Development of a preoperative predictive model for major complications following adult spinal deformity surgery

Justin K. Scheer, MD,1 Justin S. Smith, MD, PhD,2 Frank Schwab, MD,3 Virginie Lafage, PhD,3 Christopher I. Shaffrey, MD,2 Shay Bess, MD,4 Alan H. Daniels, MD,5 Robert A. Hart, MD,6 Themistocles S. Protopsaltis, MD,4 Gregory M. Mundis Jr., MD,7 Daniel M. Scuibba, MD,8 Tamir Ailon, MD,2 Douglas C. Burton, MD,9 Eric Klineberg, MD,10 Christopher P. Ames, MD,11 and the International Spine Study Group

1School of Medicine, University of California, San Diego, La Jolla, California; 2Department of Neurosurgery, University of Virginia Health System, Charlottesville, Virginia; 3Spine Service, Hospital for Special Surgery, New York; 4Department of Orthopaedic Surgery, NYU Hospital for Joint Diseases, New York, New York; 5Department of Orthopaedic Surgery, Brown University, Providence, Rhode Island; 6Department of Orthopaedic Surgery, Oregon Health & Science University, Portland, Oregon; 7Scripps Health, La Jolla, California; 8Department of Neurosurgery, The Johns Hopkins Hospital, Baltimore, Maryland; 9Department of Orthopaedic Surgery, University of Kansas Medical Center, Kansas City, Kansas; 10Department of Orthopaedic Surgery, University of California, Davis; and 11Department of Neurological Surgery, University of California, San Francisco, California

OBJECTIVE The operative management of patients with adult spinal deformity (ASD) has a high complication rate and it remains unknown whether baseline patient characteristics and surgical variables can predict early complications (intraoperative and perioperative [within 6 weeks]). The development of an accurate preoperative predictive model can aid in patient counseling, shared decision making, and improved surgical planning. The purpose of this study was to develop a model based on baseline demographic, radiographic, and surgical factors that can predict if patients will sustain an intraoperative or perioperative major complication.

METHODS This study was a retrospective analysis of a prospective, multicenter ASD database. The inclusion criteria were age ≥ 18 years and the presence of ASD. In total, 45 variables were used in the initial training of the model including demographic data, comorbidities, modifiable surgical variables, baseline health-related quality of life, and coronal and sagittal radiographic parameters. Patients were grouped as either having at least 1 major intraoperative or perioperative complication (COMP group) or not (NOCOMP group). An ensemble of decision trees was constructed utilizing the C5.0 algorithm with 5 different bootstrapped models. Internal validation was accomplished via a 70/30 data split for training and testing each model, respectively. Overall accuracy, the area under the receiver operating characteristic (AUROC) curve, and predictor importance were calculated.

RESULTS Five hundred fifty-seven patients were included: 409 (73.4%) in the NOCOMP group, and 148 (26.6%) in the COMP group. The overall model accuracy was 87.6% correct with an AUROC curve of 0.89 indicating a very good model fit. Twenty variables were determined to be the top predictors (importance ≥ 0.90 as determined by the model) and included (in decreasing importance): age, leg pain, Oswestry Disability Index, number of decompression levels, number of interbody fusion levels, Physical Component Summary of the SF-36, Scoliosis Research Society (SRS)–Schwab coronal curve type, Charlson Comorbidity Index, SRS activity, T-1 pelvic angle, American Society of Anesthesiologists grade, presence of osteoporosis, pelvic tilt, sagittal vertical axis, primary versus revision surgery, SRS pain, SRS total, use of bone morphogenetic protein, use of iliac crest graft, and pelvic incidence–lumbar lordosis mismatch.

CONCLUSIONS A successful model (87% accuracy, 0.89 AUROC curve) was built predicting major intraoperative or
Predictive model for major complications following ASD surgery

The surgical management of adult spinal deformity (ASD) can provide significant improvements in pain, disability, and health-related quality of life (HRQOL). However, these procedures are technically demanding and are associated with a high complication rate. The patient population suitable for these complicated surgeries continues to increase, including patients of advanced age. The reported complication rates in the literature are varied and range from 14% to 71%. It has been demonstrated that complication rates increase for patients undergoing revision surgery. Despite the abundant literature characterizing complication rates and the types of complications in ASD surgery, there is currently no model to predict which patients may develop complications following the surgical correction of ASD.

The ability to accurately identify these patients preoperatively constitutes a significant challenge, yet an accurate predictive model could be beneficial for both the patient and surgeon. The results of such a model could aid in the discussion between the patient and surgeon of whether to pursue a surgical intervention and, for those who opt for surgery, in adjusting the goals of surgery within the context of potential complication development. Additionally, the surgeon can plan accordingly for the operation and may employ additional techniques and/or preventative measures to potentially reduce the risk of the patient developing these complications. This patient-specific approach at complication avoidance could reduce the overall complication rate and thus potentially decrease patient morbidity.

With the advent of modern advanced predictive analytics techniques, one can now create accurate, patient-specific, predictive models with high accuracy, which can provide very useful information to aid in clinical decision making. Although traditional statistical methods can also be very clinically useful, these methods tend to be limited for use in developing patient-specific predictive models. Furthermore, they are generally designed to test specific hypotheses, have many assumptions that need to be satisfied before use, and use patient group means, not accounting for individual changes. Modern predictive modeling algorithms are very different because they can identify patterns in the data, allowing for accurate predictions without the need for a hypothesis. Thus, patient-specific models can be developed to provide valuable, detailed information, which can then be applied when discussing the risks of surgery with a patient. The goal of this study was to develop a model based on baseline demographic, radiographic, and surgical factors that could predict the patients likely to sustain a major intraoperative or perioperative complication.

Methods

Patient Population

This study is a retrospective review of a prospective multicenter ASD database, which is composed of patients from 11 sites across the US. All patients were enrolled in an IRB-approved protocol by each site. Inclusion criteria for the databases were age ≥ 18 years and the presence of spinal deformity, as defined by any coronal Cobb angle ≥ 20°, sagittal vertical axis (SVA) ≥ 5 cm, pelvic tilt (PT) ≥ 25°, or thoracic kyphosis (TK) ≥ 60°. Exclusion criteria included spinal deformity of a neuromuscular etiology and presence of active infection or malignancy.

Data Collection, Radiographic Assessment, and HRQOL

The demographic and clinical data collected included patient age, sex, body mass index (BMI), number of comorbidities, Charlson Comorbidity Index (CCI), preoperative anemia, history of depression, osteoporosis, American Society of Anesthesiologists (ASA) physical status classification, as well as all intraoperative and perioperative complications. Surgical data collected included primary versus revision surgery, single versus staged procedures, posterior fusion rod diameter and material, the uppermost instrumented vertebra, the lowermost instrumented vertebra, use of direct spinal decompression, number of decompression levels, number of Smith-Petersen osteotomies, presence of a 3-column osteotomy (pedicle subtraction osteotomy or vertebral column resection), number of interbody fusions, use of an iliac graft, use of recombinant human bone morphogenetic protein (BMP), and number of posterior vertebral levels fused. The surgical variables were included under the assumption that the same factors could be derived from a preoperative surgical plan and that their inclusion would yield a more complete preoperative predictive model.

Full-length (36-inch cassette) posteroanterior and lateral spine radiographs at baseline and 6 weeks follow-up were analyzed using validated software (SpineView, ENSAM, Laboratory of Biomechanics). Only baseline radiographic measures were included in the model and the 6-week radiographs were used for complication determination. All radiographic measures were performed at a central location based on standard techniques and included: maximum coronal Cobb angles of thoracic and lumbar curves (grouped by < 30°, 30°–60°, and > 60°); coronal C-7 plumb line; TK (T4–12; Cobb angle between superior endplate of T-4 and inferior endplate of T-12); lumbar lordosis (Cobb angle between superior endplate of L-1 and superior endplate of S-1); SVA (offset of C-7 plumb line relative to S-1); PT; the mismatch between pelvic incidence and lumbar lordosis (PI-LP); and the T-1 pelvic angle (TIPA); the angle between the line from the
femoral head axis to the centroid of T-1 and the line from the femoral head axis to the middle of the S-1 endplate). Based on the above radiographic parameters, patients were additionally stratified by the SRS-Schwab ASD classification.35

Standardized HRQOL measures were recorded at baseline and included the Oswestry Disability Index (ODI), SF-36, and Scoliosis Research Society-22r questionnaire (SRS-22r). Two standard summary scores were calculated based on the SF-36, the Physical Component Summary (PCS) and the Mental Component Summary (MCS). The SRS-22r provides a total score and multiple subdomains, including 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.

Patients were grouped as either having at least 1 major intraoperative or perioperative complication (COMP group) or not (NOCOMP group). Perioperative complications were defined as those occurring within 6 weeks of surgery. Major and minor complications were classified according to the study of Glassman et al.18 All of the patients included had a minimum of 6 weeks of follow-up to capture any perioperative complications. The complication categories included cardiopulmonary, electrolyte, gastrointestinal, implant, infection, musculoskeletal, neurological, operative, radiographic, renal, vascular, wound, and other.

Statistical Analysis and Predictive Model Construction
Continuous variables were described with means and standard deviations. Baseline variables were compared between the groups. Normality of data was determined using the Shapiro-Wilk test. Comparison of baseline means between the groups included the Student t-test or Wilcoxon rank-sum tests where appropriate. Frequency analyses for categorical variables were conducted via Pearson’s χ² analysis. All statistical analyses were conducted using commercially available software (SPSS version 22, IBM Inc.) 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 such as mean and median imputation.1 Once a complete data set was constructed, an ensemble of decision trees was constructed with a binary target variable that included patients who sustained at least 1 major intraoperative or perioperative complication, as defined above (code = 1), or not having any major intraoperative or perioperative complications (code = 0). The decision-tree algorithm was C5.0 and 5 different bootstrapped models were built.1 Internal validation was accomplished via a 70/30 data split for training and testing the model, respectively.1 Final overall predictions from the models were combined and chosen by voting with random selection for tied votes. Overall accuracy and the area under the receiver operating characteristic (AUROC) curve were calculated as well as predictor importance as determined by the model. The model was built using commercially available software (SPSS Modeler version 16, IBM Inc.).

Results
Patient Population
A total of 557 operative patients were available and included in the study. Of those, 409 did not sustain a major complication (NOCOMP group, 73.4%), and 148 had at least 1 intraoperative or perioperative major complication (COMP group, 26.6%). From the total, 390 patients (70%) were used for model training and 167 (30%) for testing the model. The percentage split was determined randomly but was within the acceptable splitting percentage options for predictive modeling.1 There were 439 women (78.8%) and 181 men (21.2%) and the mean age was 57.5 ± 15.3 years with a mean BMI of 27.6 ± 8.6 kg/m² (Table 1). The COMP group was significantly older, had a greater mean BMI, higher mean CCI, higher mean ASA score, and a higher proportion of patients with osteoporosis (p < 0.05 for all, Table 1). The COMP group had a significantly higher proportion of patients with baseline SRS-Schwab coronal curve type of L (47.6% vs 30.8%, p < 0.05, Table 1).

Surgical Data
The COMP group had a larger proportion of patients who underwent a revision (41.9%) versus primary surgery (31.5%; p = 0.023) and who had a direct decompression (25.7% vs 20.3%, p = 0.0034; Table 2). In addition, the COMP group had a significantly higher proportion of patients who underwent an interbody fusion, iliac crest graft, and BMP use (p < 0.05 for all, Table 2). The mean number of posterior levels fused was statistically similar between groups (p > 0.05, Table 2). Both groups also had

| TABLE 1. Demographic data of all the patients as well as each group |
|-------------------------|-----------------|-----------------|-----------------|-----------------|
| Variable                | All Patients    | COMP            | NOCOMP          | p Value         |
| No. of patients         | 557             | 148             | 409             |                 |
| Mean age ± SD (yrs)     | 57.5 ± 15.3     | 61.5 ± 12.3     | 56 ± 16.1       | 0.0004          |
| Females/males           | 439/118         | 119/29          | 320/89          | 0.5806          |
| Mean BMI ± SD           | 27.6 ± 8.6      | 28.6 ± 6.6      | 27.3 ± 9.3      | 0.0013          |
| Mean CCI ± SD           | 1.5 ± 1.7       | 1.9 ± 1.7       | 1.4 ± 1.6       | 0.0009          |
| Mean ASA score ± SD     | 2.3 ± 0.7       | 2.4 ± 0.6       | 2.3 ± 0.7       | 0.0061          |
| Precip anemia (%)       | 52 (9.3)        | 12 (8.1)        | 40 (9.8)        | 0.5491          |
| Osteoporosis (%)        | 70 (12.6)       | 28 (18.9)       | 42 (10.3)       | 0.0065          |
| Depression (%)          | 131 (23.5)      | 35 (23.6)       | 96 (23.5)       | 0.9653          |
| SRS-Schwab coronal curve type (%) |                  |                 |                 | 0.0047          |
    | N                      | 33.6            | 30.3            | 34.8            |
    | T                      | 5.9             | 4.8             | 6.3             |
    | L                      | 35.3            | 47.6            | 30.8            |
    | D                      | 25.1            | 17.2            | 28.0            |

SRS-Schwab coronal curve types: N = patients with no coronal curve > 30° (i.e., no major coronal deformity); T = patients with a thoracic major curve > 30° (apical level of T-9 or higher); L = patients with a lumbar or thoracolumbar major curve > 30° (apical level of T-10 or lower); and D = patients with a double major curve, with each curve > 30°. Boldface type indicates statistically significant differences between the COMP and NOCOMP groups.
statistically similar proportions of patients with different rod diameters and materials (p > 0.05 for both, Table 2). The proportion of patients who suffered major complications and specific complication subtypes are presented in Table 3.

**Radiographic Data**

Baseline radiographic parameters of the COMP and NOCOMP groups are shown in Table 4. The COMP group had significantly greater mean baseline SVA (77.7 mm vs 61.2 mm, p = 0.015) and T1PA (25.0° vs 21.4°, p = 0.014) than the NOCOMP group (p < 0.05 for both). Both groups had similar baseline mean PT, PI-LL, TK, C-7 coronal plumb line, and proportion of patients in each of the maximum coronal Cobb angle groups (p > 0.05 for all).

**HRQOL Data**

The COMP group had significantly worse baseline mean scores for ODI, PCS, SRS activity, SRS pain, SRS total, and leg pain NRS (p < 0.05 for all, Table 5). All other HRQOL variables were statistically similar between the groups (p > 0.05 for all).

**Model Results**

The overall model accuracy was 87.6% correct with an AUROC of 0.89 indicating a very good model fit. In total, 20 variables were determined to be the top predictors (importance ≥ 0.90 out of 0–1 as determined by the model) and included (in decreasing importance): age, leg pain NRS, ODI, number of levels decompressed, number of levels with an interbody fusion, PCS, SRS-Schwab coronal curve type, CCI, SRS activity, T1PA, ASA grade, presence of osteoporosis, PT, SVA, primary version revi-
This allows for increased accuracy at the cost of decrease in interpretability (transparency). The exact rules governing how the predictions are made are unavailable because the computer calculates all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtains odds/hazard ratios with the computer calculating all of the predictions, in contrast to logistic regression, in which one obtai
selected for a given clinical question/hypothesis and type of data. Using the wrong statistical test can result in erroneous results. Predictive modeling is more flexible due to the fact that it relies on the available data. Decision trees were used in this setting because of the advantages stated above, not because of the type of data and not because of a certain hypothesis. Additionally, multiple types of models can be constructed to assess which one has the greatest utility for the aims of the model, hence the ensemble of decision trees that was used. One can combine models that are similar (or different) to further enhance the goals of the model. When creating a model one must balance accuracy, generalizability, and transparency. We have created a model that is accurate and theoretically generalizable, yet not transparent. The type of model being built, and the balance of the 3 primary characteristics, depends on how it will be used when deployed in clinical practice.

There are few predictive models reported in the spine literature that Osorio et al. has identified, and for the ones that do exist, logistic regression is a popular technique that produces sets of odds ratios for developing the outcome of interest. Logistic regression is commonly used in prediction analysis because it is simple, easy to interpret and apply, and transparent. Chapman and colleagues consolidated a number of logistic regression models for complications following “spine surgery” from 6 different peer-reviewed papers into a web-based predictive model. This was an impressive feat, but in contrast to the decision trees used in the present study, there are specific limitations to logistic regression. There are a number of assumptions that must be satisfied to apply logistic regression as mentioned above and they generally identify variables that are “predictors” without a patient-specific interpretation. Our predictive model was constructed in 1 setting as opposed to 6 different studies with all the variables in 1 database. It is more efficient, updated with current predictive algorithms, is patient specific, and applies directly to patients with ASD.

Similar to our methodology, Daubs and colleagues performed a decision-tree analysis and used an ensemble of 50 decision trees to predict psychological distress in spine patients. Their model was very successful using 6 variables and 188 patients, as it was 92% accurate, 92% sensitive, and 95% specific. And lastly, one of the more advanced predictive modeling techniques was deployed by Azimi and colleagues in which they created an artificial neural network (ANN) to predict 2-year surgical satisfaction in surgical patients for lumbar spinal canal stenosis undergoing surgery. They compared the ANN to logistic regression and found that the use of ANN was more accurate than the logistic regression model. These types of studies, in addition to ours, represent the beginning of the use of predictive analytics in spine surgery outcomes. As data sets get larger with time and the quality of the data increases, advanced predictive analytics will likely play a larger role in clinical decision making.

The strengths of the current study include the multicenter design and a large number of patients with ASD (n = 557). The multicenter design (11 different sites across the US) allows for better generalizability of the results. Another strength of this study is the complete preoperative and 6-week follow-up of the patients as well as the use of 45 variables. And lastly, modern predictive analytics algorithms were used to create the model, providing a patient-specific decision-tree ensemble.

However, there are a few limitations to this study, one of which includes the retrospective design that may have introduced selection or information biases. Another limitation includes combining both intra- and perioperative complications as the target variable of interest. Ideally, with greater numbers of patients, these intraoperative and perioperative complications would be separated out. This model is one of the first of its kind and sets the groundwork for advanced predictive analytics in spinal outcomes research.

Conclusions
A successful model (87% accuracy, 0.89 AUROC) was built predicting major intra- or perioperative complications. This model can provide the foundation toward improved education and point-of-care decision making for patients undergoing ASD surgery.

References


Disclosures
The International Spine Study Group Foundation (ISSGF), through which this study was conducted, is funded through research grants from DePuy Spine, Innovasis, Biomet, K2, NuVasive, and individual donations. Dr. Smith has served as a consultant to Zimmer Biomet, NuVasive, and Cerapedics; has received royalties from Zimmer Biomet; has received support of non–study-related clinical or research effort from DePuy Synthes; has received clinical or research support for the study described from DePuy Synthes; and has received fellowship support from AO Spine and the NREF. Dr. Schwab has direct stock ownership in Nemaris Inc.; has served as a consultant to Zimmer Biomet, K2M, MSD, Medicrea, and NuVasive; has received support of non–study-related clinical or research effort from SRS, AO, and DePuy Synthes (paid through ISSGF); is a patent holder for MSD and K2M; and has speaking/teaching arrangements with Zimmer Biomet, K2M, MSD, Medicrea, and NuVasive. Dr. Lafage has direct stock ownership in NuVasive; has speaking/teaching arrangements with NuVasive, Medicrea, DePuy Spine, and Nemaris Inc.; and has received support of non–study-related clinical or research effort from SRS, NIH, and DePuy Synthes (paid through ISSGF). Dr. Shaffrey has served as a consultant to Medtronic, NuVasive, Zimmer Biomet, K2M, and Stryker; has direct stock ownership in NuVasive; is a patent holder for Medtronic, NuVasive, and Zimmer Biomet; and has received support of non–study-related clinical or research effort from DePuy Synthes. Dr. Bess has served as a consultant to NuVasive, K2 Medical, and Allosource; is a patent holder for Innovasis and K2 Medical; has received clinical or research support for this study from DePuy Spine; and has received support of non–study-related clinical or research effort from Medtronic, K2 Medical, Innovasis, and Stryker. Dr. Daniels has served as a consultant to Stryker, Globus, DePuy, and Orthofix; and has received support of non–study-related clinical or research effort from Orthofix. Dr. Hart has served as a consultant to DePuy Synthes and Globus; is a patent holder for Oregon Health & Science University; and has received royalties and/or paid honoraria from SeaSpine, DePuy Synthes, and Globus. Dr. Protopsaltis has served as a consultant to Medtronic International, Globus, and Innovasis; and has received clinical or research support for this study from Zimmer Spine. Dr. Mundis Jr. has served as a consultant to NuVasive, K2M, and DePuy Synthes. Dr. Scibba has served as a consultant to Medtronic, DePuy Synthes, Orthofix, Globus, and Stryker. Dr. Burton has served as a consultant to DePuy Spine, and has received clinical or research support for this study from DePuy Spine. Dr. Klineberg has served as a consultant to DePuy and Stryker, and has received a fellowship grant and honoraria from AOSpine, and honoraria from K2M. Dr. Ames has served as a consultant to DePuy, Stryker, and Medtronic; is a patent holder for Fish & Richardson, PC; and has received royalties from Stryker and Biomet Spine.

Author Contributions
Conception and design: Scheer, Ames. Acquisition of data: Lafage. Analysis and interpretation of data: Scheer. Drafting the article: Scheer. Critically revising the article: Scheer, Smith, Schwab, Daniels, Protopsaltis, Aillon, Burton, 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.

Correspondence
Justin K. Scheer, School of Medicine, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093. email: jscheer@uic.edu.