

## Development of a validated computer-based preoperative predictive model for pseudarthrosis with 91% accuracy in 336 adult spinal deformity patients

Justin K. Scheer, MD,<sup>1</sup> Taemin Oh, MD,<sup>1</sup> Justin S. Smith, MD, PhD,<sup>2</sup> Christopher I. Shaffrey, MD,<sup>2</sup> Alan H. Daniels, MD,<sup>3</sup> Daniel M. Sciubba, MD,<sup>4</sup> D. Kojo Hamilton, MD,<sup>5</sup> Themistocles S. Protopsaltis, MD,<sup>6</sup> Peter G. Passias, MD,<sup>6</sup> Robert A. Hart, MD, MA, MHCDS,<sup>7</sup> Douglas C. Burton, MD,<sup>8</sup> Shay Bess, MD,<sup>9</sup> Renaud Lafage, MS,<sup>10</sup> Virginie Lafage, PhD,<sup>10</sup> Frank Schwab, MD,<sup>10</sup> Eric O. Klineberg, MD,<sup>11</sup> Christopher P. Ames, MD,<sup>1</sup> and the International Spine Study Group

<sup>1</sup>Department of Neurological Surgery, University of California, San Francisco, California; <sup>2</sup>Department of Neurosurgery, University of Virginia Health System, Charlottesville, Virginia; <sup>3</sup>Department of Orthopaedic Surgery, Brown University, Providence, Rhode Island; <sup>4</sup>Department of Neurosurgery, Johns Hopkins University, Baltimore, Maryland; <sup>5</sup>Department of Neurosurgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; <sup>6</sup>Department of Orthopaedic Surgery, NYU Hospital for Joint Diseases, New York, New York; <sup>7</sup>Department of Orthopaedic Surgery, Swedish Medical Center, Seattle, Washington; <sup>8</sup>Department of Orthopaedic Surgery, University of Kansas Medical Center, Kansas City, Kansas; <sup>9</sup>Presbyterian/St. Luke's Medical Center, Denver, Colorado; <sup>10</sup>Department of Orthopaedic Surgery, Hospital for Special Surgery, New York, New York; and <sup>11</sup>Department of Orthopaedic Surgery, University of California, Davis, California

**OBJECTIVE** Pseudarthrosis can occur following adult spinal deformity (ASD) surgery and can lead to instrumentation failure, recurrent pain, and ultimately revision surgery. In addition, it is one of the most expensive complications of ASD surgery. Risk factors contributing to pseudarthrosis in ASD have been described; however, a preoperative model predicting the development of pseudarthrosis does not exist. The goal of this study was to create a preoperative predictive model for pseudarthrosis based on demographic, radiographic, and surgical factors.

**METHODS** A retrospective review of a prospectively maintained, multicenter ASD database was conducted. Study inclusion criteria consisted of adult patients (age  $\geq 18$  years) with spinal deformity and surgery for the ASD. From among 82 variables assessed, 21 were used for model building after applying collinearity testing, redundancy, and univariable predictor importance  $\geq 0.90$ . Variables included demographic data along with comorbidities, modifiable surgical variables, baseline coronal and sagittal radiographic parameters, and baseline scores for health-related quality of life measures. Patients groups were determined according to their Lenke radiographic fusion type at the 2-year follow-up: bilateral or unilateral fusion (union) or pseudarthrosis (nonunion). A decision tree was constructed, and internal validation was accomplished via bootstrapped training and testing data sets. Accuracy and the area under the receiver operating characteristic curve (AUC) were calculated to evaluate the model.

**RESULTS** A total of 336 patients were included in the study (nonunion: 105, union: 231). The model was 91.3% accurate with an AUC of 0.94. From 82 initial variables, the top 21 covered a wide range of areas including preoperative alignment, comorbidities, patient demographics, and surgical use of graft material.

**CONCLUSIONS** A model for predicting the development of pseudarthrosis at the 2-year follow-up was successfully created. This model is the first of its kind for complex predictive analytics in the development of pseudarthrosis for patients with ASD undergoing surgical correction and can aid in clinical decision-making for potential preventative strategies.

<https://thejns.org/doi/abs/10.3171/2018.8.FOCUS18246>

**KEYWORDS** pseudarthrosis; adult spinal deformity; predictive model; complications; ASD; outcomes; scoliosis; sagittal malalignment

**ABBREVIATIONS** ASD = adult spinal deformity; AUC = area under the receiver operating characteristic curve; BMP = bone morphogenetic protein; HRQOL = health-related quality of life; LIV = lowermost instrumented vertebra; LL = lumbar lordosis; NRS = numeric rating scale; PI-LL = mismatch between pelvic incidence and LL; PT = pelvic tilt; SRS = Scoliosis Research Society; SVA = sagittal vertical axis; TK = thoracic kyphosis.

**SUBMITTED** May 20, 2018. **ACCEPTED** August 13, 2018.

**INCLUDE WHEN CITING** DOI: 10.3171/2018.8.FOCUS18246.

**A**DULT spinal deformity (ASD) surgery remains technically challenging and is associated with high rates of complications, one of which is pseudarthrosis.<sup>8–11,22,27</sup> The rates of pseudarthrosis in ASD have ranged from 0% to 35%,<sup>5,8–10</sup> with one comprehensive review citing pseudarthrosis as the most frequent long-term complication of those studied, reporting a rate of 7.6%.<sup>27</sup> Patients who develop pseudarthrosis are at risk for instrumentation failure and may require revision surgery. Risk factors for pseudarthrosis have been studied and include preoperative thoracolumbar kyphosis > 20°, age of more than 55 years, arthrodesis to S1 as compared to L5 or another more cephalad level, and arthrodesis of more than 12 vertebrae.<sup>8</sup>

Despite these results, however, it is very difficult to preoperatively predict which patients may develop pseudarthrosis since bone healing is a multifactorial process. Recently, there has been an attempt to use patient-reported health-related quality of life (HRQOL) outcome measures as a surrogate to predict the development of pseudarthrosis.<sup>11</sup> The authors determined that at 1 year after ASD surgery, patients with pseudarthrosis had significantly lower Scoliosis Research Society (SRS) scores and Oswestry Disability Index scores than those with union. Currently, there is no model for accurately predicting which patients are likely to develop pseudarthrosis after ASD correction by the 2-year follow-up. Such a model could be very beneficial to many stakeholders, including the surgeon, the patient, and even third-party payers in their attempts to determine payments and hospital length of stay calculations. More specifically, a predictive model could not only aid in the preoperative discussion with patients about whether or not to pursue surgery, but could also assist in the preoperative planning, which may include protective measures such as the use of various graft materials, adjustment of the uppermost instrumented vertebra, use of interbody fusion, etc. Furthermore, risk stratification and the likelihood of revision are other parameters that a model could be especially useful in predicting.

Accurate advanced predictive modeling has only recently come to the forefront of spinal surgery outcomes research.<sup>4,7,12,14,19–21,24,25,28</sup> Such models provide patient-specific predictions for a given outcome. In an effort to extend these findings, we endeavored to build a preoperative predictive model for pseudarthrosis at 2 years postoperatively by using a large, multicenter, prospective ASD database.

## Methods

### Patient Population

This study is a retrospective review of a multicenter, prospective ASD database, which is contributed to by 13 sites across the United States. All patients were enrolled into a protocol for which each site had obtained institutional review board approval. Inclusion criteria for the database were as follows: age > 18 years; the presence of spinal deformity as defined by a scoliosis Cobb angle  $\geq$  20°, sagittal vertical axis (SVA)  $\geq$  5 cm, pelvic tilt (PT)  $\geq$  25°, and/or thoracic kyphosis (TK)  $\geq$  60°; and surgery for the ASD. Exclusion criteria consisted of spinal deformity with a neuromuscular etiology and the presence of

active infection or malignancy. In addition to applying the above database inclusion criteria, we excluded study patients if they had a revision for any indication other than pseudarthrosis in order to reduce confounding of potential pseudarthrosis as a result of the revision surgery (n = 87). The goal of the model was a preoperative risk stratification assessment, and having patients who underwent a revision surgery includes postoperative data.

### Data Collection: Demographics, Radiographic Assessment, HRQOL, and Surgical Data

The demographic and clinical data collected included patient age, sex, BMI, race, number and types of comorbidities, and Charlson Comorbidity Index.<sup>3</sup> Collected surgical data included whether the index surgery was a primary or revision procedure, whether the procedure was an anterior or a posterior fusion (and the number of levels for each), the presence of a 3-column osteotomy (pedicle subtraction osteotomy or vertebral column resection), the uppermost instrumented vertebra, the lowermost instrumented vertebra (LIV), the number of posterior levels fused, and the ASD surgical invasiveness score.<sup>17</sup> The presence of (yes/no) and the number of levels were also noted for the following: direct decompression, Smith-Petersen osteotomy, and interbody fusion. Number of levels for anterior lumbar interbody fusion, posterior lumbar interbody fusion, transforaminal lumbar interbody fusion, lateral interbody fusion, as well as graft type for both posterior and interbody fusions (autograft, allograft, and bone morphogenetic protein [BMP]), were also collected. Of note, BMP use also included off-label applications.

Baseline full-length free-standing lateral spine radiographs (36-inch cassette) were analyzed using validated software<sup>2,18,23</sup> (SpineView, ENSAM, Laboratory of Biomechanics). All radiographic measures were performed at a central location based on standard techniques<sup>18</sup> and included coronal Cobb angles of thoracic and lumbar curves, maximum coronal Cobb angle, coronal plumbline (C7–S1), TK (T4–12; Cobb angle between superior endplate of T4 and inferior endplate of T12), lumbar lordosis (LL; Cobb angle between superior endplate of L1 and superior endplate of S1), SVA (C7 plumbline relative to S1), PT, and the mismatch between pelvic incidence and LL (PI-LL). The SRS-Schwab coronal curve type and sagittal modifiers were determined for all patients.<sup>26</sup> Posterolateral fusion was graded based on the Lenke fusion grading system,<sup>13</sup> which involves four tiers ranging from A to D:<sup>13</sup> type A represents a solid fusion bilaterally, type B is a solid unilateral fusion with a small contralateral aspect, type C is small thin fusion masses bilaterally with a possible crack, and type D is graft resorption bilaterally or bilateral pseudarthrosis.<sup>13</sup> Patients with type A or B at 2 years were considered to have union and those with type C or D were labeled as having nonunion. Final typing was based on consensus review by multiple fellowship-trained attending spine surgeons.

Standardized HRQOL measures included the Oswestry Disability Index, SF-36, and 22-item SRS Outcomes Questionnaire (SRS-22r). Two standard summary scores were calculated based on the SF-36, the Physical Component Summary and the Mental Component Summary,

and the ASD frailty index was calculated.<sup>15,16</sup> The SRS-22r provides a total score and multiple subdomain scores including those for activity, pain, appearance, mental, and satisfaction. A numeric rating scale (NRS) score ranging from 0 (no pain) to 10 (most unbearable pain) was collected for back and leg pain separately.

### Statistics and Predictive Model Construction

Continuous variables were described with the mean and standard deviation. Baseline variables were compared between the union and nonunion groups. Normality of data was determined using the Shapiro-Wilk test. Comparison of baseline means between the groups initially included an ANOVA or the Kruskal-Wallis test when appropriate, which was followed by pairwise comparisons using Tukey's honestly significant difference test to control for type I error or Wilcoxon rank-sum tests where appropriate. Frequency analyses for categorical variables were conducted via Pearson's chi-square analysis. All statistical analyses were conducted using commercially available software (SPSS Statistics 22, IBM Corp.), and the level of significance was set at  $p < 0.05$  for all tests.

For the predictive model, missing values within the database were imputed using standard techniques, including mean and median imputation.<sup>1</sup> Of the 82 variables used, 23 (28.0%) required some imputation (the other 59 variables were complete), and of these 23, the number of values imputed ranged from 1 (0.30%) to 22 (6.5%). Generally, imputation below 10% for a given value is accepted.<sup>1</sup> Once a complete data set was constructed, a supervised ensemble of decision trees<sup>1</sup> was constructed with the target variable being binary: (1) patients who had developed pseudarthrosis (nonunion) and (0) those who did not (union), based on the Lenke radiographic fusion grade<sup>13</sup> at the 2-year follow-up. A decision tree was constructed, and internal validation was accomplished via bootstrapped training and testing data sets using a random number generator. The decision algorithm chosen was the C5.0 algorithm, which uses gain ratio as the splitting criterion. This was selected as it can handle categorical and continuous variables simultaneously, as well as have binary outcomes for splits or multiway splits for categorical variables with multiple categories. A total of 5 leaf nodes per model was selected, and each leaf node had a minimum of 10 entries before moving on to the next leaf or that leaf would stop. For the ensemble of the decision trees, final overall predictions from the models were combined and chosen by voting within the program, and random selection was used for the predictions that had tied votes. Overall accuracy and the area under the receiver operating characteristic curve (AUC) were calculated. The model was built using commercially available software (SPSS Modeler v16, IBM Corp.).

## Results

### Patient Population

There were 603 potentially eligible patients, and 423 (70.1%) had a complete 2-year follow-up. Among the latter group, a total of 336 (79.4%) patients met the additional inclusion criterion of not having revision surgery for an in-

dication other than pseudarthrosis (87 patients excluded). This final cohort consisted of 268 (79.8%) women and 68 (20.2%) men, whose mean age was  $57.7 \pm 15.1$  years and mean BMI was  $27.2 \pm 5.8$  kg/m<sup>2</sup> (Table 1). Two hundred thirty-one (68.8%) patients were graded as having fusion at 2 years (union), and 105 (31.2%) patients were graded as having pseudarthrosis (nonunion). There was a significantly higher proportion of smokers in the nonunion group than in the union group (12.4% vs 5.6%, respectively,  $p = 0.038$ ), and the nonunion group was significantly younger than the union group on average ( $54.5 \pm 17.3$  vs  $59.1 \pm 13.7$  years, respectively,  $p = 0.042$ ). All other demographic variables were not statistically different between the groups ( $p > 0.05$  for all).

### HRQOL

The average baseline HRQOL scores were not statistically different between the groups ( $p > 0.05$ ) with the exception of the mean NRS leg pain score, which was significantly higher in the union group than in the nonunion group ( $4.9 \pm 3.2$  vs  $4 \pm 3.3$ ,  $p = 0.015$ ; Table 2). Eight patients with pseudarthrosis required reoperation for it. Of these 8 patients, 3 had a prior fusion and 1 had a prior diagnosis of pseudarthrosis.

### Radiographic Assessment

Of the baseline radiographic parameters, the mean PT ( $24.1^\circ \pm 10.2^\circ$  vs  $21.9^\circ \pm 11.6^\circ$ ,  $p = 0.040$ ), PI-LL ( $16.8^\circ \pm 20.7^\circ$  vs  $11.7^\circ \pm 23.5^\circ$ ,  $p = 0.031$ ), and SVA ( $65.9 \pm 68.6$  vs  $53.2 \pm 81.6$  mm,  $p = 0.023$ ; Table 3) were significantly higher in the union group. All other parameters were not statistically different between the groups ( $p > 0.05$ ).

### Surgical Data

Univariate analysis of the surgical data demonstrated multiple statistically significant differences (Table 4). The union group had a higher proportion of patients undergoing revision surgery ( $p = 0.042$ ). The mean ASD surgical invasiveness score<sup>17</sup> was not statistically different between the two groups, and neither was the proportion of patients that underwent anterior and posterior fusions, the mean number of posterior levels fused, or the distribution of the uppermost instrumented vertebra ( $p > 0.05$  for all). The distribution of the LIV in the union group was significantly shifted more toward the lower vertebral segments such as sacroiliac fixation, whereas the nonunion group had more thoracic and lumbar LIV ( $p = 0.001$ ). Regarding the use of posterior graft material, the union group had a significantly higher proportion of patients with BMP ( $p = 0.000$ ). In contrast, the nonunion group had a significantly higher proportion of patients having an interbody allograft ( $p = 0.005$ ). The distribution of the posterior rod diameter was significantly different, with the union group having a higher proportion of patients with larger diameter rods, such as 6.0 and 6.35 mm, than the nonunion group ( $p = 0.017$ ). However, the distribution of rod material was not statistically different between the groups ( $p > 0.05$ ). For the surgical techniques, the union group had a significantly larger proportion of patients that underwent direct decompression ( $p = 0.006$ ) and a larger mean number of levels

**TABLE 1. Demographic information and baseline SRS-Schwab coronal classification for patients who underwent surgery for ASD**

Variable	All Patients	Union Group	Nonunion Group	p Value
No. of patients	336	231	105	NA
Age in yrs	57.7 ± 15.1	59.1 ± 13.7	54.5 ± 17.3	<b>0.042</b>
F/M	268:68	178:53	90:15	
Race, no. (%)				
Asian	1 (0.3%)	1 (0.4%)	0 (0%)	0.514
Black	14 (4.2%)	8 (3.5%)	6 (5.7%)	
Hispanic	6 (1.8%)	4 (1.7%)	2 (1.9%)	
White	312 (92.9%)	217 (93.9%)	95 (90.5%)	
Other	3 (0.9%)	1 (0.4%)	2 (1.9%)	
BMI in kg/m <sup>2</sup>	27.2 ± 5.8	27.5 ± 5.9	26.5 ± 5.5	0.119
CCI	1.6 ± 1.8	1.7 ± 1.8	1.5 ± 1.6	0.553
Min of 1 comorbidity, no. (%), range of comorbidities)	241 (71.7%, 1–9)	161 (69.7%, 1–9)	80 (76.2%, 1–9)	0.216
Osteoporosis, no. (%)	46 (13.7%)	26 (11.3%)	20 (19.0%)	0.060
Smoker, no. (%)	26 (7.7%)	13 (5.6%)	13 (12.4%)	<b>0.038</b>
Frailty index	3.2 ± 1.7	3.2 ± 1.7	3.2 ± 1.7	0.692
Preop SRS-Schwab coronal curve,* no. (%)				
Type N	102 (30.4%)	73 (31.6%)	29 (27.6%)	0.280
Type T	14 (4.2%)	7 (3.0%)	7 (6.7%)	
Type L	121 (36.0%)	87 (37.7%)	34 (32.4%)	
Type D	99 (29.5%)	64 (27.7%)	35 (33.3%)	

CCI = Charlson Comorbidity Index; NA = not applicable.

Values presented as the mean ± 1 standard deviation or as number of patients (%), unless indicated otherwise. Boldface type indicates a statistically significant difference between the union and nonunion groups.

\* Type N = patients with no coronal curve greater than 30° (i.e., no major coronal deformity); type T = patients with a thoracic major curve of greater than 30° (apical level of T9 or higher); type L = patients with a lumbar or thoracolumbar major curve greater than 30° (apical level of T10 or lower); and type D = patients with a double major curve, with each curve greater than 30°.

**TABLE 2. Preoperative HRQOL scores for patients who underwent surgery for ASD**

Measure	All Patients	Union Group	Nonunion Group	p Value
ODI	42.2 ± 18.7	42.2 ± 18.4	42.2 ± 19.5	0.929
PCS	32.4 ± 10.2	32 ± 10.3	33.3 ± 9.9	0.145
MCS	45.4 ± 14	45.2 ± 14.2	45.9 ± 13.7	0.675
SRS activity	3 ± 0.9	2.9 ± 0.9	3 ± 0.9	0.160
SRS pain	2.5 ± 0.8	2.5 ± 0.9	2.4 ± 0.8	0.880
SRS appearance	2.5 ± 0.8	2.5 ± 0.8	2.4 ± 0.7	0.919
SRS mental	3.4 ± 0.9	3.4 ± 1	3.5 ± 0.9	0.883
SRS satisfaction	2.8 ± 1	2.8 ± 1	2.7 ± 1	0.484
SRS total	2.8 ± 0.7	2.8 ± 0.7	2.8 ± 0.7	0.673
NRS back pain	7.1 ± 2.1	7 ± 2.2	7.2 ± 2.1	0.620
NRS leg pain	4.6 ± 3.3	4.9 ± 3.2	4 ± 3.3	<b>0.015</b>

MCS = Mental Component Summary score of the SF-36; ODI = Oswestry Disability Index; PCS = Physical Component Summary score of the SF-36.

Values presented as the mean ± 1 standard deviation, unless indicated otherwise. Boldface type indicates a statistically significant difference between the union and nonunion groups.

**TABLE 3. Preoperative radiographic parameters included in the model for patients surgically treated for ASD**

Parameter	All Patients	Union Group	Nonunion Group	p Value
Coronal PL (mm)	35.1 ± 32.4	35.2 ± 33.4	34.9 ± 30.4	0.860
PT (°)	23.4 ± 10.7	24.1 ± 10.2	21.9 ± 11.6	<b>0.040</b>
PI-LL (°)	15.2 ± 21.7	16.8 ± 20.7	11.7 ± 23.5	<b>0.031</b>
TK (°)	32.7 ± 18	32.2 ± 16.6	33.7 ± 20.7	0.919
SVA (mm)	62 ± 73	65.9 ± 68.6	53.2 ± 81.6	<b>0.023</b>
Max Cobb angle				
<30°	105 (31.3%)	78 (33.8%)	27 (25.7%)	0.302
30°–60°	161 (47.9%)	108 (46.8%)	53 (50.5%)	
>60°	70 (20.8%)	45 (19.5%)	25 (23.8%)	

PL = plumbline.

Values expressed as mean ± 1 standard deviation or as number (%), unless indicated otherwise. Boldface type indicates a statistically significant difference between the union and nonunion groups.

**TABLE 4. Surgical data for all patients, the union, and the nonunion groups**

Surgical Data	All Patients	Union Group	Nonunion Group	p Value
No. of patients	336	231	105	NA
Revision procedure, no. of patients (%)	119 (35.4%)	90 (39.0%)	29 (27.6%)	<b>0.042</b>
Mean surgical invasiveness score	92.4 ± 36.6	92.6 ± 36.5	92 ± 36.9	0.933
Ant fusion, no. of patients (%)	4 (1.2%)	1 (0.4%)	3 (2.9%)	0.070
Pst fusion, no. of patients (%)	335 (99.7%)	231 (100%)	104 (99.0%)	0.127
Mean no. of pst levels fused	11.3 ± 4.2	11.3 ± 4.2	11.3 ± 4.4	0.647
UIV category				
T1–6	161 (47.9%)	103 (44.6%)	58 (55.2%)	0.171
T7–12	137 (40.8%)	99 (42.9%)	38 (36.2%)	
Lumbar	38 (11.3%)	29 (12.6%)	9 (8.6%)	
LIV category, no. of patients (%)				
Thoracic	8 (2.4%)	3 (1.3%)	5 (4.8%)	<b>0.001</b>
Lumbar	77 (22.9%)	42 (18.2%)	35 (33.3%)	
Sacroiliac	251 (74.7%)	186 (80.5%)	65 (61.9%)	
Pst fusion graft types, no. of patients (%)				
Allograft	244 (72.6%)	165 (71.4%)	79 (75.2%)	0.465
Autograft	257 (76.5%)	178 (77.1%)	79 (75.2%)	0.717
BMP	241 (71.7%)	182 (78.8%)	59 (56.2%)	<b>0.000</b>
Pst rod diameter, no. of patients (%)				
4.5 mm	1 (0.3%)	0 (0%)	1 (1.0%)	<b>0.017</b>
4.75 mm	0 (0%)	0 (0%)	0 (0%)	
5.5 mm	219 (65.2%)	140 (60.6%)	79 (75.2%)	
6.0 mm	30 (8.9%)	24 (10.4%)	6 (5.7%)	
6.35 mm	86 (25.6%)	67 (29.0%)	19 (18.1%)	
Pst rod material, no. of patients (%)				
Cobalt chrome	192 (57.1%)	132 (57.1%)	60 (57.1%)	0.440
Stainless steel	79 (23.5%)	58 (25.1%)	21 (20.0%)	
Titanium	64 (19.0%)	40 (17.3%)	24 (22.9%)	
Other	1 (0.3%)	1 (0.4%)	0 (0%)	
Decompression, no. of patients (%)	212 (63.1%)	157 (68.0%)	55 (52.4%)	<b>0.006</b>
Mean no. of decompression levels	0.7 ± 0.6	0.8 ± 0.6	0.6 ± 0.6	<b>0.008</b>
SPO, no. of patients (%)	190 (56.5%)	136 (58.9%)	54 (51.4%)	0.203
Mean no. of SPO levels	3 ± 3.2	3 ± 3.1	2.9 ± 3.4	0.479
3-column osteotomy, no. of patients (%)	63 (18.8%)	47 (20.3%)	16 (15.2%)	0.259
Mean no. of 3-column osteotomy levels	0.2 ± 0.4	0.2 ± 0.4	0.2 ± 0.4	0.255
IBF, no. of patients (%)	205 (61.0%)	141 (61.0%)	64 (61.0%)	0.988
Mean no. of IBF levels	1.5 ± 1.7	1.4 ± 1.6	1.5 ± 1.7	0.883
IBF graft types, no. of patients (%)				
IBF allograft*	68 (20.2%)	37 (16.0%)	31 (29.5%)	<b>0.005</b>
IBF autograft	112 (33.3%)	81 (35.1%)	31 (29.5%)	0.315
IBF BMP	107 (31.8%)	83 (35.9%)	24 (22.9%)	<b>0.015</b>

Ant = anterior; IBF = interbody fusion; pst = posterior; SPO = Smith-Petersen osteotomy; UIV = uppermost instrumented vertebra. Values presented as the mean ± 1 standard deviation, unless indicated otherwise. Boldface type indicates a statistically significant difference between the union and nonunion groups.

\* Percentages calculated based on total number of patients for each group.

decompressed ( $p = 0.008$ ). The proportion of patients undergoing Smith-Petersen osteotomies, 3-column osteotomies, and interbody fusions as well as the mean number of levels for all three of the procedures were all statistically similar between the two groups ( $p > 0.05$  for all).

### Model Results

For the training data set ( $n = 336$ ), 317 patients were predicted correctly with 19 incorrect, resulting in an accuracy of 94.4% and an AUC of 0.97. The testing data set included 126 (37.5%) patients, and of those, 115 were

**TABLE 5. Twenty-one variables used in a model to preoperatively predict postoperative pseudarthrosis**

Predictor Importance	Variable
Overall	
1	Pst LIV
2	Pst BMP use (y/n)
3	Max coronal Cobb angle
4	NRS leg pain
5	Pst allograft (y/n)
6	SRS-Schwab PT modifier
7	Decompression (y/n)
8	Prior fusion
9	Pst UIV
10	SRS-Schwab GA modifier
11	Age
12	Osteotomy (y/n)
13	Pst rod diameter
14	No. of decompression levels
15	Heart disease
16	SRS-Schwab PI-LL modifier
17	Sex
18	Osteoporosis
19	BMI
20	Interbody graft BMP use (y/n)
21	No. of ant levels fused
Modifiable variables	
1	Pst LIV
2	Pst BMP use (y/n)
5	Pst allograft (y/n)
7	Decompression (y/n)
9	Pst UIV
12	Osteotomy (y/n)
13	Pst rod diameter
14	No. of decompression levels
20	Interbody graft BMP use (y/n)
19	BMI
21	No. of ant levels fused
Nonmodifiable variables	
3	Max coronal Cobb angle
4	Leg pain NRS
6	SRS-Schwab PT modifier
8	Prior fusion
10	SRS-Schwab GA modifier
11	Age
15	Heart disease
16	SRS-Schwab PI-LL modifier
17	Sex
18	Osteoporosis

GA = global alignment; n = no; y = yes.

The 21 variables used in the model appear in order of univariable importance on the left. The same variables are separated by modifiable and non-modifiable variables also in order of univariable importance, respectively. Note that the importance of these variables may change throughout the decision tree process.

predicted correctly with 11 incorrect. The testing model accuracy was 91.3% correct with an AUC of 0.94, indicating a very good model fit. The 21 variables are listed in decreasing order of importance in Table 5. The top three variables were posterior LIV, posterior BMP, and maximum Cobb angle.

## Discussion

Predictive modeling is coming to the forefront of spinal surgery outcomes research. It offers the ability to create accurate, patient-specific models that have the potential to impact clinical decision-making.<sup>19</sup> There is a paucity of literature on this topic as it is so new to the field. Clinical data sets are now becoming large enough to be amenable to modern advanced predictive algorithms, and the present study demonstrated that an accurate, internally validated predictive model for the development of pseudarthrosis at 2 years after ASD correction can be successfully created.

There are a few key points regarding our model and predictive models in general. First, it is unclear how the variables affect the given outcome because the current model is an ensemble of multiple decision trees, resulting in variables having a different role depending on which decision tree they are in and where in the tree they arise. For example, as patient data are filtered through the model, one variable may be predictive of the outcome in a given subset of patients and then not in another subset because as one moves down the tree, the subsets of patients being analyzed change. The baseline differences discussed in the *Results* are provided for educational purposes for the patient cohort. These differences are considered in the models, but ultimately the true effect of these variables is difficult to determine. When constructing predictive models, one must weigh accuracy, generalizability, and transparency to strike an acceptable balance as it is difficult to have all three equally in one model.<sup>1</sup> In our case, transparency was sacrificed for accuracy and generalizability. Additionally, these advanced computer models are not, at the moment, easily translatable to our practices as they are computer-generated. The ultimate goal is to incorporate them into a website, a mobile (i.e., tablet or phone) application, or eventually the electronic medical record. They could then be used for calculating the risk and/or benefit of developing the outcome of interest (complications, clinical success or failure, cost, etc.) in real time as a point-of-care clinical decision aid. In this context, when formulating the surgical plan and in discussion with the patient, alterations could be made to optimize the risk-benefit equation for each patient. These models are not designed or intended to replace the clinician. The surgeon is still required to evaluate the information from the model in the context of the patient and to make a decision that is best for the patient. Lastly, the present model utilized only preoperative and surgical data for model creation because the goal is to use it as an aid for preoperative risk stratification. The surgical variables were added so that the surgeon using the model in the clinic, for example, could alter them and get immediate feedback on how changes in the surgical plan would affect the possibility of developing pseudarthrosis in a specific patient.

For the present study, the details and rationale for the

specific use of decision trees, as well as predictive analytics in general, are described elsewhere,<sup>20,24</sup> but in short, some of the desirable properties of decision trees include the following: 1) ease of construction, 2) ability to incorporate continuous and categorical data, 3) capacity to handle hundreds of variables, and 4) feasibility with missing data. Furthermore, an ensemble of 5 different decision trees was used to increase the accuracy and generalizability of the model. The data were split into training and testing data sets (all bootstrapped, resulting in different patient populations in each model) to increase the validity of the model. Thus, for categorical outcomes such as developing pseudarthrosis versus being fused, this type of model algorithm works very well.

A number of studies have investigated the risk factors for pseudarthrosis in ASD;<sup>5,8-11</sup> to our knowledge, however, there is no study on a predictive model for pseudarthrosis in the spine literature. Reported risk factors for pseudarthrosis include the following:<sup>5,6,8-10</sup> preoperative thoracolumbar kyphosis > 20°; age > 50 or 55 years, depending on the study; fusion to the sacrum versus more cephalad fusion; fusion of more than 12 vertebral segments; smoking; osteoarthritis of the hip; thoracoabdominal approach; SVA > 5 cm at 8 weeks postoperatively; prior pseudarthrosis; pseudarthrosis at the pedicle subtraction osteotomy site; prior decompression in the lumbar region; prior radiation to the lumbar region; and presence of inflammatory and/or neurological disorders. In the present study, age and smoking were both statistically significant, findings in line with the literature. However, in terms of radiographic assessment, the SVA, PI-LL, and PT were all significantly higher in the union group. One might postulate that the patients with higher sagittal malalignment were more likely to be fused to the sacrum or have long segment fusions (> 12 vertebra), which would both be in line with developing pseudarthrosis. As regards the surgical parameters, LIV was significantly different between the groups, with the union group having more sacroiliac fusion, a finding counter to the literature. The present study included many surgical details, and significant differences between the groups for revision proportion and decompression (yes/no and mean number of levels) were difficult to interpret. Yet, they are important in the model based on the data. More union patients had BMP used posteriorly and placed in the interbody space, as well as had larger diameter posterior rods.

The study strengths include its multicenter design and large number of ASD patients (336 total). The multicenter design with 13 different sites across the United States allows for better generalizability of the results and the model. The study also used the complete 2-year follow-up of the patients as well as multiple variables for model building. And lastly, our model conforms to the guidelines of modern predictive analytics algorithms, allowing for a patient-specific predictive model.

Nonetheless, there are limitations to this study, including its retrospective design, which may have introduced selection or information biases. Other limitations are the fusion grading system and the determination of pseudarthrosis. We attempted to be as systematic and objective as possible in using the Lenke system and a consensus-

based review. However, it was not perfect, and dividing the groups into types A and B and types C and D might have introduced some error. We believe this is the best that can be achieved at the moment, and the model still has clinical value. And lastly, some may consider the lack of information regarding how the variables affect the outcome to be a limitation. As discussed above, however, modern complex models generally do not have great transparency. And in the future, if these models begin to take hold, transparency will not be an issue as the models will all be computer-based on a mobile device or integrated into the electronic medical record. This model is the first of its kind and adds to the growing collection of new predictive models in spine surgery outcomes to improve the quality of care for future patients.

## Conclusions

A successful model predicting the development of pseudarthrosis was built (accuracy of 91.3% with an AUC of 0.94). Of the 82 variables evaluated, the top 21 covered a wide range of areas including patient demographics, preoperative alignment, comorbidities, and surgical use of graft material. This model is the first of its kind for complex predictive analytics in the development of pseudarthrosis following ASD surgery and can aid in clinical decision-making for potential preventative strategies.

## References

- Abbott D: **Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst**. Indianapolis: John Wiley & Sons, 2014
- Champaign 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
- Charlson M, Szatrowski TP, Peterson J, Gold J: Validation of a combined comorbidity index. *J Clin Epidemiol* **47**:1245-1251, 1994
- Chen HN, Tsai YF: A predictive model for disability in patients with lumbar disc herniation. *J Orthop Sci* **18**:220-229, 2013
- Dickson DD, Lenke LG, Bridwell KH, Koester LA: Risk factors for and assessment of symptomatic pseudarthrosis after lumbar pedicle subtraction osteotomy in adult spinal deformity. *Spine (Phila Pa 1976)* **39**:1190-1195, 2014
- How NE, Street JT, Dvorak MF, Fisher CG, Kwon BK, Paquette S, et al: Pseudarthrosis in adult and pediatric spinal deformity surgery: a systematic review of the literature and meta-analysis of incidence, characteristics, and risk factors. *Neurosurg Rev* [epub ahead of print], 2018
- Ialenti MN, Lonner BS, Verma K, Dean L, Valdevit A, Errico T: Predicting operative blood loss during spinal fusion for adolescent idiopathic scoliosis. *J Pediatr Orthop* **33**:372-376, 2013
- Kim YJ, Bridwell KH, Lenke LG, Cho KJ, Edwards CC II, Rinella AS: Pseudarthrosis in adult spinal deformity following multisegmental instrumentation and arthrodesis. *J Bone Joint Surg Am* **88**:721-728, 2006
- Kim YJ, Bridwell KH, Lenke LG, Rhim S, Cheh G: Pseudarthrosis in long adult spinal deformity instrumentation and fusion to the sacrum: prevalence and risk factor analysis of 144 cases. *Spine (Phila Pa 1976)* **31**:2329-2336, 2006
- Kim YJ, Bridwell KH, Lenke LG, Rinella AS, Edwards C II: Pseudarthrosis in primary fusions for adult idiopathic

- scoliosis: incidence, risk factors, and outcome analysis. **Spine (Phila Pa 1976)** **30**:468–474, 2005
11. Klineberg E, Gupta M, McCarthy I, Hostin R: Detection of pseudarthrosis in adult spinal deformity: the use of health-related quality-of-life outcomes to predict pseudarthrosis. **Clin Spine Surg** **29**:318–322, 2016
  12. Lee MJ, Cizik AM, Hamilton D, Chapman JR: Predicting medical complications after spine surgery: a validated model using a prospective surgical registry. **Spine J** **14**:291–299, 2014
  13. Lenke LG, Bridwell KH, Bullis D, Betz RR, Baldus C, Schoenecker PL: Results of in situ fusion for isthmic spondylolisthesis. **J Spinal Disord** **5**:433–442, 1992
  14. Mathai KM, Kang JD, Donaldson WF, Lee JY, Buffington CW: Prediction of blood loss during surgery on the lumbar spine with the patient supported prone on the Jackson table. **Spine J** **12**:1103–1110, 2012
  15. Miller EK, Neuman BJ, Jain A, Daniels AH, Ailon T, Sciubba DM, et al: An assessment of frailty as a tool for risk stratification in adult spinal deformity surgery. **Neurosurg Focus** **43**(6):E3, 2017
  16. Miller EK, Vila-Casademunt A, Neuman BJ, Sciubba DM, Kebaish KM, Smith JS, et al: External validation of the adult spinal deformity (ASD) frailty index (ASD-FI). **Eur Spine J** **27**:2331–2338, 2018
  17. Neuman BJ, Ailon T, Scheer JK, Klineberg E, Sciubba DM, Jain A, et al: Development and validation of a novel adult spinal deformity surgical invasiveness score: analysis of 464 patients. **Neurosurgery** **82**:847–853, 2018
  18. O'Brien MF, Kuklo TR, Blanke K, Lenke L (eds): **Spinal Deformity Study Group Radiographic Measurement Manual**. Memphis, TN: Medtronic Sofamor Danek, 2005
  19. Oh T, Scheer JK, Smith JS, Hostin R, Robinson C, Gum JL, et al: 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. **Neurosurg Focus** **43**(6):E2, 2017
  20. Osorio JA, Scheer JK, Ames CP: Predictive modeling of complications. **Curr Rev Musculoskelet Med** **9**:333–337, 2016
  21. Passias PG, Oh C, Jalai CM, Worley N, Lafage R, Scheer JK, et al: Predictive model for cervical alignment and malalignment following surgical correction of adult spinal deformity. **Spine (Phila Pa 1976)** **41**:E1096–E1103, 2016
  22. Raizman NM, O'Brien JR, Poehling-Monaghan KL, Yu WD: Pseudarthrosis of the spine. **J Am Acad Orthop Surg** **17**:494–503, 2009
  23. 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 Reparatrice Appar Mot** **89**:218–227, 2003 (Fr)
  24. Scheer JK, Osorio JA, Smith JS, Schwab F, Lafage V, Hart RA, et al: Development of validated computer-based preoperative 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
  25. 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
  26. 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
  27. Sciubba DM, Yurter A, Smith JS, Kelly MP, Scheer JK, Goodwin CR, et al: A comprehensive review of complication rates after surgery for adult deformity: a reference for informed consent. **Spine Deform** **3**:575–594, 2015
  28. Tetreault LA, Kopjar B, Vaccaro A, Yoon ST, Arnold PM, Massicotte EM, et al: A clinical prediction model to determine outcomes in patients with cervical spondylotic myelopathy undergoing surgical treatment: data from the prospective, multi-center AOSpine North America study. **J Bone Joint Surg Am** **95**:1659–1666, 2013

## Disclosures

Dr. Smith has been a consultant for Zimmer Biomet, NuVasive, K2M, Cerapedics, and AlloSource; has received royalties from Zimmer Biomet; has received fellowship support from NREF and AOSpine; and has received support from Deputy Synthes/ISSG for non-study-related clinical or research effort. Dr. Shaffrey has been a consultant for Medtronic, NuVasive, and Zimmer-Biomet; holds patents with Medtronic, NuVasive, and Zimmer-Biomet; and has direct stock ownership in NuVasive. Dr. Daniels has been a consultant for Spineart, Stryker, and Orthofix. Dr. Sciubba has been a consultant for Medtronic, Deputy-Synthes, Stryker, NuVasive, K2M, and Baxter. Dr. Protosaltis has been a consultant for K2M, Medicea, NuVasive, and Innovaxis and has received clinical or research support from the Cervical Spine Research Society for the study described. Dr. Passias has been a consultant for Medicea and SpineWave; has received support from CSRS for non-study-related clinical or research effort; has received funds from Zimmer-Biomet, Globus, and Aesculap for speaking or teaching; and has served on the scientific advisory board for AlloSource. Dr. Burton holds a patent with DePuy, has been a consultant for AlloSource, and has received clinical or research support from DePuy and Pfizer for the study described. Dr. Lafage has direct stock ownership in Nemaris Inc., has received compensation from AOSpine and DePuy Spine for speaking or teaching arrangements, and has received support from DePuy Spine, NuVasive, K2M, Stryker (through ISSGF), SRS, and NASS for non-study-related clinical or research effort. Dr. Schwab has direct stock ownership in Nemaris Inc.; has been a consultant for Zimmer-Biomet, NuVasive, MSD, Medicea, and K2M; has received compensation from Zimmer-Biomet, K2M, NuVasive, and MSD for speaking or teaching arrangements; and has received support from DePuy Spine, Stryker, K2M, and NuVasive for non-study-related clinical or research effort. Dr. Klineberg has been a consultant for DePuy Synthes, Stryker, Springer, Trevana, and AlloSource; has received honoraria from K2M and AOSpine; and has received a fellowship grant from AOSpine. Dr. Ames has been a consultant for DePuy Synthes, Medtronic, Stryker, Medicea, and K2M; and has received royalties from Stryker, Biomet Spine, DePuy Synthes, NuVasive, and Next Orthosurgical.

## Author Contributions

Conception and design: Scheer, Ames. Acquisition of data: Daniels, Sciubba, R Lafage, V Lafage. Analysis and interpretation of data: Scheer, Daniels, Sciubba. Drafting the article: Scheer. Critically revising the article: Scheer, Oh, Smith, V Lafage, Ames. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Scheer. Statistical analysis: Scheer. Study supervision: Ames.

## Supplemental Information

### Previous Presentations

Parts of this work were previously presented at the North American Spine Society meeting held in Orlando, FL, on October 25–28, 2017.

## Correspondence

Justin K. Scheer: University of California, San Francisco, CA. justin.scheer@ucsf.edu.