion criteria for the database (Cobb > 20°) and the SRS-Schwab classification (Cobb > 30°). These patients were removed from the following analysis.

### Analysis by domains

At baseline, significantly larger Diff<sub>norm</sub> values were found for Pain (−3.08 ± 1.44 MCID) and Activity (−3.27 ± 2.38 MCID) domains compared with Appearance (−2.14 ± 0.96 MCID) and Mental (−1.54 ± 2.24 MCID) domains (p < .001) (Table 5). The principal domains of complaint,

Table 4

Distribution of patients by curve type at baseline.

Curve type	Number	Percentage
U	7	3
T	10	4
L	22	10
D	21	9
TS	5	2
LS	59	26
DS	51	23
S	48	22

D, double; DS, double sagittal; L, (thoraco)-lumbar; LS, (thoraco)-lumbar/sagittal; S, sagittal only; T, thoracic; TS, thoracic/sagittal; U, unclassified.

Table 5

Difference by domains with normative values ( $\text{Diff}_{\text{norm}}$ ) expressed in minimum clinically important difference (MCID) for the entire population, at baseline, 2-year follow-up, and in terms of improvement.

	SRS score							
	Activity		Pain		Appearance		Mental	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
$\text{Diff}_{\text{norm}}$								
Preoperation	-3.27	2.38	-3.08	1.44	-2.14	0.96	-1.54	2.24
2-year follow-up	-1.64	2.75	-1.44	1.89	-0.56	1.15	-0.58	2.19
Improvement in MCID	1.63	2.26	1.64	1.84	1.58	1.23	0.97	1.98

SRS, Scoliosis Research Society.

defined as the domain presenting the largest  $\text{Diff}_{\text{norm}}$ , were Activity (43% of the population) and Pain (34%).

Postoperatively, the mean MCID differences were  $-1.64 \pm 2.75$  MCID for Activity domain,  $-1.44 \pm 1.89$  MCID for Pain,  $-0.56 \pm 1.15$  MCID for Appearance, and  $-0.58 \pm 2.19$  MCID for Mental. The average improvement for the Activity, Pain, and Appearance domains was between 1.58 and 1.64 MCID. On average, patients experienced the smallest improvement in the Mental domain ( $+0.97 \pm 1.98$  MCID).

The Activity domain had the highest rate of deterioration (14%), with a rate of optimal effectiveness equal to 36% (Fig. 4). The Pain domain had a rate of optimal effectiveness of 41% and had the greatest percentage of patients without improvement (23%). Appearance was the domain with the greatest rate of optimal improvement (56%) and smallest rate of deterioration (4%). Mental domain had the greatest percentage of consistent patients (39%) and the smallest percentage of optimal results (23%).

#### Clinical condition of patients at baseline

The threshold used in the baseline clinical group for Pain and Activity domains was based on the percentiles for the  $\text{Diff}_{\text{norm}}$  for the entire population at baseline (Table 6).

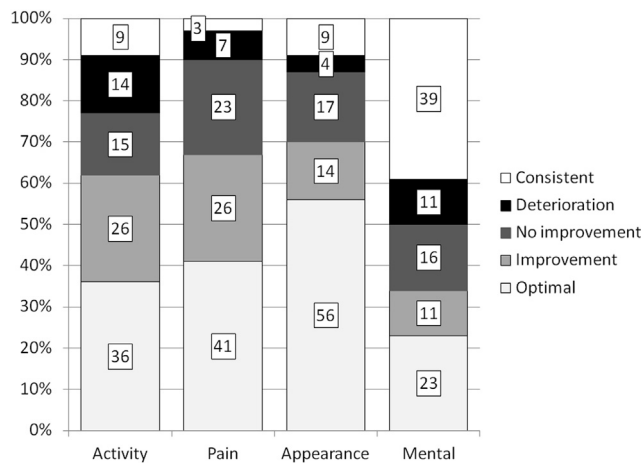


Fig. 4. Percentage of treatment effectiveness by SRS domains (Activity, Pain, Appearance, and Mental).

Thirty-six percent ( $n = 79$ ) of the patients were classified as being in the “worst” state, 28% ( $n = 63$ ) as “severe,” 19% ( $n = 43$ ) as “poor,” and 17% ( $n = 38$ ) as “moderate” (Fig. 2).

“Worst” ( $59 \pm 10$  years) and “severe” ( $59 \pm 14$  years) patients were significantly older than “moderate” patients ( $44 \pm 17$  years,  $p = .001$ ) (Table 7). “Worst” patients also had more severe Charlson comorbidity scores than “poor” and “moderate” patients ( $p = .014$ ). “Moderate” patients had a significantly lower BMI ( $23.0 \pm 3.5$ ) than “poor” and “worst” (respectively,  $27.3 \pm 6.7$  and  $29.2 \pm 5.9$ ;  $p = .001$ ) patients. Patients presenting with a history of previous spine surgery were more likely to be classified in the “worst” category ( $p = .004$ ). The proportion of women classified in the “moderate” category was significantly greater than the proportion of men ( $p = .001$ ).

The sagittal deformity only group (S) had the largest proportion in the “worst” category (54%), whereas no patients fell into the “moderate category” (Table 8). The group with double curve without sagittal deformity (D) had the smallest percentage of patients in the “worst category” and the greatest percentage of patients in the “moderate category.”

#### Overall treatment effectiveness

Thirty-three percent of the patients perceived an “optimal” improvement, 25% a “satisfactory” improvement, and 17% a “mediocre” improvement, while 24% did not perceive an improvement or deteriorated (Table 9). Only nine patients had less than 1 MCID of  $\text{Diff}_{\text{norm}}$  in Pain and Activity domains at baseline and no deterioration at the

Table 6

Percentile for the difference by domains with normative values ( $\text{Diff}_{\text{norm}}$ ) expressed in minimum clinically important difference (MCID) for the entire population at baseline.

Percentiles	$\text{Diff}_{\text{norm}}$ at baseline			
	Activity	Pain	Appearance	Mental
25th	-5	-4	-3	-3
50th	-3	-3	-2	-1
75th	-2	-2	-1	0



Table 7

Comparisons among the baseline clinical classification groups in terms of demographic data, gender, comorbidities, and history of prior surgery.

Baseline clinical classification	Demographic						Prior surgery,* %		% Female	% Male
	Age		BMI		Charlson score					
	Mean	SD	Mean	SD	Mean	SD	Yes	No		
Moderate	44*	17	23.0*	3.5	0.55*	0.95	11	22	19	6
Poor	52	18	26.2	4.3	0.88	0.91	15	23	19	20
Severe	59*	14	27.3*	6.7	1.30*	1.49	24	31	28	29
Worst	59*	10	29.2*	5.9	2.10*	1.79	51	24	34	45

BMI, body mass index; SD, standard deviation.

\*  $p < .05$ .

Table 8

Distribution of patients by curve type and by baseline SRS-Schwab classification.

SRS-Schwab classification	Curve type						
	T	L	D	TS	LS	DS	S
Moderate	40	18	43	20	12	22	0
Poor	20	23	33	0	15	24	15
Severe	0	32	19	40	29	33	31
Worst	40	27	5	40	44	22	54

D, double; DS, double sagittal; L, (thoraco)-lumbar; LS, (thoraco)-lumbar/sagittal; S, sagittal only; SRS, Scoliosis Research Society; T, thoracic; TS, thoracic/sagittal; U, unclassified.

Table 9

Distribution and assessment of demographic data based on global clinical treatment classification.

Overall treatment effectiveness classification	Distribution (%)	Demographic						Prior surgery, %		% Female	% Male
		Age		BMI		Charlson score					
		Mean	SD	Mean	SD	Mean	SD	Yes	No		
No improvement or deterioration	24	53	13	26.5	6.2	1.37	1.56	20	28	25	21
Mediocre	17	57	14	30.0	5.6	1.76	1.59	24	12	16	27
Satisfactory	25	58	15	27.8	5.3	1.70	1.81	26	24	25	27
Optimal	33	55	16	26.1	5.8	1.08	1.27	30	36	34	27

BMI, body mass index; SD, standard deviation.

Table 10

Cross-comparison of patient distribution based on preoperative curve type (SRS-Schwab classification) and global treatment outcome classification.

	Curve type						
	T	L	D	TS	LS	DS	S
No improvement or deterioration	22	25	30	25	29	25	15
Mediocre	33	5	5	50	14	17	25
Satisfactory	0	40	25	0	19	23	38
Optimal	44	30	40	25	38	35	22

D, double; DS, double sagittal; L, (thoraco)-lumbar; LS, (thoraco)-lumbar/sagittal; S, sagittal only; T, thoracic; TS, thoracic/sagittal; U, unclassified.

2-year follow-up (group “consistent” for Pain and Activity; they were not included in the following analysis).

There were no significant differences between the clinical treatment effectiveness groups based on age, gender, Charlson comorbidity score, or history of previous surgery (Table 9). Patients with a “mediocre” treatment effectiveness had a significantly greater BMI than the patients with an “optimal” treatment effectiveness.

Depending on curve type (Table 10), 15% to 30% of the patients did not perceive an improvement or reported symptoms indicating deterioration. Between 22% and 44% perceived an optimal improvement. Patients with only sagittal deformity had the lowest percentage of optimal improvement.

Patients’ distribution was significantly different when comparing the baseline clinical classification and the global

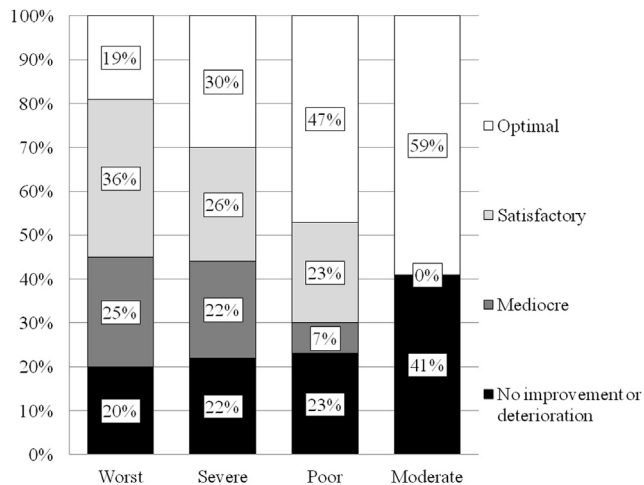


Fig. 5. Cross-comparison of patient distribution based on baseline clinical classification and overall treatment effectiveness classification.

clinical treatment effectiveness classification ( $p < .001$ , Fig. 5). Twenty percent of the “worst” patients had no improvement or had deteriorated, whereas 41% of “moderate” patients had no improvement or had deteriorated. Only 19% of “worst” patients had “optimal” results, and 59% of “moderate” patients had “optimal” results. Worst patients had also the greatest percentage of “satisfactory” results.

## Discussion

Using MCID, this study aimed to evaluate the clinical effectiveness of spine surgery for ASD using a normative population comparison group (stratified by age and sex).

Assessing surgical efficacy using age- and sex-matched normative data avoids bias in the analysis due to gender and age, especially in this sample population with a large age range (19–84 years). It also assists in setting realistic expectations among age groups. Soroceanu et al. [21] highlighted that patient outcomes were strongly correlated with the preoperative expectations. However, the main limitation of this approach relates to the distribution of normative data, with standard deviations (0.31–0.79) for each gender and age group approaching MCID thresholds (0.375 and 0.8). To our knowledge, comparison of HRQOL scores with normative data has only been reported in a study published by Blondel et al. [5] that aimed to evaluate the clinical impact of restoring sagittal alignment.

A unique feature of this study’s methodology is the use of MCID. Although MCIDs have been used in several studies for the clinical assessment of ASD patients [7–9], MCID values are traditionally used as thresholds but not as a scale that would require validation. Moreover MCID was developed to differentiate between statistical improvement from clinical improvement but was not validated for deterioration. To our knowledge, no studies exist that evaluates

MCID thresholds for deterioration. This deficiency was noted by Gum et al. [22] on a population of 722 patients who received posterior lumbar fusion for back and leg pain. The authors concluded that “a threshold for clinical deterioration was difficult to identify” and that “patients may interpret the absence of change as deterioration.” In light of these findings, our approach consisted of combining patients with no improvement with ones who sustained a perceived deterioration.

Nevertheless, we believe that a common MCID scale can permit an easier and more relevant clinical interpretation. It also permits comparison of different domains of the SRS instrument, which would not otherwise be possible. Another benefit of the presented methodology is that it combined the evaluation of change in clinical outcomes to the analysis of postoperative outcomes with regards to an age- and sex-matched normative population.

## Comparisons between patients and the normative population

At baseline, for 77% of the patients, the Pain and Disability domains were the two domains most highly affected. These results are in line with those reported by Bess et al. [23], who postulated that pain and disability were the main drivers of surgical treatment in the ASD population. Those results explained the focus on those domains in the baseline classification, as well as for the global treatment efficiency classification. For the baseline classification, thresholds for Pain and Activity domains were chosen based on percentiles and simple thresholds.

## Treatment effectiveness

The evaluation of treatment effectiveness took into account the most affected preoperative and postoperative domains (Pain and Activity), clinical improvement, and deterioration from baseline to follow-up. Thirty-three percent of the patients had a clinical improvement in their deficient domains that was sufficient to reach less than 1 MCID difference from normative data. However, 24% had no clinical improvement or were clinically deteriorated. The greatest percentage of deterioration was observed in the Activity domain (14%).

The Mental and Appearance domains have been identified as second-order compared with Pain and the Activity. The two domains were treated independently and were not incorporated into the overall classification so that clinical evaluation is not further complicated. The Appearance domain was the one most likely to show an “optimal” improvement (56%) and the least likely to deteriorate (4%), whereas the Mental domain was the least likely to show an “optimal” improvement (23%).



### *Baseline classification versus treatment effectiveness*

Two different aspects of the relationship between clinical baseline state and treatment effectiveness could be distinguished. Only 19% of patients with the greatest clinical deficit at baseline reached an optimal result postoperatively (for the other groups, the percentage was between 30% and 59%). However, 36% of these patients had a “satisfactory” result, potentially reflecting that for these patients the aim of reaching a score matching the normative population postoperatively may not be realistic. Alternatively, only 20% of these patients did not have an improvement or deteriorated, whereas 41% with a “moderate” deficit did not perceive an improvement or were deteriorated, highlighting two potential issues when indicating patients for surgery: First, performing surgery on patients with a less measurable negative impact may inherently limit the possible benefits that can be derived from intervention. Second, surgeries that were performed at an earlier time point in the disease process may also inherently suffer from an inability to demonstrate large differences in clinical improvement.

### *Demographic data*

In the ASD population, prior studies demonstrate that older patients reported worse clinical scores than younger patients but did not specifically assess whether this difference could be attributed to spinal deformity or age [23-25]. Baseline clinical classification demonstrated that patients with “severe” and “worst” health status at baseline were older, indicating that the negative impact of spinal deformity was greater among older patients. However, in terms of treatment effectiveness, no association was found with age, which is consistent with findings in other studies [24,25].

In the present study, patients with greater pain and disability had a greater BMI and higher Charlson comorbidity scores. BMI and comorbidities independently decrease quality of life. Like age, the analysis of clinical treatment effectiveness indicated that patients with a greater BMI or high levels of comorbidities can still expect a significant clinical improvement from surgery. However, the role of comorbidities and BMI seems to be controversial. Smith et al. [14] found in a multicenter spinal deformity database with 2-year follow-up that a greater BMI is a predictor of poor outcome. Daubs et al. [24] did not find a correlation between patient outcomes after spine surgery and comorbidities in an elderly ASD population, whereas Slover et al. [26] found that comorbidities impact the change in clinical improvement.

### *Revision versus primary surgery*

Patients with previous spine surgery had poorer HRQOLs preoperatively. However, no significant difference was found in terms of treatment effectiveness when

compared with primary cases. In contrast, Djurasovic et al. [27] found that revision cases showed only modest improvement and that 38% of the patients reach the MCID threshold. One possible explanation of these different conclusions is the difference in methodology used.

### *Deformity pattern*

Preoperatively, results demonstrated that patients with sagittal deformity had poorer HRQOL than patients with pure coronal deformity. These results are consistent with published reports regarding the importance of sagittal plane deformity [5,28]. Indeed, great variability was found between curve type and treatment effectiveness; depending on the type of curve, the percentage of optimal improvement was between 22% and 44%.

### *Limitations*

The follow-up rate at 2 years was about 70%, which is a weakness of our study. As mentioned previously, the use of MCID as a scale to quantify improvement and deterioration has not been validated to date. Finally, the relatively small number of patients included in the present study did not permit further analysis such as stratification by deformity pattern, age, complication, indication, or technique. Therefore, this study focused on the preoperative condition of the patients, but further analysis is warranted to investigate other potential factors that could drive surgical outcomes.

### *Conclusion*

Using MCID and normative scores, this study presents an innovative way to evaluate the clinical effectiveness of treatment for ASD patients. Thirty-three percent of patients achieved a health status that was less than 1 MCID from normative population values at 2 years after surgery, whereas 24% did not perceive improvement or experienced deterioration. Clinical state at baseline was a dual factor of treatment effectiveness. Patients with a moderate deficit had a greater chance of having an optimal result but also a greater chance of not perceiving an improvement. On the other hand, patients presenting with the poorest HRQOLs were more likely to perceive an improvement but were also less likely to reach optimal levels of functionality.

### *Key points:*

1. Using SRS score, 33% of patients with adult spinal deformity reached normative values at 2 years after surgery and 24% did not perceive improvement or deteriorated.
2. At baseline, for 77% of the patients, the Pain and Disability domains were the two domains most highly affected.

3. Patients with a moderate deficit had a greater chance of having an optimal result but also a greater chance of not perceiving an improvement.
4. Stated alternatively, patients presenting with the poorest HRQOLs were more likely to perceive an improvement but were also less likely to reach optimal improvement.

## References

- [1] Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2940–52; discussion 2952.
- [2] Asher MA, Lai SM, Glattes RC, et al. Refinement of the SRS-22 Health-Related Quality of Life questionnaire Function domain. *Spine (Phila Pa 1976)* 2006;31:593–7.
- [3] Baldus C, Bridwell KH, Harrast J, et al. Age-gender matched comparison of SRS instrument scores between adult deformity and normal adults: are all SRS domains disease specific? *Spine (Phila Pa 1976)* 2008;33:2214–8.
- [4] Berven S, Deviren V, Demir-Deviren S, et al. Studies in the modified Scoliosis Research Society Outcomes Instrument in adults: validation, reliability, and discriminatory capacity. *Spine (Phila Pa 1976)* 2003;28:2164–9; discussion 2169.
- [5] Blondel B, Schwab F, Ungar B, et al. Impact of magnitude and percentage of global sagittal plane correction on health-related quality of life at 2-years follow-up. *Neurosurgery* 2012;71:341–8; discussion 348.
- [6] Berven S, Deviren V, Demir-Deviren S. Minimal clinically important difference in adult spinal deformity: how much change is significant? Paper presented at: 12th International Meeting on Advanced Spine Technology; July 7–9, 2005; Banff Springs, Alberta, Canada.
- [7] Smith JS, Klineberg E, Schwab F, et al. Change in classification grade by the SRS-Schwab Adult Spinal Deformity Classification predicts impact on health-related quality of life measures: prospective analysis of operative and non-operative treatment. *Spine (Phila Pa 1976)* 2013;38:1663–71.
- [8] 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;14:67–77.
- [9] Smith JS, Singh M, Klineberg E, et al. Surgical treatment of pathological loss of lumbar lordosis (flatback) in patients with normal sagittal vertical axis achieves similar clinical improvement as surgical treatment of elevated sagittal vertical axis. *J Neurosurg Spine* 2014;21:160–70.
- [10] Baldus C, Bridwell K, Harrast J, et al. The Scoliosis Research Society Health-Related Quality of Life (SRS-30) age-gender normative data: an analysis of 1346 adult subjects unaffected by scoliosis. *Spine (Phila Pa 1976)* 2011;36:1154–62.
- [11] Terran J, Schwab F, Shaffrey CI, et al. The SRS-Schwab adult spinal deformity classification: assessment and clinical correlations based on a prospective operative and nonoperative cohort. *Neurosurgery* 2013;73:559–68.
- [12] Schwab FJ, Smith VA, Biserni M, et al. Adult scoliosis: a quantitative radiographic and clinical analysis. *Spine (Phila Pa 1976)* 2002;27:387–92.
- [13] Moal B, Schwab F, Ames CP, et al. Radiographic outcomes of adult spinal deformity correction: a critical analysis of variability and failures across deformity patterns. *Spine Deform* 2014;2:219–25.
- [14] Smith JS, Shaffrey CI, Glassman SD, et al. Clinical and radiographic parameters that distinguish between the best and worst outcomes of scoliosis surgery for adults. *Eur Spine J* 2013;22:402–10.
- [15] Glassman SD, Bridwell K, Dimar JR, et al. The impact of positive sagittal balance in adult spinal deformity. *Spine (Phila Pa 1976)* 2005;30:2024–9.
- [16] Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- [17] Champain S, Benchikh K, Nogier A, et al. Validation of new clinical quantitative analysis software applicable in spine orthopaedic studies. *Eur Spine J* 2006;15:982–91.
- [18] Rillardon L, Levassor N, Guigui P, et al. Validation of a tool to measure pelvic and spinal parameters of sagittal balance. *Rev Chir Orthop Reparatrice Appar Mot* 2003;89:218–27.
- [19] Bess S, Schwab F, Lafage V, et al. Classifications for adult spinal deformity and use of the Scoliosis Research Society-Schwab Adult Spinal Deformity Classification. *Neurosurg Clin N Am* 2013;24:185–93.
- [20] Liu Y, Liu Z, Zhu F, et al. Validation and reliability analysis of the new SRS-Schwab classification for adult spinal deformity. *Spine (Phila Pa 1976)* 2013;38:902–8.
- [21] Sorocanu A, Ching A, Abdu W, et al. Relationship between preoperative expectations, satisfaction, and functional outcomes in patients undergoing lumbar and cervical spine surgery: a multicenter study. *Spine (Phila Pa 1976)* 2012;37:E103–8.
- [22] Gum JL, Glassman SD, Carreon LY. Clinically important deterioration in patients undergoing lumbar spine surgery: a choice of evaluation methods using the Oswestry Disability Index, 36-Item Short Form Health Survey, and pain scales: clinical article. *J Neurosurg Spine* 2013;19:564–8.
- [23] Bess S, Boachie-Adjei O, Burton D, et al. Pain and disability determine treatment modality for older patients with adult scoliosis, while deformity guides treatment for younger patients. *Spine (Phila Pa 1976)* 2009;34:2186–90.
- [24] Daubs MD, Lenke LG, Cheh G, et al. Adult spinal deformity surgery: complications and outcomes in patients over age 60. *Spine (Phila Pa 1976)* 2007;32:2238–44.
- [25] Smith JS, Shaffrey CI, Glassman SD, et al. Risk-benefit assessment of surgery for adult scoliosis: an analysis based on patient age. *Spine (Phila Pa 1976)* 2011;36:817–24.
- [26] Slover J, Abdu WA, Hanscom B, et al. The impact of comorbidities on the change in Short-Form 36 and Oswestry scores following lumbar spine surgery. *Spine (Phila Pa 1976)* 2006;31:1974–80.
- [27] Djurasovic M, Glassman SD, Howard JM, et al. Health-related quality of life improvements in patients undergoing lumbar spinal fusion as a revision surgery. *Spine (Phila Pa 1976)* 2011;36:269–76.
- [28] Glassman SD, Berven S, Bridwell K, et al. Correlation of radiographic parameters and clinical symptoms in adult scoliosis. *Spine (Phila Pa 1976)* 2005;30:682–8.