ORIGINAL ARTICLE

Development of a model to predict the probability of incurring a complication during spine surgery

Pascal Zehnder[1] · Ulrike Held[1] · Tim Pigott[2] · Andrea Luca[3] · Markus Loibl[4] · Raluca Reitmeir[4] · Tamás Fekete[4] · Daniel Haschtmann[4] · Anne F. Mannion[4]

Abstract

Purpose Predictive models in spine surgery are of use in shared decision-making. This study sought to develop multivariable models to predict the probability of general and surgical perioperative complications of spinal surgery for lumbar degenerative diseases.

Methods Data came from EUROSPINE's Spine Tango Registry (1.2012–12.2017). Separate prediction models were built for surgical and general complications. Potential predictors included age, gender, previous spine surgery, additional pathology, BMI, smoking status, morbidity, prophylaxis, technology used, and the modified Mirza invasiveness index score. Complete case multiple logistic regression was used. Discrimination was assessed using area under the receiver operating characteristic curve (AUC) with 95% confidence intervals (CI). Plots were used to assess the calibration of the models.

Results Overall, 23′714/68′111 patients (54.6%) were available for complete case analysis: 763 (3.2%) had a general complication, with ASA score being strongly predictive (ASA-2 OR 1.6, 95% CI 1.20–2.12; ASA-3 OR 2.98, 95% CI 2.19–4.07; ASA-4 OR 5.62, 95% CI 3.04–10.41), while 2534 (10.7%) had a surgical complication, with previous surgery at the same level being an important predictor (OR 1.9, 95%CI 1.71–2.12). Respectively, model AUCs were 0.74 (95% CI, 0.72–0.76) and 0.64 (95% CI, 0.62–0.65), and calibration was good up to predicted probabilities of 0.30 and 0.25, respectively.

Conclusion We developed two models to predict complications associated with spinal surgery. Surgical complications were predicted with less discriminative ability than general complications. Reoperation at the same level was strongly predictive of surgical complications and a higher ASA score, of general complications. A web-based prediction tool was developed at https://sst.webauthor.com/go/fx/run.cfm?fx=SSTCalculator.

Keywords Spine surgery · Patient outcome · Complications · Prediction model · Degenerative spine

Introduction

Patients today are, correctly, much more involved in decision-making regarding their treatment than they used to be [1]. Physicians and surgeons treating spinal disorders should be able to make evidence-based predictions regarding the outcome of their treatments, based on reliable prognostic information. In the last few years, a number of statistical prediction models have been developed to predict the outcome of spine surgery [2–8], mainly focusing on the benefits of the intervention regarding pain relief, quality of life improvement and/or return to work.

In view of recent developments in shared decision-making, not only the benefits but also the risks associated with different treatment modalities must be clearly communicated to the patient. For this reason, risk calculators have also been developed to predict complication rates [4, 9–13].

Lee et al. were the first to create a predictive model assessing the risk of medical complications following spine surgery and to develop an online tool for its clinical use [10]. Their study was based on a population of 1476 patients, split into two subsets for internal and cross-validation. Although successful, they acknowledged that the accuracy of such predictive models would be improved with greater power.

✉ Pascal Zehnder
pascal.zehnder@uzh.ch

1 University of Zürich, Zürich, Switzerland

2 The Walton Centre NHS Foundation Trust, Liverpool, UK

3 IRCCS Galeazzi, Milan, Italy

4 Schulthess Klinik, Zürich, Switzerland

Also, their model only evaluated the risk for general medical complications and lacked a surgical complication counterpart. Later on, they developed a model for surgical site infection—one of many possible surgical complications—on the same patient population [9]. The model showed reasonable success as far as its discriminative capacity was concerned, but no information was provided regarding how well the predictions were aligned with the observed outcomes (i.e. "calibration").

Kasparek et al. [14] sought to validate the general-medical complication model of Lee et al. [10]. In total, 44 patients developed a medical complication in a population of 273 patients undergoing spinal surgery. The model demonstrated adequate prediction of the medical complication risk group (low, medium, high), but the authors conceded that, with only 273 patients, the analysis may have been underpowered. Janssen et al. [15] investigated the validity of the surgical site infection model of Lee et al. [9] in a population of 898 patients undergoing thoracolumbar spine surgery. They demonstrated a Nagelkerke's R2 of 0.01, indicating poor external predictive strength.

Kim et al. [16, 17] used artificial neural networks in addition to classic logistic regression methods to identify risk factors for various types of complication in two subsets of spine patients: those undergoing elective adult spinal deformity surgery [16] and those undergoing posterior lumbar spine fusion [17]. They trained their multivariable models using data from the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) database, and model performance was compared with that of a model containing just the American Society for Anesthesiology (ASA) score as predictor. The areas under the receiver operating characteristic curves (AUC) were better for the multivariable models (0.54–0.84, depending on the complication in question) than for ASA alone (0.37–0.52). This hence showed promising results in the advancing field of artificial neural network models. However, no online tool was developed for its further use or the assessment of its validity in the clinical setting. Indeed, neural networks are notoriously more challenging to use for the development of decision-support systems, since the most important input variables are more difficult to identify than they are in regression models.

The aforementioned studies focused on medical complications based on small databases, were conducted as single-centre studies, had poor external validity, or used novel statistical/machine-learning approaches to produce models that do not easily lend themselves to further validation by others.

In the present study our aim was to combine previously identified individual predictors of outcome, e.g. number of previous spine surgeries [18, 19], age [20], ASA score [21], complexity of the surgery [22], BMI [23], and smoking status [24, 25] in a multivariable model to predict complications, using a large multicentre dataset from the EUROSPINE "Spine Tango" registry [26]. We sought to develop two models and a web-based tool for predicting the likelihood of incurring a perioperative complication in connection with spine surgery: one focusing on general medical complications and the other, on surgical complications.

Methods

Source of data

This was a retrospective multicentre registry data-based study of prospectively collected data within the EURO-SPINE "Spine Tango" Registry [26]. The analysis of these routinely collected, anonymised data was approved by the Swiss Ethics Committees for research involving humans [27].

The registry currently includes over 120 '000 surgical patient cases from several spine centres from more than 20 countries. Medical history and surgical details are documented by the surgeon using the Spine Tango surgery form, as are surgical and general medical complications arising between admission and discharge.

Participants

Inclusion and exclusion criteria

Patients included in the present study were operated on between January 2012 and December 2017. All included patients had spine surgery for degenerative disorders of the lumbar spine in one of the participating EUROSPINE Spine Tango centres. Patients aged between 18 and 95 were identified on the basis of a "main pathology" documented as "degenerative disease", and "level of intervention" as "thoracolumbar", "thoraco-lumbo-sacral", ''lumbar'', "lumbosacral" or ''sacral'' on the Spine Tango 2011 Surgery form. All patient cases with missing values for the variables indicated below were excluded from the analysis (see later).

Outcome

We defined two independent binary (presence/absence) outcome variables describing perioperative complications. First, the occurrence of a surgical complication arising either intraoperatively (nerve root damage, spinal cord damage, dura lesion, vascular injury, fractures of vertebral structures) and/or postoperatively before discharge (epidural or other hematoma, radiculopathy, CSF leak/pseudomeningocele, motor/sensory dysfunction, bowel/bladder dysfunction, wound infection, implant malposition, implant failure, wrong level). Second, the occurrence of a general complication arising either intraoperatively (anaesthesiological,

cardiovascular, pulmonary, thromboembolism, death) and/or postoperatively before discharge (cardiovascular, pulmonary, cerebral, kidney/urinary, liver/gastrointestinal, thromboembolism, death).

Predictors

The selection of predictors was based on evidence of their predictive capacity in the current literature [18–25, 28]. All the predictor variables are collected systematically in the registry and are typically known or decided upon preoperatively, on the "admission" section of the Spine Tango Form. They included: age; sex (male, female); whether patients had had previous spine surgery (if so, whether at the same or on a different spinal level); body mass index (BMI, < 20, 20–25, 26–30, 31–35, > 35 kg/m^2); smoking status (no, yes); morbidity state [using the American Society of Anaesthesiologists Physical Status Score (ASA; scored 1–5)]; the modified Mirza Invasiveness Index (representing the invasiveness of the planned surgical procedure) [29, 30]; the presence of an additional pathology (other than degenerative disease, such as fracture/trauma, non-degenerative deformity, infection, tumour, etc.); planned preoperative prophylaxis (infection, thromboembolism, ossification); and planned intraoperative technology (conventional or minimally invasive spine surgery (MISS)/less invasive spine surgery(LISS)).

Sample size

Sample size considerations for prediction models need to take into account whether the outcome number of events is large enough for fitting multiple predictor models, taking all relevant predictors into account simultaneously. The generally accepted rule is to have at least ten events per variable [30–32]. Given that, in the planned study, 10 independent predictor variables were to be evaluated simultaneously, this implied at least 100 events (occurrence of a complication) observed for each of the outcomes, a condition that was met for the generation of separate models to predict general medical complications and surgical complications (see later).

Missing data

A complete case analysis was used for the development of the prediction models. As such, cases with any missing data regarding the previously noted predictors or outcomes were excluded from the analysis. The results were reanalysed at a later stage in a sensitivity analysis based on imputed data.

Statistical analysis methods

Descriptive statistics included median and interquartile ranges or mean and standard deviations (SD) for continuous variables and counts and percentages of total for categorical variables. The two binary outcome variables, general medical perioperative complications and surgical perioperative complications, were addressed with multiple logistic regression models fitted to each outcome.

Predictor models with many prognostic indicators tend to fit the data used in the study optimally, but predictions for new subjects perform less well. This problem is known as overfitting. To address overfitting in our models, we used shrinkage, a technique in which the regression coefficients of the prediction models are multiplied by a global shrinkage factor (a real number < 1), leading to a reduction in their values ("shrinkage") towards zero. We used the *dfbeta* method [33] to derive a separate global shrinkage factor for each of the two models. This method is equivalent to leave-one-out cross-validation (LOOCV), which is a common but less efficient approach that is difficult to apply to large datasets [33] such as the present one. Original and shrunken regression coefficients were presented as odds ratios (OR) with 95% confidence intervals (CI).

Model performance

To assess model performance, we evaluated how well the predicted probability of a complication corresponded to the actual observed complication rate, by assessing the model's discriminative ability and calibration. Discrimination was examined using the area under the receiver operating characteristic curve (AUROC, or c-statistic) with 95% CI. In the ROC curve, sensitivity is plotted against 1-specificity. In general, an AUC of 0.5 suggests no better discrimination than tossing a coin, 0.6–0.7 is considered possibly helpful, 0.7–0.8 is considered acceptable, 0.8–0.9 is considered excellent, and more than 0.9 is considered outstanding [34].

Calibration of a prediction model measures the agreement between observed outcomes and predictions. In the present study, internal calibration of the two models was assessed using calibration plots. Internal calibration refers to agreement between observed and predicted probabilities in the sample in which the model was developed, showing how well the model represents the observed reality and whether it tends to over- or underestimate the probability of an event [35–37].

Sensitivity analysis

A sensitivity analysis was performed to assess the potential bias in the results as a consequence of using complete case analysis. For this reason, a single dataset was imputed based on the assumption that data were missing at random (MAR). The imputed dataset was obtained from doing a single imputation, in which missing values for our predictors (e.g. missing smoking status, BMI or ASA score) were

filled using multivariate imputation based on chained equations [38, 39]. After single imputation the coefficients of the two models, corresponding AUC's and their calibration plots were re-estimated, using the aforementioned methods.

All analyses were conducted using *R for Windows* [40], using the packages *openxlsx, tableone, pROC, tidyverse, shrink* [33]*, boot, biostatUZH, mice* [39] and *gbm*. The work was carried out following the concept of reproducible research, and the R-code is available upon request [41]. The results of the study were reported according to the TRIPOD guidelines [42].

Results

Participants

Figure 1 shows the flowchart for patient inclusion in the study. In total, 68′111 cases were registered in the database at the time of data export. Of these, 54′452 were degenerative cases, of which 43′557 included the lumbar spine (according to the definition above). Selecting patients between age 18 and 95 years resulted in 43′461 cases. Using a complete case approach to the analysis, the final sample size was $N = 23′714$, with the total number of cases excluded being 19′747 (45.4%). The variables with the most missing data were smoking status (39.4%), morbidity status (18.2%) and BMI (15.4%). The baseline characteristics of the final study group are shown in Table 1. The mean age was 58.9 (15.7 SD) years, and 11,450 (48.3%) were males. In total, 16,921 (71.4%) patients had had no previous surgery. Patients were most frequently (9176; 38.7%) in the BMI category 26–30 kg/m². Most participants, 18,799 (79.3%), did not smoke. The most common morbidity state was ASA-2 (12,941 (54.6%) patients). The median Mirza score was 2 (interquartile range 1 to 7). Most participants had infection prophylaxis (22,832; 96.3%), and the majority had thromboembolism prophylaxis (17,754; 74.9%). There were more conventional technologies used (9437; 39.8%) than MISS/LISS (2502; 13.7%); the most commonly used technology was microscope (14,268; 60.2%).

Details regarding the incidence of the different types of complication are shown in Table 2 for general medical complications and Table 3 for surgical complications. Overall, 763/23,714 (3.2%) patients had a general medical complication and 2534/23,714 (10.7%), a surgical complication, indicating that sufficient events were observed to be able to fit the multiple prediction models. The most common intraoperative general complication was of a cardiovascular nature (25; 0.11%) and the most common postoperative general complication, postoperative kidney/urinary problems, being reported in 200 (0.84%) cases. The most common intraoperative surgical complication was dural tear, being reported in
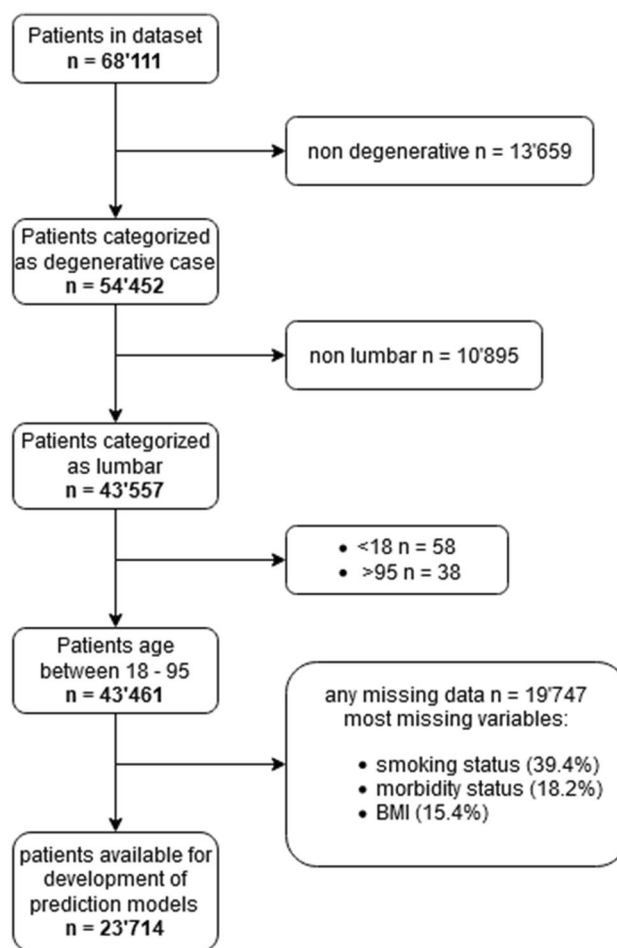

Fig. 1 Patient selection flowchart

1638 (6.91%) cases, with motor dysfunction being the most common postoperative surgical complication 168 (0.71%).

Model development

Model predicting general medical complications

The calculated shrinkage factor for the general medical complication model was 0.98, indicating that not much overfitting was present. The odds ratios, their 95% confidence intervals (CI) and the p values for the shrunken prediction model for general medical complications are shown in Table 4. Higher age (OR 1.03, 95% CI 1.03–1.04 per year) was associated with greater odds of having a complication. An ASA score of 2 or more was also associated with greater odds of having a complication, with the effect being more marked the higher the ASA score (ASA 2, OR 1.6, 95% CI 1.2–2.12; ASA 3, OR 2.98, 95% CI 2.19–4.07; ASA 4, OR 5.62, 95% CI 3.04–10.41), as were more complex procedures according to the modified Mirza score (OR 1.03, 95% CI 1.02–1.04 per point increase) and conventional

Table 1 Baseline characteristics of the study group

Variable	Number
N	23,714
Age (mean (SD)) years	58.9 (15.7)
Gender = male (%)	11,450 (48.3)
Any additional pathology (%)	1949 (8.2)
Previous surgery (%)	
No previous surgery	16,921 (71.4)
Surgery at the same level	3567 (15.0)
Surgery at a diff. level	3226 (13.6)
BMI kg/m^2 (%)	
< 20	906 (3.8)
20 – 25	7677 (32.4)
26 – 30	9176 (38.7)
31 – 35	4243 (17.9)
> 35	1712 (7.2)
Smoker = no (%)	18,799 (79.3)
Morbidity (%)	
ASA1	6369 (26.9)
ASA2	12,941 (54.6)
ASA3	4296 (18.0)
ASA4	108 (0.5)
Modified Mirza index (median [IQR])	2.0 [1.0, 7.0]
*Prophylaxis**	
No prophylaxis (%)	380 (1.6)
Infection prophylaxis (%)	22,832 (96.3)
Thromboembolism prophylaxis (%)	17,754 (74.9)
Ossification prophylaxis (%)	116 (0.5)
*Technology**	
Conventional (%)	9437 (39.8)
MISS/LISS (%)	3237 (13.7)
Loops (%)	2502 (10.6)
Endoscope (%)	39 (0.2)
CASS (%)	126 (0.5)
Microscope (%)	14,268 (60.2)
Neuromonitoring (%)	1115 (4.7)
Other (%)	666 (2.8)

*Multiple responses possible

Table 2 General complication counts and percentages in the group of 23,714 patients

Variable	Number
N	23,714
Total number patients with any general complication (%)	763 (3.2)
*Intraoperative general complications**	
None (%)	23,648 (99.72)
Anaesthesiological (%)	21 (0.09)
Cardiovascular (%)	25 (0.11)
Pulmonary (%)	5 (0.02)
Thromboembolism (%)	3 (0.01)
Death (%)	0 (0)
Other (%)	16 (0.07)
*Postoperative general medical complications**	
None (%)	23,006 (97.01)
Cardiovascular (%)	146 (0.62)
Pulmonary (%)	82 (0.35)
Cerebral (%)	42 (0.18)
Kidney/urinary (%)	200 (0.84)
Liver/GI (%)	88 (0.37)
Thromboembolism (%)	22 (0.09)
Death (%)	7 (0.03)
Other (%)	201 (0.85)

*Multiple responses possible

at the same level (OR 1.9, 95% CI 1.71–2.12), BMI over 35 (OR 1.29, 95% CI 1.00–1.67), an ASA score of 3 (OR 1.23, 95% CI 1.06–1.43), a higher modified Mirza score (OR 1.01, 95% CI 1.00–1.01), any additional spine pathology (OR 1.3, 95% CI 1.14–1.49), ossification prophylaxis (OR 2.18, 95% CI 1.38–3.46) and conventional surgical technology having been used (OR 1.12, 95% CI 1.02–1.22) were each associated with an increased odds of having a surgical complication. Male gender (OR 0.81, 95% CI 0.75–0.89) was associated with decreased odds of incurring a surgical complication, as was using thromboembolism prophylaxis (OR 0.85, 95% CI 0.77–0.94).

Model performance

The ROC's for the models can be seen in Fig. 2 (general medical complications) and Fig. 3 (surgical complications). The AUC for the model for general complications was 0.74 (95% CI: 0.72–0.76), while that for surgical complications was 0.64 (95% CI: 0.62–0.65) after shrinkage. The calibration plots for the two models are shown in Fig. 4 (general complications) and Fig. 5 (surgical complications). In the calibration curve for the general complications model, the observed values agreed well with the predicted values up to a predicted probability of 0.3. However, beyond this

surgical technology having been used (OR 1.32, 95% CI 1.12–1.54). Using infection prophylaxis (OR 0.59, 95% CI 0.37–0.92) was associated with reduced odds of a general medical complication.

Model predicting surgical complications

The calculated shrinkage factor for the surgical complication model was 0.97. The odds ratios, their confidence intervals (95% CI) and the p values for the shrunken prediction model for surgical complications are shown in Table 5. Higher age (OR 1.02, 95% CI 1.01–1.02), previous spine surgery

Table 3 Surgical complication counts and percentages in the group of 23,714 patients

Variable	Number
N	23,714
Total number of patients with any surgical complication (%)	2534 (10.7)
*Intraoperative surgical complications**	
None (%)	21,891 (92.31)
Nerve root damage (%)	89 (0.38)
Spinal cord damage (%)	9 (0.04)
Dural tear (%)	1638 (6.91)
Vascular injury (%)	14 (0.06)
Fracture of vertebral structures (%)	25 (0.11)
Other (%)	95 (0.40)
*Postoperative surgical complications**	
None (%)	22,748 (95.93)
Epidural hematoma (%)	152 (0.64)
Other hematoma (%)	55 (0.23)
Radiculopathy (%)	157 (0.66)
CSF leak/pseudomeningocele (%)	87 (0.37)
Motor dysfunction (%)	168 (0.71)
Sensory dysfunction (%)	141 (0.59)
Bowel/bladder dysfunction (%)	90 (0.38)
Wound infection superficial (%)	87 (0.37)
Wound infection deep (%)	60 (0.25)
Implant malposition (%)	44 (0.19)
Implant failure (%)	17 (0.07)
Wrong level (%)	5 (0.02)
Other (%)	167 (0.70)

*Multiple responses possible

Table 4 Shrunken regression coefficients for the general complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	*p* value
(Intercept)	0	0–0.01	< 0.0001
Age years	1.03	1.03–1.04	< 0.0001
Gender = male	0.92	0.79—1.07	0.26
Previous surgeries			
At the same level	0.99	0.79–1.23	0.91
At a different level	1.12	0.93–1.36	0.24
BMI kg/m²			
20–25	1.16	0.77–1.73	0.47
26–30	0.92	0.62–1.38	0.70
31–35	0.97	0.64–1.48	0.89
> 35	1.11	0.7–1.76	0.65
Smoker = no	0.91	0.75–1.11	0.37
Morbidity			
ASA2	1.6	1.2–2.12	0.001
ASA3	2.98	2.19–.07	< 0.0001
ASA4	5.62	3.04–10.41	< 0.0001
Mirza score	1.03	1.02–1.04	< 0.0001
Any additional pathology	1.24	1–1.55	0.051
Prophylaxis			
No prophylaxis	0.5	0.2–1.27	0.15
Infection prophylaxis	0.59	0.37–0.92	0.02
Thromboembolism prophylaxis	1.17	0.97–1.41	0.11
Ossification prophylaxis	1.89	0.76–4.68	0.17
Technology used			
Conventional technology	1.32	1.12–1.54	0.0007
MISS/LISS	0.91	0.71–1.17	0.46

point, higher predicted probability values corresponded to much lower observed values, and the confidence intervals increased markedly. The same pattern was seen for the surgical complications model, but there the inflection point was reached at a predicted probability of about 0.25.

Sensitivity analysis

The estimated coefficients for the predictors in the dataset with single imputation are shown in "Appendix" (Tables 8 and 9). Shrinkage factors calculated for the single imputation method were 0.985 for general medical complications and 0.980 for surgical complications. Recalculating the AUC for the single imputation dataset resulted in an AUC of 0.75 (95% CI: 0.74–0.76) for general medical complications and an AUC of 0.64 (95% CI: 0.63–0.65) for surgical complications. These results were then compared to the complete case dataset. The ROC curves and calibration plots for the sensitivity analysis can also be found in "Appendix" (Figs. 6, 7, 8, 9).

Discussion

Summary and Interpretation

We developed two models to predict general medical and surgical complications during spine surgery, based on the data collected within the EUROSPINE Spine Tango registry over a period of 6 years. The issue of overfitting was addressed by using shrinkage.

The ASA grade had the largest odds ratio for the probability of incurring a general complication, with a higher grade increasing the odds. The effect of the ASA grade on the incidence of general and surgical complications in spine surgery has been shown in numerous previous studies [21, 44–47].

Previous spine surgery at the same vertebral level had the largest per point odds ratio for an increased probability of a surgical complication. Nonetheless, very high values in continuous variables such as age or the modified Mirza score, which also increased the odds of a complication,

Table 5 Shrunken regression coefficients for the surgical complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	*p* value
(Intercept)	0.05	0.03–0.07	<0.0001
Age years	1.02	1.01–1.02	<0.0001
Gender = male	0.81	0.75–0.89	<0.0001
Previous surgeries			
At the same level	1.9	1.71–2.12	<0.0001
At a different level	1	0.88–1.13	0.97
BMI kg/m²			
20–25	0.96	0.77–1.21	0.74
26–30	0.99	0.79–1.23	0.91
31–35	1.11	0.88–1.4	0.39
>35	1.29	1.00–1.67	0.05
smoker = no	0.99	0.89–1.1	0.86
Morbidity			
ASA2	1.07	0.95–1.21	0.24
ASA3	1.23	1.06–1.43	0.007
ASA4	1.02	0.57–1.82	0.95
Mirza score	1.01	1–1.01	<0.0001
Any additional pathology	1.3	1.14–1.49	0.0001
Prophylaxis			
No prophylaxis	0.7	0.44–1.1	0.12
Infection prophylaxis	0.8	0.61–1.05	0.11
Thromboembolism prophylaxis	0.85	0.77–0.94	0.001
Ossification prophylaxis	2.18	1.38–3.46	0.0009
Technology used			
Conventional technology	1.12	1.02–1.22	0.018
MISS/LISS	1.11	0.97–1.25	0.12

could potentially exceed the per point odds ratio of previous spine surgery. The effect of previous spine surgery is already known and has been quantified before in data from the same registry [18, 19], and higher age and/or more invasive procedures have also been shown in previous studies to be associated with a greater likelihood of incurring a surgical complication [22, 43]. Our model also showed an increase in odds for surgical complications when additional non-degenerative spine pathologies are present. A possible explanation for this could be that the additional pathology makes the intervention more difficult and non-standard methods have to be used.

The models developed in the present study can be used when discussing a possible surgical intervention in a shared decision-making situation, along with models predicting the possible benefits or expected average outcomes, where models/baseline variables to predict individual outcome are not available. By using the regression coefficients from the model, one can estimate the risk of having either a surgical

or medical complication in an individual patient. For example, a male patient, age 40 y, body mass index between 20 and 25 kg.m⁻², non-smoker, without previous spinal surgery, ASA score of 1, undergoing L5–S1 posterior discectomy without posterior fusion (Mirza Score = 1), no additional pathologies, using prophylaxis for infection and thromboembolism have a calculated risk of 5.11% for having a surgical complication. The same patient has a calculated risk of 0.87% of suffering a medical complication.

In comparison, a female patient age 59 y, body mass index between 26 and 30 kg.m⁻², smoker, previous surgery at the same level, ASA score of 2, undergoing L4–S1 posterolateral fusion with pedicle screws and no decompression (Mirza score = 6), no additional pathologies, using infection and thromboembolism prophylaxis have a calculated risk of 16.8% of experiencing a surgical complication and a calculated risk of 2.76% of developing a medical complication. This preoperative knowledge might help patients and surgeons alike to decide upon their next treatment steps that could be taken pre- or perioperatively to minimise the risk of a complication, e.g. smoking cessation, weight loss, less invasive surgery. Based on the predictor models, a freely available, web-based prediction tool has been developed https://sst.webauthor.com/go/fx/run.cfm?fx=SSTCalculator.

To assess the performance of our models in terms of discrimination, we calculated ROC curves and their AUC's. The model predicting medical complications showed an AUC of 0.74, which can be considered as acceptable, and which was similar to the AUC of 0.76 reported for the model of Lee et al. [10]. The discriminative ability of the model predicting surgical complications was less good, with an AUC of just 0.64, but might still be considered possibly helpful. A possible explanation for the lower AUC for the surgical complications model might be that surgical complications in general could be less predictable and might be more dependent on the surgeon skill and experience rather than the patient's baseline characteristics that have been included in our model, although previous studies have shown that the chosen variables do have a predictive value for complications in degenerative spine surgery. In comparison, general complications could depend more on the physical status of the patient and hence be better able to be predicted.

Internal calibration was assessed using calibration plots. In the calibration plot for the general complications model we saw that the observed probabilities agreed well with the estimated probabilities for low predicted probabilities, and thus, the model shows a relatively satisfactory goodness-of-fit in that region. Beyond a predicted probability of approximately 0.3, however, the predicted values showed a clear overestimation of the reality and thus led to an overprediction of complication risks. The same applied to the model for surgical complications, which tended towards overestimation from a probability of about 0.25 onwards. One

Fig. 2 ROC curve general complication model. AUC 0.74 (95% CI: 0.72–0.76)
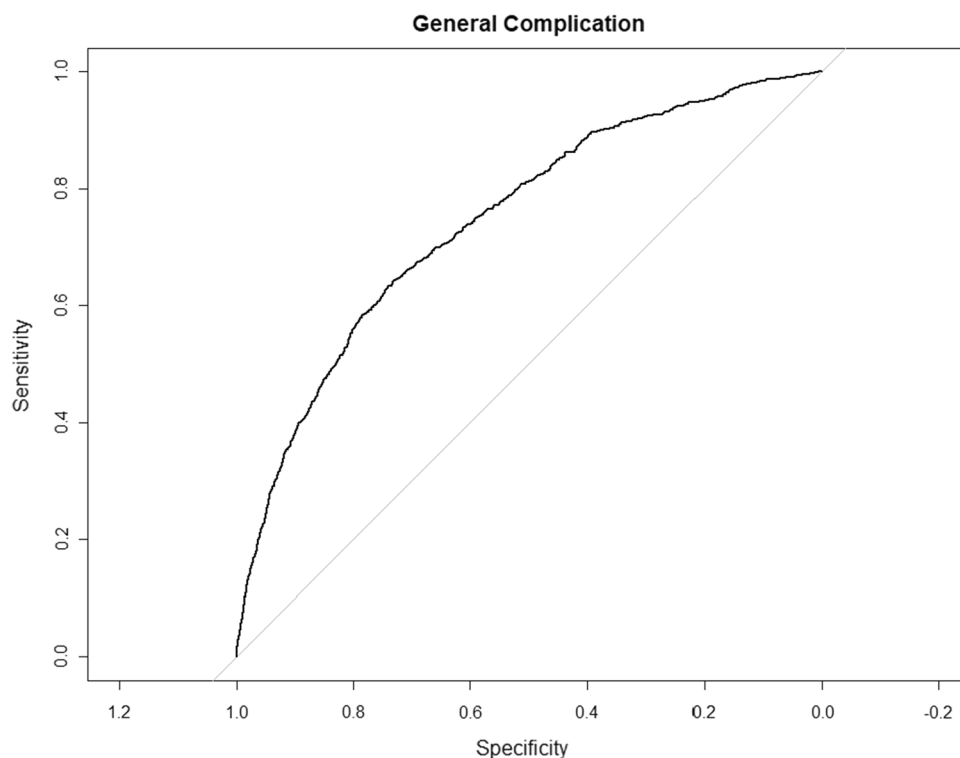
Fig. 3 ROC curve surgical complication model. AUC 0.64 (95% CI: 0.62–0.65)
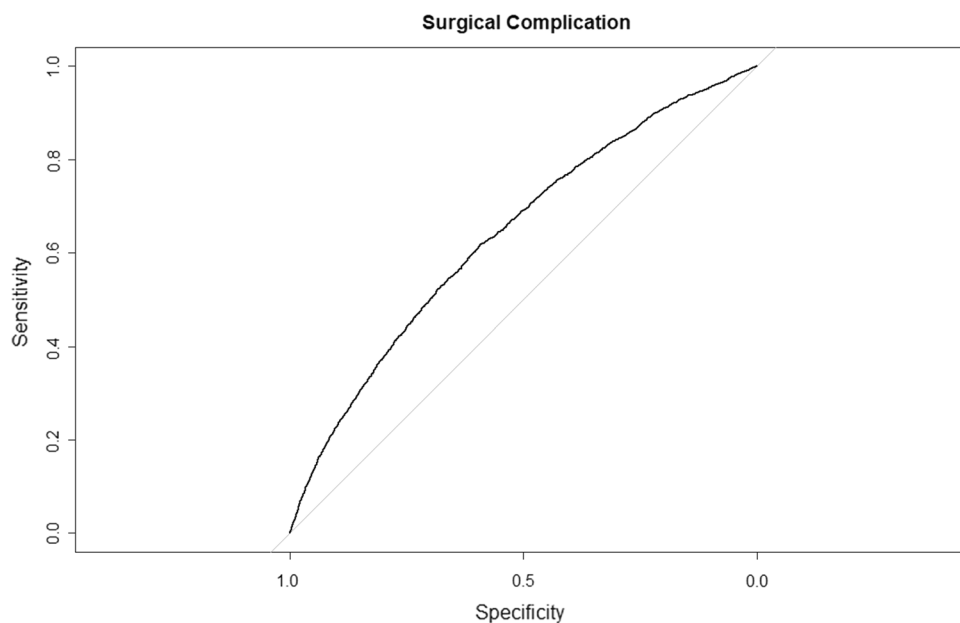

possible explanation for this could be that the complication rates in our dataset were low and therefore the risk of, and accuracy in predicting, an unfavourable event was also very low. In other words, in reality there were no cases with high risk and thus high-risk cases could not be used to train the model. Either way, this phenomenon should be considered,

if a comparatively high probability above 0.25–0.30 is given during the risk assessment, when using these models.

The aforementioned findings were confirmed in sensitivity analyses using single imputation for missing data. In these we found similar tendencies regarding the influence of
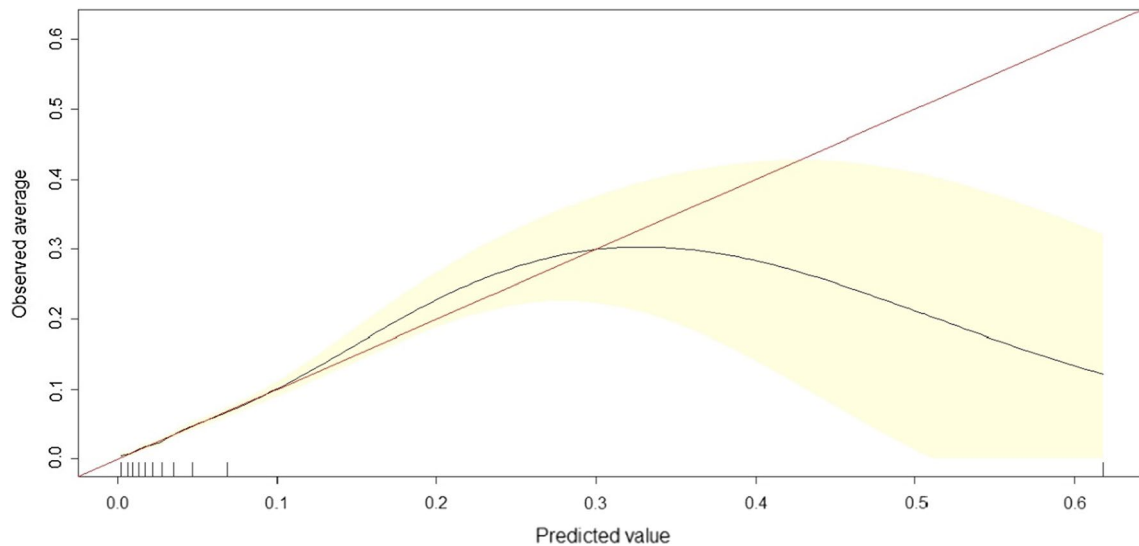
Fig. 4 Calibration plot for general complication model. The y-axis describes the observed average probability of complications; x-axis describes the models corresponding to predicted values. The red line indicates optimal calibration; the black line represents the models' calibration with confidence limits given by the yellow area

Calibration plot surgical complications
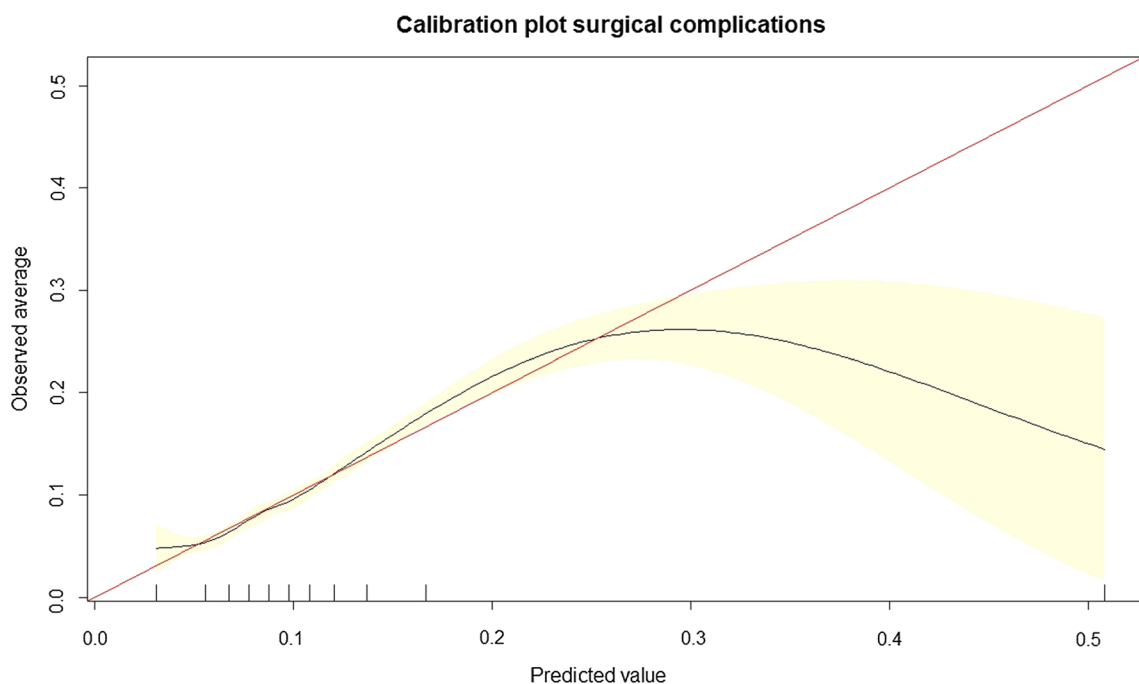
Fig. 5 Calibration plot surgical complication model. The y-axis describes the observed average probability of complications; x-axis describes the models corresponding to predicted values. The red line indicates optimal calibration; the black line represents the models 'calibration with confidence limits as yellow area

the different predictor variables, the calculated AUCs for the ROC analyses and the calibration curves.

Tendencies for overfitting of our models appeared to be very small, as our shrinkage factors were close to 1. This suggests that the models may be applicable to populations outside of the given study group, although this should be verified by external validation.

Reported complication rates in spine surgery for degenerative disease range from 3.7 to 16%, depending on the definition of complication used and the technique being

focused on [48–50]. In our population we found 3.2% medical complications and 10.7% surgical complications. A problem in the current literature is the reliability of complication reporting since there are no generally acknowledged reporting standards [51]. Overall, we found no notable discrepancy with the previously reported complication rates in our population group. When the number of cases in the registry has increased sufficiently to include a variable reflecting, for example, the type or geographical region of the contributing centres, it may be possible to calibrate/customise the model to accommodate differences in the thresholds for reporting.

Strengths and limitations

We view the use of a large multicentre database for model development as a relative advantage regarding the prediction of complications in lumbar spine surgery, when compared with other models previously developed. The model developed by McGirt et al. [4] used a single-centre database. The ACS-NSQIP [13] was developed using a heterogeneous patient population, and by only accepting one Current Procedural Terminology (CPT) code for each risk calculation, it may have underestimated the complexity of the surgery [52, 53], since many spine surgery cases comprise multiple procedures in one operation. In our model, complex spine surgery is taken into account using the modified Mirza index [29, 54]. Ratliff et al. [12] used the large MarketScan dataset, which included mostly younger patients. As the authors themselves conceded, this can be seen as a major limitation, since an increasing number of spine patients are elderly, due to the effects of degenerative disease on the aging spine [55]. A drawback of Lee et al.'s predictive model [9, 10] is that it was developed using only 1476 patients, which may not be a large enough sample. In addition to the above, our prediction of surgical complications included a range of different types, both intraoperatively and prior to discharge, rather than only surgical site infection [9], although our model was not able to specifically predict the likelihood of incurring any particular surgical complication or its severity. Including the modified Mirza index as a measure of surgical invasiveness obviated the need to sub-divide the data to produce separate predictor models for different spine surgical procedures of varying complexity and hence improved the overall power of the results.

A limitation of our study was that the prediction models were not validated using an external dataset. Although we tried to avoid overfitting of our models by using shrinkage, this is not a substitute for an actual external validation.

Another limitation of the study was the complete case analysis approach (i.e. we only included patients for whom there were no missing data), which could have introduced selection bias. The missing data are the result of the incomplete filling out of the Spine Tango surgery forms by the participating surgeons. We tried to address this limitation by performing a sensitivity analysis using single imputation and found our results to be robust. In the present study there was no analysis of complications occurring after hospital discharge, since the collected data were focused on complications during the hospital stay, as recorded on the Tango Surgery form. Finally, surgeons and patients using these models should keep in mind that the developed models tend to overestimate risks when assessing higher risk situations, as discussed in detail above.

We also emphasise that not only the risks, but also the benefits of surgery should be communicated adequately to the patient, such that both can be taken into account when making decisions about surgery. As such, any risk calculations using our model should ideally go hand in hand with estimates of pain and functional improvement when planning surgery.

Implications for future research

To ensure their broader applicability, our models should be externally validated. Also, separate prediction models could be developed for different spine pathologies, which would further increase their applicability. Analogous models could also be built looking at the risk of complications after discharge using the data from the Spine Tango Follow-Up form, which documents complications arising after hospitalisation up to many years' follow-up. Similarly, models predicting the likelihood of clinically relevant improvements in patient-oriented outcomes should be developed.

Conclusion

We were able to build two predictor models that can be used to predict the probability of incurring a complication during or shortly after spine surgery (before discharge). Of the two models, general complications were able to be predicted with greater discriminative ability than surgical complications. Reoperations at the same level were a predominant predictive factor for surgical complications, and a higher ASA score showed the highest odds for general complications. Complication rates were in the expected range, as reported in the literature. A freely available, web-based prediction tool has been developed for the purposes of further testing and validation https://sst.webauthor.com/go/fx/run.cfm?fx=SSTCalculator

The findings of this study are relevant for patient counselling and informed and shared decision making.

Appendix

Unshrunken baseline models

See Tables 6, 7.

Table 6 Unshrunken regression coefficients for the general complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	*p* value
(Intercept)	0.002	0.001–0.005	<0.0001
Age years	1.034	1.027–1.041	<0.0001
Gender=male	0.915	0.786–1.065	0.252
Previous surgeries			
At the same level	0.987	0.789–1.225	0.906
At a different level	1.126	0.925–1.365	0.23
BMI kg/m^2			
20–25	1.163	0.787–1.784	0.469
26–30	0.923	0.624–1.417	0.701
31–35	0.969	0.643–1.51	0.884
>35	1.115	0.708–1.798	0.647
smoker=no	0.91	0.744–1.119	0.362
Morbidity			
ASA2	1.614	1.221–2.166	0.001
ASA3	3.064	2.249–4.226	<0.0001
ASA4	5.863	3.061–10.724	<0.0001
Mirza score	1.03	1.024–1.037	<0.0001
Any additional pathology	1.249	0.998–1.553	0.049
Prophylaxis			
No prophylaxis	0.491	0.174–1.2	0.142
Infection prophylaxis	0.579	0.376–0.941	0.019
Thromboembolism prophylaxis	1.174	0.971–1.428	0.103
Ossification prophylaxis	1.917	0.667–4.349	0.165
Technology used			
Conventional technology	1.325	1.129–1.556	0.001
MISS/LISS	0.908	0.702–1.162	0.453

Table 7 Unshrunken regression coefficients for the surgical complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	p value
(Intercept)	0.045	0.03–0.066	<0.0001
Age years	1.018	1.014–1.021	<0.0001
Gender = male	0.809	0.743–0.882	<0.0001
Previous surgeries			
At the same level	1.945	1.747–2.164	<0.0001
At a different level	0.998	0.877–1.133	0.974
BMI kg/m^2			
20–25	0.961	0.768–1.215	0.735
26–30	0.987	0.789–1.246	0.907
31–35	1.112	0.881–1.417	0.379
> 35	1.303	1.007–1.696	0.046
smoker = no	0.99	0.887–1.107	0.86
Morbidity			
ASA2	1.077	0.954–1.217	0.232
ASA3	1.239	1.062–1.447	0.007
ASA4	1.02	0.544–1.777	0.948
Mirza score	1.01	1.005–1.015	<0.0001
Any additional pathology	1.313	1.145–1.502	<0.0001
Prophylaxis			
No prophylaxis	0.69	0.432–1.088	0.114
Infection prophylaxis	0.792	0.604–1.057	0.102
Thromboembolism prophylaxis	0.847	0.768–0.936	0.001
Ossification prophylaxis	2.241	1.371–3.517	0.001
Technology used			
Conventional technology	1.12	1.021–1.229	0.016
MISS/LISS	1.11	0.975–1.261	0.112

Sensitivity analysis

See Tables 8, 9 and Figs. 6, 7, 8, 9.

Table 8 Sensitivity analysis: shrunken coefficients for the general complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	*p* value
(Intercept)	0	0	<0.0001
Age years	1.03	1.02–1.04	<0.0001
Gender=male	0.94	0.83–1.06	0.32
Previous surgeries			
At the same level	0.91	0.76–1.1	0.33
At a different level	1.18	1.01–1.39	0.043
BMI kg/m²			
20–25	1.07	0.76–1.5	0.70
26–30	0.84	0.6–1.18	0.33
31–35	0.92	0.65–1.31	0.65
>35	0.97	0.66–1.42	0.86
Smoker=no	1.02	0.86–1.21	0.83
Morbidity			
ASA2	1.84	1.47–2.31	<0.0001
ASA3	3.63	2.83–4.67	<0.0001
ASA4	5.97	3.46–10.32	<0.0001
Mirza score	1.03	1.03–1.04	<0.0001
Any additional pathology	1.26	1.05–1.52	0.015
Prophylaxis			
No prophylaxis	0.81	0.41–1.61	0.55
Infection prophylaxis	1.5	1.04–2.15	0.03
Thromboembolism prophylaxis	1.22	1.05–1.43	0.012
Ossification prophylaxis	2.07	1.01–4.25	0.048
Technology used			
Conventional technology	1.22	1.07–1.39	0.004
MISS/LISS	0.8	0.65–1	0.05

Table 9 Sensitivity analysis: shrunken coefficients for the surgical complications model

	Odds Ratio (OR)	95% Confidence interval (CI)	*p* value
(Intercept)	0.03	0.02–0.03	<0.0001
Age years	1.01	1.01–1.02	<0.0001
Gender=male	0.8	0.75–0.86	<0.0001
Previous surgeries			
At the same level	1.93	1.77–2.1	<0.0001
At a different level	1.09	0.98–1.21	0.097
BMI kg/m²			
20–25	1.02	0.85–1.22	0.86
26–30	0.96	0.8–1.15	0.66
31–35	1.08	0.9–1.3	0.40
>35	1.31	1.07–1.61	0.008
smoker=no	0.99	0.9–1.07	0.75
Morbidity			
ASA2	1.26	1.15–1.38	<0.0001
ASA3	1.42	1.26–1.6	<0.0001
ASA4	1.04	0.61–1.77	0.88
Mirza score	1.01	1.01–1.02	<0.0001
Any additional pathology	1.33	1.19–1.49	<0.0001
Prophylaxis			
No prophylaxis	1.11	0.84–1.47	0.47
Infection prophylaxis	1.33	1.11–1.6	0.002
Thromboembolism prophylaxis	0.97	0.9–1.05	0.44
Ossification prophylaxis	2.05	1.4–3.01	0.0002
Technology used			
Conventional technology	0.93	0.86–1	0.044
MISS/LISS	0.9	0.82–1	0.058

General Complication
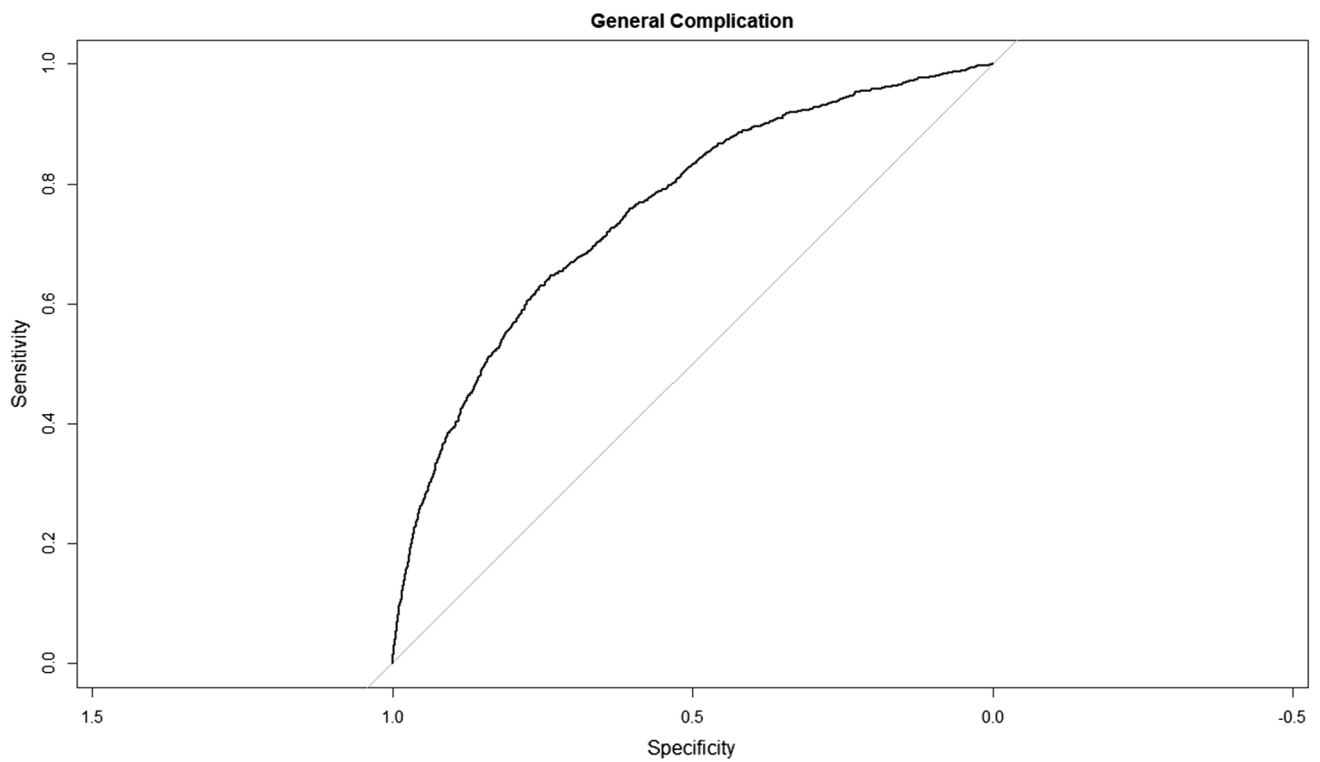
Fig. 6 Sensitivity analysis: ROC curve for general complications

Surgical Complication
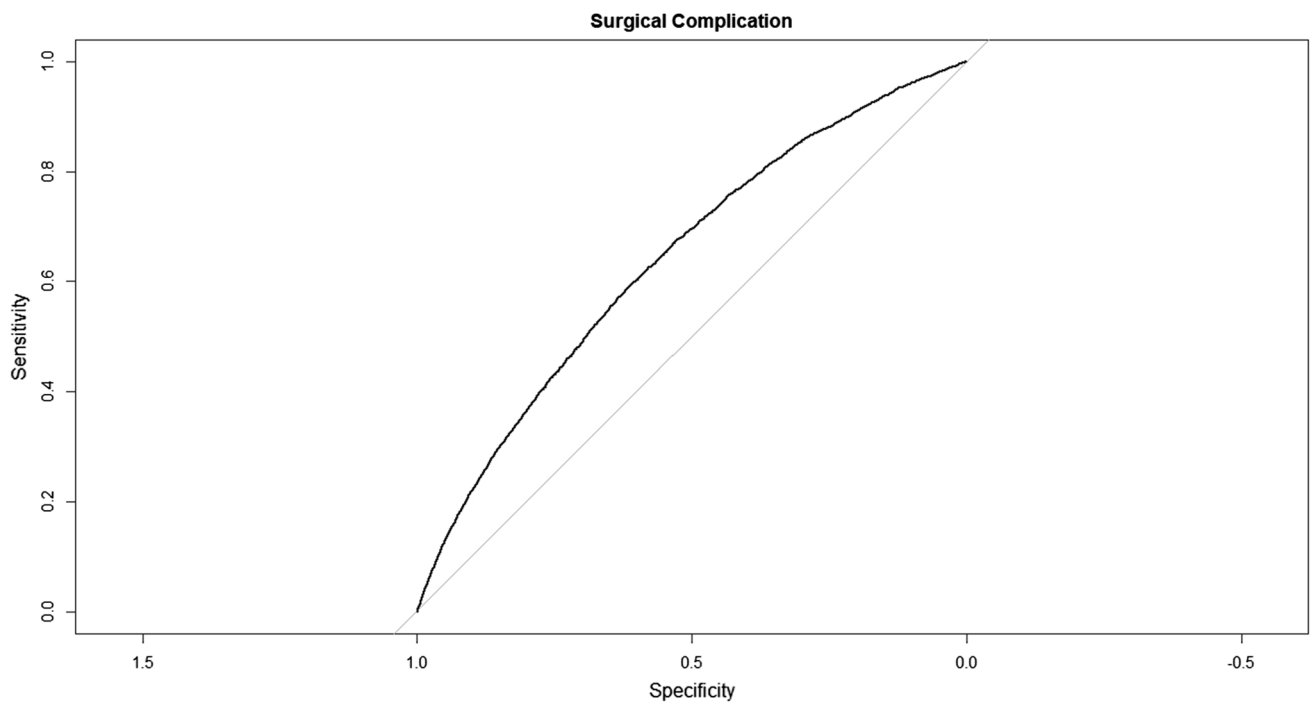
Fig. 7 Sensitivity analysis: ROC curve for surgical complications
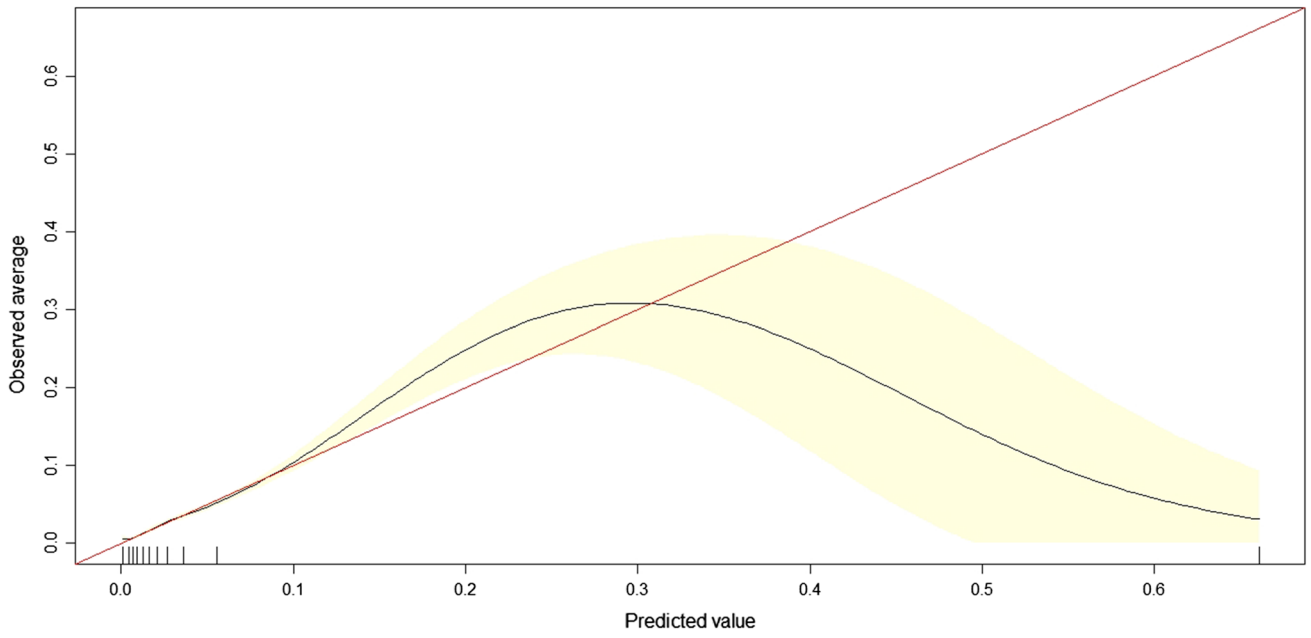
Fig. 8 Sensitivity analysis: calibration plot for general complications
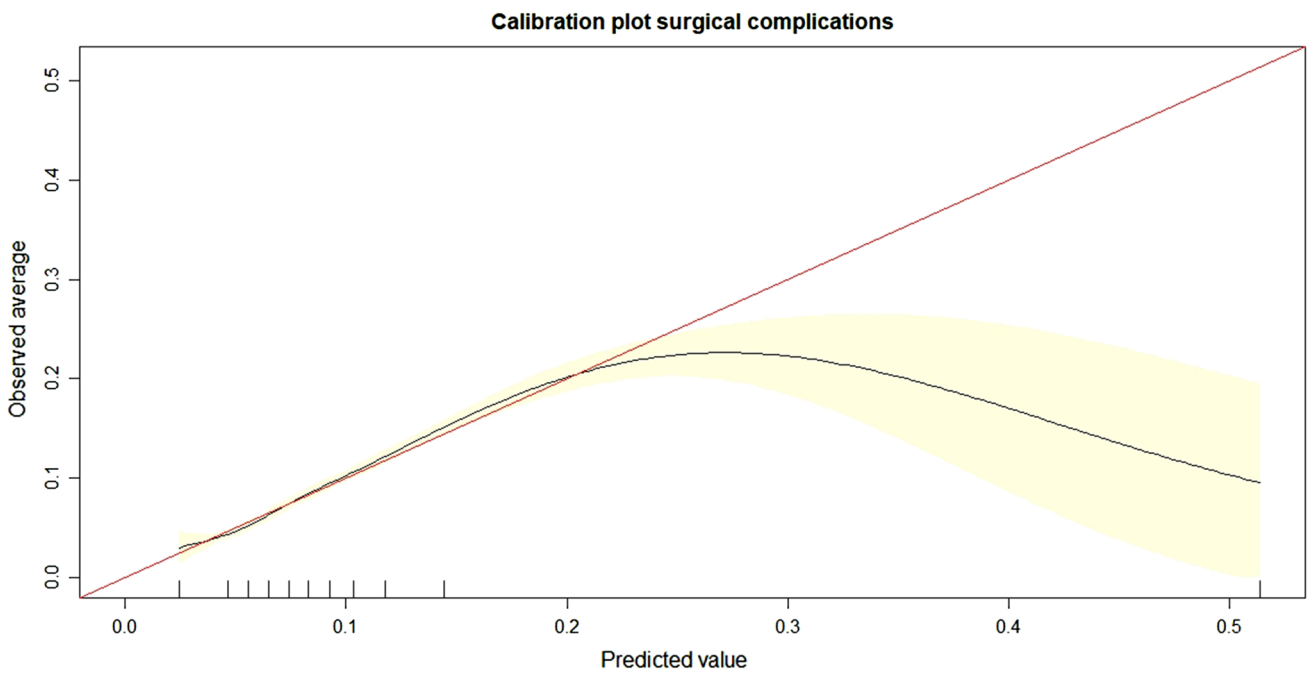
Fig. 9 Sensitivity analysis: calibration plot for surgical complications

Compliance with ethical standards

References

1. Boss EF, Mehta N, Nagarajan N et al (2016) Shared decision making and choice for elective surgical care: a systematic review. Otolaryngol-Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg 154:405–420. https://doi.org/10.1177/0194599815620558

2. Khor S, Lavallee D, Cizik AM et al (2018) Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. JAMA Surg 153:634–642. https://doi.org/10.1001/jamasurg.2018.0072

3. Tetreault LA, Kopjar B, Vaccaro A et al (2013) 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. https://doi.org/10.2106/JBJS.L.01323

4. McGirt MJ, Sivaganesan A, Asher AL, Devin CJ (2015) Prediction model for outcome after low-back surgery: individualized likelihood of complication, hospital readmission, return to work, and 12-month improvement in functional disability. Neurosurg Focus 39:E13. https://doi.org/10.3171/2015.8.FOCUS15338

5. Lee BH, Yang J-H, Lee H-M et al (2016) Surgical outcome predictor in degenerative lumbar spinal disease based on health related quality of life using euro-quality 5 dimensions analysis. Yonsei Med J 57:1214–1221. https://doi.org/10.3349/ymj.2016.57.5.1214

6. Held U, Burgstaller JM, Wertli MM et al (2018) Prognostic function to estimate the probability of meaningful clinical improvement after surgery - Results of a prospective multicenter observational cohort study on patients with lumbar spinal stenosis. PLoS ONE. https://doi.org/10.1371/journal.pone.0207126

7. Chen H-N, Tsai Y-F (2013) A predictive model for disability in patients with lumbar disc herniation. J Orthop Sci Off J Jpn Orthop Assoc 18:220–229. https://doi.org/10.1007/s00776-012-0354-1

8. Spratt KF, Keller TS, Szpalski M et al (2004) A predictive model for outcome after conservative decompression surgery for lumbar spinal stenosis. Eur Spine J 13:14–21. https://doi.org/10.1007/s00586-003-0583-2

9. Lee MJ, Cizik AM, Hamilton D, Chapman JR (2014) Predicting surgical site infection after spine surgery: a validated model using a prospective surgical registry. Spine J 14:2112–2117. https://doi.org/10.1016/j.spinee.2013.12.026

10. Lee MJ, Cizik AM, Hamilton D, Chapman JR (2014) Predicting medical complications after spine surgery: a validated model

using a prospective surgical registry. Spine J 14:291–299. https://doi.org/10.1016/j.spinee.2013.10.043

11. Bekelis K, Desai A, Bakhoum SF, Missios S (2014) A predictive model of complications after spine surgery: the national surgical quality improvement program (NSQIP) 2005–2010. Spine J Off J North Am Spine Soc 14:1247–1255. https://doi.org/10.1016/j.spinee.2013.08.009

12. Ratliff JK, Balise R, Veeravagu A et al (2016) Predicting occurrence of spine surgery complications using "Big Data" modeling of an administrative claims database. JBJS 98:824–834. https://doi.org/10.2106/JBJS.15.00301

13. Bilimoria KY, Liu Y, Paruch JL et al (2013) Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. J Am Coll Surg 217:833-842.e3. https://doi.org/10.1016/j.jamcollsurg.2013.07.385

14. Kasparek MF, Boettner F, Rienmueller A et al (2018) Predicting medical complications in spine surgery: evaluation of a novel online risk calculator. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 27:2449–2456. https://doi.org/10.1007/s00586-018-5707-9

15. Janssen DMC, van Kuijk SMJ, d'Aumerie BB, Willems PC (2018) External validation of a prediction model for surgical site infection after thoracolumbar spine surgery in a Western European cohort. J Orthop Surg 13:114. https://doi.org/10.1186/s13018-018-0821-2

16. Kim JS, Arvind V, Oermann EK et al (2018) Predicting surgical complications in patients undergoing elective adult spinal deformity procedures using machine learning. Spine Deform 6:762–770. https://doi.org/10.1016/j.jspd.2018.03.003

17. Kim JS, Merrill RK, Arvind V et al (2018) Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. Spine 43:853–860. https://doi.org/10.1097/BRS.0000000000002442

18. Zehnder P, Aghayev E, Fekete TF et al (2016) Influence of previous surgery on patient-rated outcome after surgery for degenerative disorders of the lumbar spine. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 25:2553–2562. https://doi.org/10.1007/s00586-016-4383-x

19. Zehnder P, Fekete T, Aghayev E et al (2017) Quantification of the increased risk of complications during surgery for degenerative spinal disorders in patients with prior spine surgery. Orthop Proc. https://doi.org/10.1302/1358-992X.99BSUPP_10.SBPR2016-028

20. Sobottke R, Aghayev E, Röder C et al (2012) Predictors of surgical, general and follow-up complications in lumbar spinal stenosis relative to patient age as emerged from the Spine Tango Registry. Eur Spine J 21:411–417. https://doi.org/10.1007/s00586-011-2016-y

21. Fu K-M, Smith J, Polly D et al (2011) Correlation of higher preoperative American society of anesthesiology grade and increased morbidity and mortality rates in patients undergoing spine surgery: clinical article. J Neurosurg Spine 14:470–474. https://doi.org/10.3171/2010.12.SPINE10486

22. Cizik AM, Lee MJ, Martin BI et al (2012) Using the Spine Surgical Invasiveness Index to Identify Risk of Surgical Site Infection. J Bone Joint Surg Am 94:335–342. https://doi.org/10.2106/JBJS.J.01084

23. Bono OJ, Poorman GW, Foster N et al (2018) Body mass index predicts risk of complications in lumbar spine surgery based on surgical invasiveness. Spine J Off J North Am Spine Soc 18:1204–1210. https://doi.org/10.1016/j.spinee.2017.11.015

24. Phan K, Fadhil M, Chang N et al (2018) Effect of smoking status on successful arthrodesis, clinical outcome, and complications after anterior lumbar interbody fusion (ALIF). World Neurosurg 110:e998–e1003. https://doi.org/10.1016/j.wneu.2017.11.157

25. Martin C, Gao Y, Duchman K, Pugely A (2016) The Impact of current smoking and smoking cessation on short-term morbidity risk after lumbar spine surgery. Spine 41:577–584. https://doi.org/10.1097/BRS.0000000000001281

26. EUROSPINE - Overview. https://www.eurospine.org/spine-tango.htm. Accessed 20 Jun 2020

27. Swiss association of research ethics committees. In: swissethics. https://swissethics.ch/en/. Accessed 20 Jun 2020

28. Pochon L, Kleinstück F, Porchet F, Mannion A (2015) Influence of gender on patient-oriented outcomes in spine surgery. Eur Spine J. https://doi.org/10.1007/s00586-015-4062-3

29. Mirza SK, Deyo RA, Heagerty PJ, et al (2008) Development of an index to characterize the "invasiveness" of spine surgery: validation by comparison to blood loss and operative time. Spine 33:2651–2661; discussion 2662. https://doi.org/https://doi.org/10.1097/BRS.0b013e31818dad07

30. Holzer EM, Aghayev E, O'Riordan D et al (2020) Validation of a surgical invasiveness index in patients with lumbar spinal disorders registered in the Spine Tango registry. Eur Spine J. https://doi.org/10.1007/s00586-020-06651-w

31. Concato J, Peduzzi P, Holford TR, Feinstein AR (1995) Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. J Clin Epidemiol 48:1495–1501. https://doi.org/10.1016/0895-4356(95)00510-2

32. Peduzzi P, Concato J, Feinstein AR, Holford TR (1995) Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 48:1503–1510. https://doi.org/10.1016/0895-4356(95)00048-8

33. Peduzzi P, Concato J, Kemper E et al (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49:1373–1379. https://doi.org/10.1016/s0895-4356(96)00236-3

34. Dunkler D, Sauerbrei W, Heinze G (2016) Global, Parameterwise and Joint Shrinkage Factor Estimation. J Stat Softw 69:1–19. https://doi.org/https://doi.org/10.18637/jss.v069.i08

35. Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer 5:1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

36. Austin PC, Steyerberg EW (2014) Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. Stat Med 33:517–535. https://doi.org/10.1002/sim.5941

37. Steyerberg EW, Vickers AJ, Cook NR et al (2010) Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiol Camb Mass 21:128–138. https://doi.org/10.1097/EDE.0b013e3181c30fb2

38. Hosmer DW, Lemeshow S (1980) Goodness of fit tests for the multiple logistic regression model. Commun Stat - Theory Methods 9:1043–1069. https://doi.org/10.1080/03610928008827941

39. Azur MJ, Stuart EA, Frangakis C, Leaf PJ (2011) Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 20:40–49. https://doi.org/10.1002/mpr.329

40. van Buuren S, Groothuis-Oudshoorn K (2011) Mice: multivariate imputation by chained equations in R. J Stat Softw 45:1–67

41. R: The R project for statistical computing. https://www.r-project.org/. Accessed 20 Jun 2020

42. Localio AR, Goodman SN, Meibohm A et al (2018) Statistical code to support the scientific story. Ann Intern Med 168:828–829. https://doi.org/10.7326/M17-3431

43. Collins GS, Reitsma JB, Altman DG, Moons KGM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD). Circulation 131:211–219. https://doi.org/10.1161/CIRCULATIONAHA.114.014508

44. Li Z, Liu P, Zhang C et al (2019) Incidence, prevalence, and analysis of risk factors for surgical site infection after lumbar fusion surgery: ≥2-year follow-up retrospective study. World Neurosurg 131:e460–e467. https://doi.org/10.1016/j.wneu.2019.07.207

45. Whitmore RG, Stephen JH, Vernick C et al (2014) ASA grade and Charlson Comorbidity Index of spinal surgery patients: correlation with complications and societal costs. Spine J 14:31–38. https://doi.org/10.1016/j.spinee.2013.03.011

46. Mannion AF, Fekete TF, Porchet F et al (2014) The influence of comorbidity on the risks and benefits of spine surgery for degenerative lumbar disorders. Eur Spine J 23:66–71. https://doi.org/10.1007/s00586-014-3189-y

47. Somani S, Capua JD, Kim JS et al (2017) ASA classification as a risk stratification tool in adult spinal deformity surgery: a study of 5805 patients. Glob Spine J 7:719–726. https://doi.org/10.1177/2192568217700106

48. Kobayashi K, Imagama S, Ando K et al (2017) Complications associated with spine surgery in patients aged 80 years or older: Japan association of spine surgeons with ambition (JASA) multicenter study. Glob Spine J 7:636–641. https://doi.org/10.1177/2192568217716144

49. Saleh A, Thirukumaran C, Mesfin A, Molinari RW (2017) Complications and readmission after lumbar spine surgery in elderly patients: an analysis of 2,320 patients. Spine J 17:1106–1112. https://doi.org/10.1016/j.spinee.2017.03.019

50. Flexman AM, Charest-Morin R, Stobart L et al (2016) Frailty and postoperative outcomes in patients undergoing surgery for degenerative spine disease. Spine J 16:1315–1323. https://doi.org/10.1016/j.spinee.2016.06.017

51. Altshuler M, Mueller K, MacConnell A et al (2020) Does minimally invasive spine surgery reduce the rate of perioperative medical complications? A retrospective single-center experience of 1435 degenerative lumbar spine surgeries. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc. https://doi.org/10.1007/s00586-020-06536-y

52. Nasser R, Yadla S, Maltenfort MG et al (2010) Complications in spine surgery: a review. J Neurosurg Spine 13:144–157. https://doi.org/10.3171/2010.3.SPINE09369

53. Sebastian A, Goyal A, Alvi MA et al (2019) Assessing the performance of national surgical quality improvement program surgical risk calculator in elective spine surgery: insights from patients undergoing single-level posterior lumbar fusion. World Neurosurg 126:e323–e329. https://doi.org/10.1016/j.wneu.2019.02.049

54. Merrill RK, Ibrahim JM, Machi AS, Raphael JS (2020) Analysis and review of automated risk calculators used to predict postoperative complications after orthopedic surgery. Curr Rev Musculoskelet Med 13:298–308. https://doi.org/10.1007/s12178-020-09632-0

55. Deyo RA, Mirza SK, Martin BI et al (2010) Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 303:1259–1265. https://doi.org/10.1001/jama.2010.338