

Potential of predictive computer models for preoperative patient selection to enhance overall quality-adjusted life years gained at 2-year follow-up: a simulation in 234 patients with adult spinal deformity

Taemin Oh, MD,¹ Justin K. Scheer, MD,² Justin S. Smith, MD, PhD,³ Richard Hostin, MD,⁴ Chessie Robinson, MA,⁵ Jeffrey L. Gum, MD,⁶ Frank Schwab, MD,⁷ Robert A. Hart, MD,⁸ Virginie Lafage, PhD,⁷ Douglas C. Burton, MD,⁹ Shay Bess, MD,¹⁰ Themistocles Protopsaltis, MD,⁷ Eric O. Klineberg, MD,¹¹ Christopher I. Shaffrey, MD,³ Christopher P. Ames, MD,¹ and the International Spine Study Group

¹Department of Neurological Surgery, University of California, San Francisco, California; ²Department of Neurosurgery, University of Illinois at Chicago, Illinois; ³Department of Neurosurgery, University of Virginia Health System, Charlottesville, Virginia; ⁴Department of Orthopaedic Surgery, Baylor Scoliosis Center, Plano; ⁵Baylor Scott & White Health, Center for Clinical Effectiveness, Dallas, Texas; ⁶Norton Leatherman Spine Center, Louisville, Kentucky; ⁷Department of Orthopaedic Surgery, New York University Hospital for Joint Diseases, New York, New York; ⁸Department of Orthopaedic Surgery, Oregon Health & Science University, Portland, Oregon; ⁹Department of Orthopaedic Surgery, University of Kansas Medical Center, Kansas City, Kansas; ¹⁰Rocky Mountain Hospital for Children, Denver, Colorado; and ¹¹Department of Orthopaedic Surgery, University of California, Davis, California

OBJECTIVE Patients with adult spinal deformity (ASD) experience significant quality of life improvements after surgery. Treatment, however, is expensive and complication rates are high. Predictive analytics has the potential to use many variables to make accurate predictions in large data sets. A validated minimum clinically important difference (MCID) model has the potential to assist in patient selection, thereby improving outcomes and, potentially, cost-effectiveness.

METHODS The present study was a retrospective analysis of a multiinstitutional database of patients with ASD. Inclusion criteria were as follows: age \geq 18 years, radiographic evidence of ASD, 2-year follow-up, and preoperative Oswestry Disability Index (ODI) $>$ 15. Forty-six variables were used for model training: demographic data, radiographic parameters, surgical variables, and results on the health-related quality of life questionnaire. Patients were grouped as reaching a 2-year ODI MCID (+MCID) or not (–MCID). An ensemble of 5 different bootstrapped decision trees was constructed using the C5.0 algorithm. Internal validation was performed via 70:30 data split for training/testing. Model accuracy and area under the curve (AUC) were calculated. The mean quality-adjusted life years (QALYs) and QALYs gained at 2 years were calculated and discounted at 3.5% per year. The QALYs were compared between patients in the +MCID and –MCID groups.

RESULTS A total of 234 patients met inclusion criteria (+MCID 129, –MCID 105). Sixty-nine patients (29.5%) were included for model testing. Predicted versus actual results were 50 versus 40 for +MCID and 19 versus 29 for –MCID (i.e., 10 patients were misclassified). Model accuracy was 85.5%, with 0.96 AUC. Predicted results showed that patients in the +MCID group had significantly greater 2-year mean QALYs ($p = 0.0057$) and QALYs gained ($p = 0.0002$).

CONCLUSIONS A successful model with 85.5% accuracy and 0.96 AUC was constructed to predict which patients would reach ODI MCID. The patients in the +MCID group had significantly higher mean 2-year QALYs and QALYs gained. This study provides proof of concept for using predictive modeling techniques to optimize patient selection in complex spine surgery.

<https://thejns.org/doi/abs/10.3171/2017.9.FOCUS17494>

KEY WORDS Oswestry Disability Index; minimum clinically important difference; predictive modeling; quality-adjusted life year

ABBREVIATIONS ASD = adult spinal deformity; AUC = area under the curve; BMI = body mass index; HRQOL = health-related QOL; IBF = interbody fusion; LIV = lowest instrumented vertebra; MCID = minimum clinically important difference; NRS = numerical rating scale; ODI = Oswestry Disability Index; QALY = quality-adjusted life year; QOL = quality of life; SF-36 = 36-Item Short-Form Health Survey; SPO = Smith-Petersen osteotomy; SRS = Scoliosis Research Society; UIV = uppermost instrumented vertebra.

SUBMITTED August 1, 2017. **ACCEPTED** September 13, 2017.

INCLUDE WHEN CITING DOI: 10.3171/2017.9.FOCUS17494.

ADULT spinal deformity (ASD) is a debilitating disease affecting up to one-third of the general population, and its prevalence markedly increases with age.⁵ As the population of the US ages, ASD is becoming increasingly recognized as a disease that could reach epidemic proportions.^{5,41} Importantly, the substantial burden of symptomatic ASD on patient health and well-being has been well established, which underscores the need for treating ASD safely and efficiently.^{8,13,22,34}

Although there is still some debate regarding optimum management paradigms, surgical intervention remains one of the cornerstones of therapy. Recent research led by our group and others has shown that, compared with nonoperative treatment, surgery is associated with significantly greater improvement in pain, functionality, and quality of life (QOL), as well as increasing the likelihood of reaching a minimum clinically important difference (MCID) for these improvements.^{2,3,12,19,27,29,33,35} However, this observation is not universal, because in select patients the associated postoperative complications may impact the benefits. For example, Bourghli et al. found that patients who fail to improve in QOL measures 6 months after surgery are more likely to undergo further revision surgeries for deformity correction.¹¹ Liu et al. also showed that among 239 patients treated surgically for ASD, 44% did not reach an MCID and 5% of patients actually deteriorated.¹⁹ For such patients, reoperations can contribute further to personal and financial costs, as well as putting them at additional risk for morbidity or mortality.

Optimizing patient selection has always been a critical ingredient toward maximizing chances for surgical success, but consistently predicting which patients will perform well postoperatively remains a challenging task. Multiple studies have shown that 1) on average, surgical treatment for symptomatic ASD leads to greater qualitative improvements in QOL compared with nonoperative treatment, as measured by patient responses to health-related QOL (HRQOL) questionnaires,^{12,19,33} and 2) on average, surgical treatment for symptomatic ASD leads to greater quantitative improvements in QOL, as measured by quality-adjusted life years (QALYs).^{20,27,40} However, no study to date has successfully combined these 2 findings to show that patients who are likely to demonstrate improvements in HRQOL measures also have greater improvements in QALYs.

Thus, the primary goal of this study was to develop and internally validate a preoperative, predictive model to identify which patients will reach a 2-year MCID for the Oswestry Disability Index (ODI), and to subsequently apply this model to demonstrate that these patients achieve greater QALY improvements at 2-year follow-up from surgery. By establishing such a predictive model, we aim to provide a preoperative tool that surgeons can use to help identify which patients are most likely to be successful surgical candidates. In health systems with limited economic resources, the adoption of models as decision tools may assist physicians in guiding cost-effective care. Such a model could potentially find great value when used on a mobile or internet app at point-of-care clinic visits, to facilitate discussions with patients prior to surgery.

Methods

Study Population

In this study, a retrospective review was performed on a prospective, multicenter (11 sites) database of consecutively treated patients with ASD. Inclusion criteria for this database were as follows: 1) age \geq 18 years and 2) radiographic diagnosis of ASD, defined by at least 1 of the following variables—scoliosis, Cobb angle \geq 20°, C7–S1 sagittal vertical axis \geq 5 cm, pelvic tilt \geq 25°, or thoracic kyphosis \geq 60°. Patients with active infections, malignancy, and neuromuscular spinal deformities were excluded from the database. For the present study, additional inclusion criteria were applied post hoc and included 1) complete 2-year follow-up and 2) preoperative ODI score $>$ 15. This preoperative ODI cutoff was set based on previously established MCID thresholds for the ODI (-15).^{10,17}

Data Collection—Demographics, Surgical Variables, HRQOL, Radiographic Parameters

In total, 46 variables were collected and included in the initial training of the predictive model. Demographic variables included patient age, sex, race, body mass index (BMI), Charlson Comorbidity Index, and American Society of Anesthesiologists grade, as well as the presence and number of preoperative comorbidities, including arthritis, osteoporosis, depression, and smoking history. Planned surgical variables were either binary (e.g., whether a decompression, osteotomy, Smith-Petersen osteotomy [SPO], 3-column osteotomy, or interbody fusion [IBF] were performed) or continuous (e.g., number of levels fused or decompressed, number of IBFs or SPOs, uppermost instrumented vertebra [UIV], and lowermost instrumented vertebra [LIV]). The HRQOL scores were extracted from patient responses to established questionnaires, including the ODI, 36-Item Short-Form Health Survey (SF-36), Scoliosis Research Society Outcomes Questionnaire (SRS-22r), back and/or leg pain numerical rating scale (NRS) scores, frailty index, and work status. The frailty index is a novel score for preoperative risk stratification (unpublished data), and patients were categorized as being not frail, frail, or severely frail.¹⁸ Work status was categorized with a score from 1 to 5, as follows: 1, disabled; 2, employed; 3, retired; 4, retired due to back pain; and 5, unemployed.

In accordance with existing protocols for radiographic evaluation, full-length, free-standing 36-inch cassette lateral spine radiographs were obtained preoperatively.⁶ These radiographs were analyzed using validated software (Spineview, ENSAM; Laboratory of Biomechanics)^{15,25} and techniques previously reported in the literature.^{21,31,36} Model creation was limited to the following preoperative radiographic parameters: maximum thoracic/lumbar Cobb angle groups ($<$ 30°, 30–60°, $>$ 60°); coronal C-7 plumb line; thoracic kyphosis (T4–12 Cobb angle); pelvic tilt; pelvic incidence and lumbar lordosis mismatch; C7–S1 sagittal vertical axis; and T1 pelvic angle.^{9,24,36} Patients were additionally classified according to the SRS-Schwab coronal curve classification system.³²

Patients were stratified based on whether they at least reached an MCID for ODI (+MCID) or not ($-$ MCID) at the 2-year follow-up. The MCID is a variable derived from

TABLE 1. Summary of demographic data (11 variables) in 234 patients with ASD

Variable	All Patients, n = 234	+MCID, n = 129	–MCID, n = 105	p Value
Age in yrs	59.3 ± 13.2	59.9 ± 12.7	58.5 ± 13.9	0.471
Sex, F/M	195:39 (83.3%)	108:21 (83.7%)	87:18 (82.9%)	0.860
Race				0.742
Asian	4 (1.7%)	3 (2.3%)	1 (1.0%)	
Black	13 (5.6%)	6 (4.7%)	7 (6.7%)	
Hispanic	4 (1.7%)	2 (1.6%)	2 (1.9%)	
White	199 (85.0%)	115 (89.1%)	84 (80.0%)	
Other	3 (1.3%)	1 (0.8%)	2 (1.9%)	
BMI	27.7 ± 5.8	27.7 ± 6.1	27.7 ± 5.4	0.897
ASA	2.4 ± 0.6	2.4 ± 0.6	2.4 ± 0.6	0.840
CCI	1.9 ± 1.8	1.9 ± 1.7	1.9 ± 1.8	0.957
≥1 comorbidity	178 (76.1%)	102 (79.1%)	76 (72.4%)	0.741
Arthritis	108 (46.2%)	60 (46.5%)	48 (45.7%)	0.725
Osteoporosis	59 (25.2%)	19 (14.7%)	13 (12.4%)	0.747
Depression	32 (13.7%)	38 (29.5%)	21 (20.0%)	0.166
Smoker	16 (6.8%)	8 (6.2%)	8 (7.6%)	0.571
No. of comorbidities	2.3 ± 1.8	2.3 ± 1.8	2.2 ± 1.8	0.559

ASA = American Society of Anesthesiologists; CCI = Charlson Comorbidity Index.

Comparison of demographics between patients meeting ODI MCID at 2 years (+MCID) or not (–MCID). Values are expressed as the number of patients (%) or as the mean ± SD.

established HRQOL surveys, such as the ODI, and serves as a quantitative measure for whether an intervention resulted in a clinically meaningful difference to patient outcome. Threshold MCID values have been previously established.^{10,14,16,17}

Quality-Adjusted Life Years

A QALY is both a quantitative and qualitative measure, and is calculated as a product of “health state” (ranging from 0 [death] to 1 [completely healthy]) and time spent in that specific health state.^{39,42} To calculate QALYs, health utility values were derived from Short-Form 6-Dimension (SF-6D) utility index scores at the preoperative time point and at 2 years after surgery, based on established formulas.^{27,42} The mean QALYs were subsequently discounted at the recommended rate of 3.5% per year.⁴³

Statistical Analysis

Continuous variables were reported as mean ± SD. Baseline parameters between the 2 subgroups (+MCID vs –MCID) were analyzed using the Student t-test or Pearson’s chi-square test for continuous and categorical variables, respectively. For nonparametric variables (tested using the Shapiro-Wilk test), the Mann-Whitney test was used instead. All statistical analyses were performed using SPSS v22 (IBM, Inc.). The threshold for statistical significance was $p < 0.05$.

During creation of the predictive model, missing data were imputed using standard techniques, and variables were tested for collinearity.¹ The target variable was made binary as follows: 1 for patients who achieved at least an MCID for ODI at 2-year follow-up (+MCID) or 0 for patients who did not achieve at least an MCID for ODI at

2-year follow-up (–MCID). Subsequently, an ensemble of decision trees was generated with 5 different bootstrapped models, each of which used a unique random sample of the data, using the C5.0 algorithm. The final model represents a combination of all 5 models. Internal validation was performed using a 70:30 data split for training versus testing of each model.¹ Overall accuracy as well as the area under the curve (AUC) were calculated. Predictive modeling was performed using SPSS Modeler version 16 (IBM, Inc.).

Results

Patient Population

A total of 234 patients with ASD met inclusion criteria, of whom 129 were +MCID and 105 were –MCID at 2-year follow-up. There was an overall preponderance of women (195 women, 39 men; 83.3% women), with a mean age of 59.3 ± 13.2 years and a mean BMI of 27.7 ± 5.8. Most patients presented with at least 1 comorbidity (178 patients, 76.1%), and the mean number of comorbidities was 2.3 ± 1.8. No significant differences in demographic variables were observed in the +MCID versus –MCID groups (Table 1).

Radiographic Features

All reported preoperative radiographic parameters were similar in the +MCID versus –MCID groups (Table 2), and no significant differences were appreciated ($p > 0.05$ for all).

Surgical Parameters

Most surgeries were primary (65.8%) versus revision (34.2%). Both the +MCID and –MCID cohorts under-

TABLE 2. Summary of radiographic parameters (8 variables) in 234 patients with ASD

Variable	All Patients, n = 234	+MCID, n = 129	–MCID, n = 105	p Value
Max Cobb angle groups				0.322
<30°	78 (33.3%)	40 (31.0%)	38 (36.2%)	
30–60°	108 (46.2%)	65 (50.4%)	43 (41.0%)	
>60°	43 (18.4%)	21 (16.3%)	22 (21.0%)	
SRS-Schwab coronal curve type				0.655
Type N	76 (32.5%)	43 (33.3%)	33 (31.4%)	
Type T	7 (3.0%)	5 (3.9%)	2 (1.9%)	
Type L	94 (40.2%)	48 (37.2%)	46 (43.8%)	
Type D	57 (24.4%)	33 (25.6%)	24 (22.9%)	
Coronal C-7 plumb line, mm	–2.0° ± 54.2°	–0.24° ± 56.4°	–4.3° ± 51.4°	0.472
Thoracic kyphosis	31.4° ± 18.7°	31.9° ± 19.6°	30.8° ± 17.5°	0.657
Sagittal vertical axis, mm	68.0° ± 77.6°	68.1° ± 81.9°	67.7° ± 72.3°	0.727
T1 pelvic angle	23.2° ± 13.1°	23.3° ± 13.6°	23.1° ± 12.4°	0.790
Pelvic tilt	24.4° ± 10.6°	24.4° ± 11.0°	24.5° ± 10.2°	0.590
PI-LL	16.7° ± 20.7°	16.1° ± 20.9°	17.5° ± 20.5°	0.905

Max = maximum; PI-LL = pelvic incidence and lumbar lordosis mismatch.

Comparison of radiographic parameters between patients meeting ODI MCID at 2 years (+MCID) or not (–MCID). See Schwab et al. for classification subtypes for the SRS-Schwab coronal curve types. Values are expressed as the number of patients (%) or as the mean ± SD.

went a similar number of decompressions, IBFs, SPOs, and 3-column osteotomies (Table 3, $p > 0.05$ for all). In addition, the mean number of levels decompressed, levels fused, SPO levels, and IBF levels were similar. The most common levels for UIV and LIV were T-10 and pelvis, respectively, and there were no significant differences in UIV or LIV between groups.

Baseline HRQOL

The +MCID group had a significantly higher baseline mean ODI than the –MCID group ($p = 0.038$, Table 4), indicating greater disability. No other significant differences were appreciated.

Model Results

A total of 165 patients (70.5%) were used for model training, and 69 (29.5%) were used for model testing. This percent split, although arbitrarily determined, was considered within the acceptable range for predictive modeling.¹ Model accuracy was 85.5% correct, with an AUC of 0.96, indicating good model fit. Predicted versus actual results were 50 versus 40 patients for +MCID and 19 versus 29 patients for –MCID, which translates to 10 misclassified patients. The 11 most robust predictors of reaching the goal of ≥ 1 ODI MCID at 2-year follow-up are presented in Table 5 (predictor importance ≥ 0.80). Demographic variables included baseline depression, arthritis, osteoporosis, and presence of at least 1 comorbidity; HRQOL variables included back and/or leg pain NRS and ODI scores; and surgical variables included number of levels fused, the UIV, and whether a 3-column osteotomy or decompression was performed.

Mean QALYs and QALYs Gained

The mean QALYs and QALYs gained at 2-year follow-

up were calculated. As shown in Fig. 1, when using the predicted results derived from the model, patients in the +MCID group demonstrated higher mean 2-year QALYs (1.2 ± 0.2 vs 1.1 ± 0.2 ; $p = 0.0057$). As shown in Fig. 2, the number of QALYs gained at 2 years was also greater in the +MCID group (0.2 ± 0.2 vs -0.02 ± 0.2 ; $p = 0.0002$). This effect was observed despite 10 patients being misclassified during model training.

Discussion

The overall objective for the present study was to construct an accurate, internally validated, predictive model for identifying patients with ASD who are more likely to be better surgical candidates and thus more likely to derive quantitative improvements in QOL. The results of our investigation provide a proof-of-concept predictive model that 1) predicts with 86% accuracy which patients will achieve at least an MCID for ODI at 2-year follow-up, and 2) demonstrates that patients who do so are more likely to have higher mean and cumulative QALYs gained. Our model demonstrated an AUC of 0.96, which suggests an excellent model fit, and was successfully applied to detect significant differences in QALYs despite 10/165 patients being misclassified on initial training.

Optimal preoperative patient selection is important to help maximize the chances of achieving surgical success, not only for the surgeon but also, more importantly, for the patient. Although evidence-based care drives much clinical decision making, from the surgeon's perspective there is a lack of tools currently available that can predict with robust accuracy and consistency which patients are likely to perform better. Even when it pertains to surgical planning, evidence shows that surgeons cannot accurately predict postoperative spinal alignment in 33% of cases, which suggests that there is a need to improve our arsenal

TABLE 3. Summary of surgical parameters (12 variables) in 234 patients with ASD

Variable	All Patients, n = 234	+MCID, n = 129	-MCID, n = 105	p Value
Revision surgeries	80 (34.2%)	40 (31.0%)	40 (38.1%)	0.256
Mean no. of posterior levels fused	11.3 ± 4.4	11.7 ± 4.3	10.7 ± 4.6	0.437
Decompression	162 (69.2%)	94 (72.9%)	68 (64.8%)	0.181
Mean no. of decompression levels	1.1 ± 0.3	1.1 ± 0.3	1.1 ± 0.3	0.208
Osteotomy	169 (72.2%)	89 (69.0%)	80 (76.2%)	0.221
SPO	130 (55.6%)	71 (55.0%)	59 (56.2%)	0.860
Mean no. of SPO levels	2.9 ± 3.1	3.0 ± 3.2	2.9 ± 3.0	0.542
3-column osteotomy	55 (23.5%)	25 (19.4%)	30 (28.6%)	0.099
IBF	165 (70.5%)	92 (71.3%)	73 (69.5%)	0.765
Mean no. of IBF levels	1.9 ± 1.9	1.9 ± 1.8	1.8 ± 1.9	0.908
Postop UIV				0.557
T-1	2 (0.9%)	1 (0.8%)	1 (1.0%)	
T-2	17 (7.3%)	11 (8.5%)	6 (5.7%)	
T-3	36 (15.4%)	20 (15.5%)	16 (15.2%)	
T-4	32 (13.7%)	18 (14.0%)	14 (13.3%)	
T-5	9 (3.8%)	6 (4.7%)	3 (2.9%)	
T-6	3 (1.3%)	3 (2.3%)	0 (0.0%)	
T-7	1 (0.4%)	1 (0.8%)	0 (0.0%)	
T-8	2 (0.9%)	2 (1.6%)	0 (0.0%)	
T-9	14 (6.0%)	8 (6.2%)	6 (5.7%)	
T-10	55 (23.5%)	32 (24.8%)	23 (21.9%)	
T-11	22 (9.4%)	9 (7.0%)	13 (12.4%)	
T-12	8 (3.4%)	2 (1.6%)	6 (5.7%)	
L-1	10 (4.3%)	6 (4.7%)	4 (3.8%)	
L-2	10 (4.3%)	4 (3.1%)	6 (5.7%)	
L-3	6 (2.6%)	2 (1.6%)	4 (3.8%)	
L-4	1 (0.4%)	0 (0.0%)	1 (1.0%)	
L-5	3 (1.3%)	2 (1.6%)	1 (1.0%)	
Postop LIV				0.930
T-12	3 (1.3%)	1 (0.8%)	2 (2.0%)	
L-1	2 (0.9%)	1 (0.8%)	1 (1.0%)	
L-2	3 (1.3%)	1 (0.8%)	2 (2.0%)	
L-3	4 (1.7%)	3 (2.3%)	1 (1.0%)	
L-4	13 (5.6%)	6 (4.7%)	7 (6.7%)	
L-5	18 (7.7%)	11 (8.5%)	7 (6.7%)	
S-1	19 (8.1%)	9 (7.0%)	10 (9.5%)	
S-2	2 (0.9%)	1 (0.8%)	1 (1.0%)	
Pelvis	167 (71.4%)	94 (72.9%)	73 (69.5%)	

Comparison of surgical parameters between patients meeting ODI MCID at 2 years (+MCID) or not (-MCID). Values are expressed as the number of patients (%) or as the mean ± SD.

TABLE 4. Summary of HRQOL parameters (15 variables) in 234 patients with ASD

Variable	All Patients, n = 234	+MCID, n = 129	-MCID, n = 105	p Value
ODI	46.5 ± 14.8	48.3 ± 14.2	44.3 ± 15.3	0.038*
PCS	30.0 ± 8.0	30.0 ± 7.5	30.4 ± 8.7	0.517
MCS	44.5 ± 14.1	44.0 ± 14.7	45.2 ± 13.2	0.581
SRS-22r				
Activity	2.8 ± 0.8	2.8 ± 0.8	2.8 ± 0.7	0.572
Pain	2.3 ± 0.7	2.3 ± 0.7	2.3 ± 0.8	0.993
Appearance	2.3 ± 0.7	2.3 ± 0.7	2.3 ± 0.7	0.795
Mental	3.4 ± 0.9	3.4 ± 0.9	3.4 ± 1.0	0.617
Satisfaction	2.8 ± 1.1	2.7 ± 1.1	2.9 ± 1.1	0.224
Total	2.7 ± 0.6	2.7 ± 0.6	2.7 ± 0.6	0.799
Frailty index	3.7 ± 0.9	3.6 ± 0.9	3.7 ± 0.8	0.419
Frailty index category	1.4 ± 0.5	1.3 ± 0.5	1.4 ± 0.5	0.646
Work status	2.6 ± 1.1	2.6 ± 1.2	2.6 ± 1.1	0.841
NRS				
Back pain	7.5 ± 2.0	7.3 ± 2.0	7.7 ± 1.9	0.185
Leg pain	5.0 ± 3.2	5.3 ± 3.3	4.7 ± 3.2	0.120

MCS = mental component summary of the SF-36; PCS = physical component summary of the SF-36.

Comparison of HRQOL parameters between patients meeting ODI MCID at 2 years (+MCID) or not (-MCID). Values are expressed as the mean ± SD.

* Statistically significant at p < 0.05.

as well. Ideally, surgeons and patients both would have access to a preoperative algorithm that can fulfill this function. In that regard, the model built in this current study lays the foundation for predictive analytics upon which such an algorithm can be built and used in the clinical setting prior to considering surgical intervention.

Predictive modeling has been increasingly used by our group over the past several years. The benefits and pitfalls of predictive modeling have been previously described.²³ Predictive modeling is a statistical technique that represents a departure from the more traditional regression analyses; however, it affords greater flexibility by allowing patterns within the available data to create accurate, patient-specific predictions without the need to establish hypotheses or control groups a priori.¹ Through this established methodology, we have successfully developed models for accurate prediction of postoperative complications,³⁰ proximal junction failure and proximal junction kyphosis,²⁸ and length of hospital stay.²⁶ In the present study, the same modeling techniques were used in a large cohort of surgically treated patients with ASD to predict which patients will be likely to have better outcomes, as defined by meeting MCID thresholds for ODI and QALY improvements.

Our model uses preoperative demographics, surgical variables, HRQOL responses, and radiographic variables to predict which patients will achieve clinically significant improvements in QOL (≥ 1 ODI MCID). Patient demographics (preoperative depression/arthritis/osteoporosis, presence of comorbidity), surgical variables (decompression, number of levels fused, 3-column osteotomies, UIV), and HRQOL measures (ODI, back and/or leg pain NRS

of tools for preoperative planning.⁴ From the patient's perspective, knowledge of their preoperative risk factors and how these impact their chances for a successful outcome can add invaluable information for their decision making

TABLE 5. The 11 most important variables for predictor modeling in 234 patients with ASD

Order of Importance	Variable
1	Depression
2	ODI
3	No. of posterior levels fused
4	Presence of arthritis
5	Presence of ≥ 1 comorbidity
6	Presence of osteoporosis
7	3-column osteotomy PSO/VCR
8	UIV
9	Leg pain NRS
10	Decompression (yes vs no)
11	Back pain NRS

PSO = pedicle subtraction osteotomy; VCR = vertebral column resection.

scores) were all well represented among the 11 most important predictors. Preoperative depression was the most highly ranked predictor. This is not an unexpected finding, because Smith and colleagues have already demonstrated that patients with baseline depression have significantly worse SRS-22r and ODI scores.^{37,38}

Interestingly, radiographic parameters did not rank among the most important predictors. However, it should be emphasized that this does not necessarily discount their importance; the final predictive model was built using all 46 variables, and this model is intended to be generalizable. In select patient populations, radiographic parameters may serve a more critical role, especially with longer-term follow-up.

Our model also demonstrates that patients who are predicted to meet the +MCID threshold at 2-year follow-up have significantly higher mean QALYs and QALYs gained. This finding lends greater value to our predictive model, because QALYs represent QOL per unit of time and can thus be used as a proxy to measure quantitative improvements in QOL. Much of the national rhetoric surrounding health care has emphasized the delivery of cost-effective health care, and this model provides evidence that patients predicted to derive qualitative benefits postoperatively are more likely to experience quantitative benefits as well. An argument could be made for possible selection bias, because our analysis excluded patients with a baseline ODI score of < 15 . This cohort of patients not only has an excellent baseline QOL, but is also likely to retain an excellent QOL postoperatively,⁷ potentially translating to greater mean QALYs. By excluding them in the current analysis, we could potentially be underestimating the mean QALYs in the -MCID group. However, the counterargument holds that by including patients with an ODI score of < 15 , this could lead to overestimation of the mean QALYs in the -MCID group. Furthermore, the primary goal of the current study was to develop a model that could predict which patients would meet the MCID for ODI, so patients who could not meet the MCID were excluded. Last, our analysis demonstrates a statistically significant difference in not only mean QALYs but also

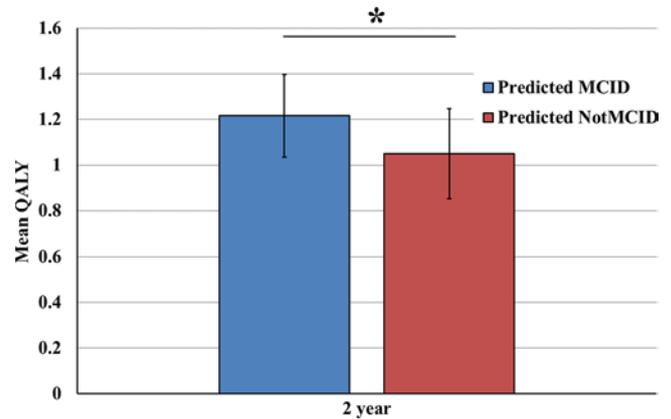


FIG. 1. Bar graph showing the mean QALYs, comparing patients in the +MCID versus -MCID group. A comparison of the mean QALYs between patients predicted to meet the ≥ 1 ODI MCID threshold at 2 years (+MCID) versus those predicted not to meet the threshold (-MCID). Asterisk indicates values meeting the statistical significance threshold $p < 0.05$.

QALYs gained, which provides further evidence that there are quantitative improvements in QOL among patients who meet the MCID ODI score.

The strengths of this study relate to its multiinstitutional nature, which lends greater generalizability to our model, the large volume of patients with ASD, and the detailed follow-up extending to 2 years postoperatively. The limitations of this study primarily relate to the retrospective design. Furthermore, although the use of predictive analytics can be an important tool that can aid in decision making, it is simply one piece of the complex algorithm for preoperative decision making. In that regard, just because patients present with one or more of the preoperative predictors identified in our model does not necessarily translate to such patients being poor surgical candidates. Third, this model cannot be applied to patients with an ODI score of < 15 at baseline, because these patients were not included in model development. Last, the “success of surgery” can-

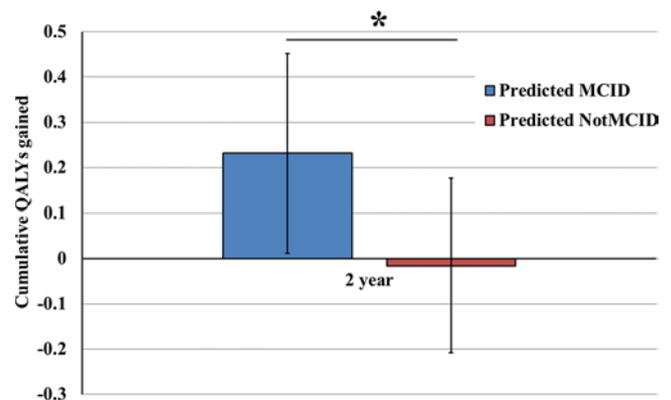


FIG. 2. Bar graph showing the cumulative QALYs gained, comparing patients in the +MCID versus -MCID group. A comparison of cumulative QALYs between patients predicted to meet the ≥ 1 ODI MCID threshold at 2 years (+MCID) versus those predicted not to meet the threshold (-MCID). Asterisk indicates values meeting the statistical significance threshold $p < 0.05$.

not always be captured using HRQOL surveys, and may incorporate elements not covered by our model (e.g., financial burden of surgery impacting QOL). With that being said, this model can serve as a launching point to facilitate point-of-care discussions between surgeons and patients prior to considering surgical intervention for ASD.

Conclusions

A predictive model with 86% accuracy and 0.96 AUC was built to predict which patients are likely to reach MCID thresholds for ODI postoperatively, and the model was used to demonstrate that patients who are better surgical candidates are also more likely to have higher QALYs, as well as to gain more QALYs. Such a model may find use in the clinical setting in the form of a mobile or internet app to quantify each individual patient's outcome risk, and can assist in the surgical decision-making process.

References

- Abbott D: **Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst**. Indianapolis: Wiley, 2014
- Acaroglu E: Decision-making in the treatment of adult spinal deformity. **EFORT Open Rev** 1:167–176, 2017
- Acaroglu E, Yavuz AC, Guler UO, Yuksel S, Yavuz Y, Domingo-Sabat M, 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** 25:2390–2400, 2016
- Ailon T, Scheer JK, Lafage V, Schwab FJ, Klineberg E, Sciubba DM, et al: Adult spinal deformity surgeons are unable to accurately predict postoperative spinal alignment using clinical judgment alone. **Spine Deform** 4:323–329, 2016
- Ames CP, Scheer JK, Lafage V, Smith JS, Bess S, Berven SH, et al: Adult spinal deformity: epidemiology, health impact, evaluation, and management. **Spine Deform** 4:310–322, 2016
- Angevine PD, Kaiser MG: Radiographic measurement techniques. **Neurosurgery** 63 (3 Suppl):40–45, 2008
- Bakhsheshian J, Scheer JK, Gum JL, Horner L, Hostin R, Lafage V, et al: Comparison of structural disease burden to health-related quality of life scores in 264 adult spinal deformity patients with 2-year follow-up: novel insights into drivers of disability. **Clin Spine Surg** 30:E124–E131, 2017
- Bess S, Line B, Fu KM, McCarthy I, Lafage V, Schwab F, 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)** 41:224–233, 2016
- Bess S, Protosaltis TS, Lafage V, Lafage R, Ames CP, Errico T, et al: Clinical and radiographic evaluation of adult spinal deformity. **Clin Spine Surg** 29:6–16, 2016
- Blondel B, Schwab F, Ungar B, Smith J, Bridwell K, Glassman S, et al: Impact of magnitude and percentage of global sagittal plane correction on health-related quality of life at 2-years follow-up. **Neurosurgery** 71:341–348, 2012
- Bourghli A, Boissiere L, Larrieu D, Vital JM, Yilgor C, Pellisé F, et al: Lack of improvement in health-related quality of life (HRQOL) scores 6 months after surgery for adult spinal deformity (ASD) predicts high revision rate in the second postoperative year. **Eur Spine J** 26:2160–2166, 2017
- Bridwell KH, Glassman S, Horton W, Shaffrey C, Schwab F, Zebala LP, 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)** 34:2171–2178, 2009
- Carreon LY, Glassman SD, Shaffrey CI, Fehlings MG, Dahl B, Ames CP, et al: Predictors of health-related quality-of-life after complex adult spinal deformity surgery: a Scolio-RISK-1 secondary analysis. **Spine Deform** 5:139–144, 2017
- Carreon LY, Sanders JO, Diab M, Sucato DJ, Sturm PF, Glassman SD: The minimum clinically important difference in Scoliosis Research Society-22 Appearance, Activity, and Pain domains after surgical correction of adolescent idiopathic scoliosis. **Spine (Phila Pa 1976)** 35:2079–2083, 2010
- Champain S, Benchikh K, Nogier A, Mazel C, Guise JD, Skalli W: Validation of new clinical quantitative analysis software applicable in spine orthopaedic studies. **Eur Spine J** 15:982–991, 2006
- Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY: Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. **Spine J** 8:968–974, 2008
- 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** 90:1839–1847, 2008
- Leven DM, Lee NJ, Kothari P, Steinberger J, Guzman J, Skovrlj B, et al: Frailty index is a significant predictor of complications and mortality after surgery for adult spinal deformity. **Spine (Phila Pa 1976)** 41:E1394–E1401, 2016
- Liu S, Schwab F, Smith JS, Klineberg E, Ames CP, Mundis G, et al: Likelihood of reaching minimal clinically important difference in adult spinal deformity: a comparison of operative and nonoperative treatment. **Ochsner J** 14:67–77, 2014
- McCarthy I, O'Brien M, Ames C, Robinson C, Errico T, Polly DW Jr, et al: Incremental cost-effectiveness of adult spinal deformity surgery: observed quality-adjusted life years with surgery compared with predicted quality-adjusted life years without surgery. **Neurosurg Focus** 36(5):E3, 2014
- O'Brien MF, Kuklo TR, Blanke K, Lenke L (eds): **Spinal Deformity Study Group Radiographic Measurement Manual**. Memphis: Medtronic Sofamor Danek, 2005
- Oh T, Lafage R, Lafage V, Protosaltis T, Challier V, Shaffrey C, et al: Comparing quality of life in cervical spondylotic myelopathy with other chronic debilitating diseases using the SF-36 survey. **World Neurosurg** 106:699–706, 2017
- Osorio JA, Scheer JK, Ames CP: Predictive modeling of complications. **Curr Rev Musculoskelet Med** 9:333–337, 2016
- Protosaltis T, Schwab F, Bronsard N, Smith JS, Klineberg E, Mundis G, et al: The T1 pelvic angle, a novel radiographic measure of global sagittal deformity, accounts for both spinal inclination and pelvic tilt and correlates with health-related quality of life. **J Bone Joint Surg Am** 96:1631–1640, 2014
- Rillardon L, Levassor N, Guigui P, Wodecki P, Cardinne L, Templier A, et al: [Validation of a tool to measure pelvic and spinal parameters of sagittal balance.] **Rev Chir Orthop Repar Appar Mot** 89:218–227, 2003 (Fr)
- Scheer JK, Ailon TT, Smith JS, Hart R, Burton DC, Bess S, et al: Predictive modeling of length of hospital stay following adult spinal deformity correction: analysis of 653 patients with an accuracy of 75% within 2 days. **Neurosurgery** 63 (Suppl 1):166–167, 2016 (Abstract)
- Scheer JK, Hostin R, Robinson C, Schwab F, Lafage V, Burton DC, et al: Operative management of adult spinal deformity results in significant increases in QALYs gained compared to non-operative management: analysis of 479 patients with minimum 2-year follow-up. **Spine (Phila Pa 1976)** [epub ahead of print], 2016
- Scheer JK, Osorio JA, Smith JS, Schwab F, Lafage V, Hart RA, et al: Development of validated computer-based preop-

- erative predictive model for proximal junction failure (PJF) or clinically significant PJK with 86% accuracy based on 510 ASD patients with 2-year follow-up. **Spine (Phila Pa 1976)** **41**:E1328–E1335, 2016
29. Scheer JK, Smith JS, Clark AJ, Lafage V, Kim HJ, Rolston JD, et al: Comprehensive study of back and leg pain improvements after adult spinal deformity surgery: analysis of 421 patients with 2-year follow-up and of the impact of the surgery on treatment satisfaction. **J Neurosurg Spine** **22**:540–553, 2015
 30. Scheer JK, Smith JS, Schwab F, Lafage V, Shaffrey CI, Bess S, et al: Development of a preoperative predictive model for major complications following adult spinal deformity surgery. **J Neurosurg Spine** **26**:736–743, 2017
 31. Scheer JK, Tang JA, Smith JS, Acosta FL Jr, Protosaltis TS, Blondel B, et al: Cervical spine alignment, sagittal deformity, and clinical implications: a review. **J Neurosurg Spine** **19**:141–159, 2013
 32. Schwab F, Ungar B, Blondel B, Buchowski J, Coe J, Deinlein D, et al: Scoliosis Research Society-Schwab adult spinal deformity classification: a validation study. **Spine (Phila Pa 1976)** **37**:1077–1082, 2012
 33. Smith JS, Lafage V, Shaffrey CI, Schwab F, Lafage R, Hostin R, 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** **78**:851–861, 2016
 34. Smith JS, Line B, Bess S, Shaffrey CI, Kim HJ, Mundis G, et al: The health impact of adult cervical deformity in patients presenting for surgical treatment: comparison to United States population norms and chronic disease states based on the EuroQuol-5 Dimensions Questionnaire. **Neurosurgery** **80**:716–725, 2017
 35. Smith JS, Shaffrey CI, Berven S, Glassman S, Hamill C, Horton W, et al: Operative versus nonoperative treatment of leg pain in adults with scoliosis: a retrospective review of a prospective multicenter database with two-year follow-up. **Spine (Phila Pa 1976)** **34**:1693–1698, 2009
 36. Smith JS, Shaffrey CI, Fu KM, Scheer JK, Bess S, Lafage V, et al: Clinical and radiographic evaluation of the adult spinal deformity patient. **Neurosurg Clin N Am** **24**:143–156, 2013
 37. Smith JS, Shaffrey CI, Glassman SD, Carreon LY, Schwab FJ, Lafage V, et al: Clinical and radiographic parameters that distinguish between the best and worst outcomes of scoliosis surgery for adults. **Eur Spine J** **22**:402–410, 2013
 38. Smith JS, Shaffrey CI, Lafage V, Schwab F, Scheer JK, Protosaltis T, et al: Comparison of best versus worst clinical outcomes for adult spinal deformity surgery: a retrospective review of a prospectively collected, multicenter database with 2-year follow-up. **J Neurosurg Spine** **23**:349–359, 2015
 39. Smith MD, Drummond M, Brixner D: Moving the QALY forward: rationale for change. **Value Health** **12 (Suppl 1)**:S1–S4, 2009
 40. Terran J, McHugh BJ, Fischer CR, Lonner B, Warren D, Glassman S, et al: Surgical treatment for adult spinal deformity: projected cost effectiveness at 5-year follow-up. **Ortho J** **14**:14–22, 2014
 41. Wang MC, Kreuter W, Wolfa CE, Maiman DJ, Deyo RA: Trends and variations in cervical spine surgery in the United States: Medicare beneficiaries, 1992 to 2005. **Spine (Phila Pa 1976)** **34**:955–963, 2009
 42. Weinstein MC, Torrance G, McGuire A: QALYs: the basics. **Value Health** **12 (Suppl 1)**:S5–S9, 2009
 43. Whitehead SJ, Ali S: Health outcomes in economic evaluation: the QALY and utilities. **Br Med Bull** **96**:5–21, 2010
- study was conducted, is received through research grants from DePuy Spine and individual donations.
- Dr. Ames is a consultant for Stryker, Medtronic, and DePuy; is a patent holder with Fish & Richardson, P.C.; and receives royalties from Stryker and Biomet Spine. Dr. Bess is a consultant for K2M and AlloSource; is a patent holder with K2M; receives clinical or research support for the study described (includes equipment or material) from DePuy Synthes, NuVasive, and K2M; and receives support of a non-study-related clinical or research effort that he oversees from Medtronic, Biomet, and Orthofix. Dr. Burton receives clinical or research support for the study described (includes equipment or material) from DePuy, and he is a patent holder with DePuy. Dr. Gum is a consultant for Medtronic, DePuy, Alphatec, Stryker, Acuity, Pacmed (ended January 2015), K2M, and Nuvasive; received travel funds from the Fischer Owen fund (ended December 31, 2016); and receives research funding from Norton Healthcare. Funds were given only to his institution by Integra, Intellirod Spine, Inc., Pfizer, and the ISSG. NuVasive supplied database support only. Dr. Hostin is a consultant for DePuy Spine, and he receives support of a non-study-related clinical or research effort that he oversees from NuVasive, Seeger, DJO, DePuy Spine, and K2M. Dr. Klineberg is a consultant for DePuy, Stryker, Trevena, and Springer. He receives honoraria from K2M and AOSpine, as well as a fellowship grant from AOSpine. Dr. Lafage has direct stock ownership in Nemaris, Inc.; is a consultant for NuVasive; receives support of a non-study-related clinical or research effort that she oversees from DePuy Spine, NuVasive, K2M, SRS, and Stryker (paid through the ISSGF); and has speaking and/or teaching arrangements with DePuy Spine, AO, and Medtronic Sofamor Danek. Dr. Protosaltis is a consultant for Medtronic, NuVasive, Globus, and Innovaxis, and he receives support of a non-study-related clinical or research effort that he oversees from Zimmer Biomet and the Cervical Spine Research Society. Dr. Schwab has direct stock ownership in Nemaris, Inc.; receives support of a non-study-related clinical or research effort that he oversees from DePuy Spine, K2M, Stryker, and NuVasive (grants paid through the ISSGF); is a consultant for Zimmer-Biomet, Medtronic Sofamor Danek, K2M, and NuVasive; and receives grants from Zimmer Biomet, Medtronic Sofamor Danek, K2M, and NuVasive. Dr. Shaffrey has direct stock ownership in NuVasive; is a patent holder with NuVasive, Medtronic, and Zimmer Biomet; is a consultant for NuVasive, Medtronic, K2M, and Stryker; and receives support of a non-study-related clinical or research effort that he oversees from DePuy Synthes, NuVasive, Medtronic, and K2M. Dr. Smith is a consultant for Zimmer Biomet, NuVasive, K2M, and AlloSource; receives support of a non-study-related clinical or research effort that he oversees from DePuy Synthes/ISSG; receives clinical or research support for the study described (includes equipment or material) from DePuy Synthes/ISSG; receives fellowship funding from the NREF and AOSpine; and receives royalties from Zimmer Biomet.

Author Contributions

Conception and design: Scheer. Acquisition of data: Scheer, Smith, Hostin, Robinson, Gum, Schwab, Hart, Lafage, Burton, Bess, Protosaltis, Klineberg, Shaffrey, Ames. Analysis and interpretation of data: all authors. Drafting the article: all authors. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Oh. Statistical analysis: Oh, Scheer, Lafage.

Correspondence

Taemin Oh, Department of Neurological Surgery, University of California, San Francisco, 505 Parnassus Ave., M-779, San Francisco, CA 94143. email: taemin.oh@ucsf.edu.

Disclosures

Funding for the ISSG Foundation (ISSGF), through which this