


## Original Article

### Corresponding Author


Kyung Hyun Kim

 <https://orcid.org/0000-0002-1338-5523>

Department of Neurosurgery, Spine and Spinal Cord Institute, Gangnam Severance Hospital, Yonsei University College of Medicine, 211 Eonju-ro, Gangnam-gu, Seoul 06273, Korea  
Email: NSKHK@yuhs.ac

### Co-corresponding Author

Sang Hyun Kim

 <https://orcid.org/0000-0003-2643-7035>

Department of Neurosurgery, Ajou University Hospital, Ajou University School of Medicine, 164 World cup-ro, Yeongtong-gu, Suwon 16499, Korea  
Email: shkim709@aumc.ac.kr

Received: August 30, 2022

Revised: December 26, 2022

Accepted: December 28, 2022

See the commentary on "Predicting Mechanical Complications After Adult Spinal Deformity Operation Using a Machine Learning Based on Modified Global Alignment and Proportion Scoring With Body Mass Index and Bone Mineral Density" via <https://doi.org/10.14245/ns.2346296.148>.



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2023 by the Korean Spinal Neurosurgery Society

# Predicting Mechanical Complications After Adult Spinal Deformity Operation Using a Machine Learning Based on Modified Global Alignment and Proportion Scoring With Body Mass Index and Bone Mineral Density

Sung Hyun Noh<sup>1,2</sup>, Hye Sun Lee<sup>3</sup>, Go Eun Park<sup>3</sup>, Yoon Ha<sup>4</sup>, Jeong Yoon Park<sup>5</sup>, Sung Uk Kuh<sup>4</sup>, Dong Kyu Chin<sup>5</sup>, Keun Su Kim<sup>5</sup>, Yong Eun Cho<sup>5</sup>, Sang Hyun Kim<sup>1</sup>, Kyung Hyun Kim<sup>5</sup>

<sup>1</sup>Department of Neurosurgery, Ajou University College of Medicine, Suwon, Korea

<sup>2</sup>Department of Neurosurgery, Yonsei University College of Medicine, Seoul, Korea

<sup>3</sup>Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea

<sup>4</sup>Department of Neurosurgery, Spine and Spinal Cord Institute, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

<sup>5</sup>Department of Neurosurgery, Spine and Spinal Cord Institute, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

**Objective:** This study aimed to create an ideal machine learning model to predict mechanical complications in adult spinal deformity (ASD) surgery based on GAPB (modified global alignment and proportion scoring with body mass index and bone mineral density) factors.

**Methods:** Between January 2009 and December 2018, 238 consecutive patients with ASD, who received at least 4-level fusions and were followed-up for  $\geq 2$  years, were included in the study. The data were stratified into training ( $n = 167$ , 70%) and test ( $n = 71$ , 30%) sets and input to machine learning algorithms, including logistic regression, random forest gradient boosting system, and deep neural network.

**Results:** Body mass index, bone mineral density, the relative pelvic version score, the relative lumbar lordosis score, and the relative sagittal alignment score of the global alignment and proportion score were significantly different in the training and test sets ( $p < 0.05$ ) between the complication and no complication groups. In the training set, the area under receiver operating characteristics (AUROCs) for logistic regression, gradient boosting, random forest, and deep neural network were 0.871 (0.817–0.925), 0.942 (0.911–0.974), 1.000 (1.000–1.000), and 0.947 (0.915–0.980), respectively, and the accuracies were 0.784 (0.722–0.847), 0.868 (0.817–0.920), 1.000 (1.000–1.000), and 0.856 (0.803–0.909), respectively. In the test set, the AUROCs were 0.785 (0.678–0.893), 0.808 (0.702–0.914), 0.810 (0.710–0.910), and 0.730 (0.610–0.850), respectively, and the accuracies were 0.732 (0.629–0.835), 0.718 (0.614–0.823), 0.732 (0.629–0.835), and 0.620 (0.507–0.733), respectively. The random forest achieved the best predictive performance on the training and test dataset.

**Conclusion:** This study created a comprehensive model to predict mechanical complications after ASD surgery. The best prediction accuracy was 73.2% for predicting mechanical complications after ASD surgery. This information can be used to prevent mechanical complications during ASD surgery.

**Keywords:** Machine learning, Adult spinal deformity, Mechanical complication, Body mass index, Bone mineral density, Random forest

## INTRODUCTION

Adult spinal deformity (ASD) is a disorder that is globally prevalent.<sup>1</sup> It is characterized by significant low back/leg pain, stooping, and poor health-related quality of life (HRQoL) in patients with ASD compared with the general population. Although spinal surgery for correcting ASD is invasive, it is effective in symptomatic cases where conservative treatment is often unsuccessful.<sup>2</sup> However, the surgical correction of ASD is a difficult procedure that is known to have a high risk of complications during the surgery and postoperative period.<sup>3</sup> The estimated incidence of morbidity and mortality due to surgical correction is 31.3% and 0.5%, respectively.<sup>3</sup> Since there are many complications of ASD surgery, there are some ideal surgical target parameters such as Scoliosis Research Society-Schwab classification and age-adjusted alignment goals.<sup>4,5</sup> There are also formulas, such as the global alignment and proportion (GAP) score, which predict mechanical complications after ASD surgery, and the modified global alignment and proportion scoring with body mass index and bone mineral density (GAPB) system, which combines body mass index (BMI) and bone mineral density (BMD) with the GAP score.<sup>6,7</sup>

Most studies have been performed using simple statistical techniques such as linear regression and logistic regression, and in practice, they provide information on mean values that do not properly reflect the characteristics of the population. However, in the past few years, the medical field has increasingly adopted computational techniques that allow the processing of large amounts of data and the creation of complex mathematical models that describe the relationships between different variables. The idea behind artificial intelligence is to create a system that mimics the natural ability of humans to continuously learn as they access new data and apply it to new situations in the future. Our research team reported that GAPB predicts mechanical complications better than other systems related to ASD.<sup>6</sup> This study aimed to create an ideal machine learning model to predict mechanical complications in ASD surgery based on the GAPB system.

## MATERIALS AND METHODS

### 1. Patient Population

This was a retrospective analysis of surgically treated patients with ASD enrolled from 2009 to 2017. This study was approved by the Institutional Review Board (IRB) of the Ajou University Hospital (IRB No. 2022-0546-008). Written informed consent

was obtained from all participants. The inclusion criteria were as follows: patients who underwent ASD surgery to correct sagittal imbalance; the presence of one of the following radiological criteria, including coronal Cobb angle  $>20^\circ$ , sagittal vertical axis  $>5$  cm, pelvic tilt (PT)  $>25^\circ$ , and/or thoracic kyphosis  $>60^\circ$ , and/or pelvic incidence minus lumbar lordosis (PI-LL)  $>10^\circ$ ; use of posterior spinal fixation and instruments with ASD surgery at  $\geq$  level 4; and patients with a follow-up period of  $\geq 2$  years. The exclusion criteria were patients with ASD due to syndrome, autoimmune disease, infection, tumor, or other pathological conditions. Between January 2009 and December 2017, 491 patients with ASD underwent ASD surgery at our hospital. Among them, 253 patients with a follow-up period of  $<2$  years, patients without corrective surgery for ASD, and those with a surgical level of  $\leq 3$  were excluded. Between January 2009 and December 2017, 238 consecutive patients with sagittal imbalance who underwent ASD surgery were ultimately included in the study.

### 2. Data Collection

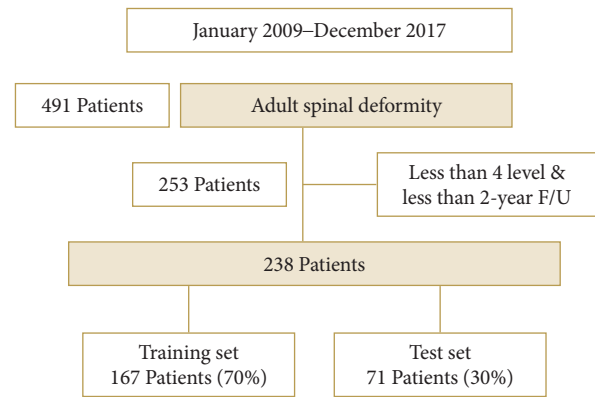
Demographic data, radiologic parameters, surgical characteristics, HRQoL data were collected for all 238 patients included in the electronic medical records. Demographic data included age, sex, BMI, BMD, and GAP score variables. Yilgor et al.<sup>7</sup> created the GAP score. The overall goal of the GAP score is to achieve patient-specific spine-pelvic alignment guidance, and the GAP score predicts mechanical complications. After that, Noh et al.<sup>8,9</sup> made GAPB including BMI and BMD in GAP. Factors frequently used to predict mechanical complications after ASD were used to create an artificial intelligence model.

The following sagittal alignment parameters were measured: PI, PT, lumbar vertebral lordosis (LL [L1-S1]), PI-LL, and global tilt. Radiographic measures included preoperative, postoperative, and final follow-up alignment parameters. We defined mechanical complications after ASD surgery as the following (proximal junctional kyphosis, proximal junctional failure, distal junctional failure, distal junctional kyphosis, rod fracture, implant-related complications) and investigated their prevalence. Proximal junctional kyphosis was defined as a  $\geq 10^\circ$  increase in kyphosis between upper Instrumented vertebra (UIV) and UIV+2 between the early postoperative and 2-year follow-up radiographs. Proximal junctional failure was defined as a fracture of UIV or UIV+1, withdrawal of the instrument in UIV, and/or sagittal subluxation. Distal junctional kyphosis/failure referred to a  $\geq 10^\circ$  increase in kyphosis angle between lowest instrumented vertebra (LIV) and LIV-1, and/or withdrawal of the apparatus from the LIV. Rod breakage referred to

single or double rod breakage. Implant-related complications included other radiographic implant-related complications such as screw loosening, breakage, pullout, or interbody graft, hook, or screw leave. HRQoL was measured using the Oswestry Disability Index, the Scoliosis Research Society-22 Spinal Malformation Questionnaire, and Short Form-36.

**3. Prediction Models and Evaluation**

The patients were randomly divided into training (n = 167, 70%) and test (n = 71, 30%) datasets (Fig. 1). The training set was used to develop the model, and the test set was used to evaluate the model. Among the models that can be implemented



**Fig. 1.** Flowchart of the patients in our study. F/U, follow-up.

**Table 1.** Patient demographics

Variable	Training set (n = 167)			Test set (n = 71)		
	Complication		p-value	Complication		p-value
	No (n = 96)	Yes (n = 71)		No (n = 42)	Yes (n = 29)	
Age (yr)	67.18 ± 7.07	68.63 ± 7.99	0.215	66.40 ± 6.64	67.72 ± 7.49	0.438
Sex						
Male	12	9		8	5	
Female	84	62	0.973	34	24	0.847
BMI (kg/m <sup>2</sup> )	23.54 ± 2.92	24.70 ± 3.05	0.013*	23.39 ± 2.53	25.02 ± 2.79	0.013*
BMD (T-score)	-1.57 ± 0.85	-2.60 ± 0.98	< 0.001*	-1.85 ± 0.83	-2.41 ± 0.86	0.007*
GAP score						
Relative pelvic version score			0.003*			0.002*
0	41	17		24	5	
1	10	6		4	3	
2	35	25		11	11	
3	10	23		3	10	
Relative lumbar lordosis score						
0	42	13		19	5	
2	45	21		18	9	
3	9	37	< 0.001*	5	15	0.001*
Lordosis distribution index score						
0	54	33		26	13	
1	9	6		3	5	
2	15	16		7	1	
3	18	16	0.547	6	10	0.045*
Relative spinopelvic alignment score						
0	35	15		19	4	
1	51	31		17	8	
3	10	25	< 0.001*	6	17	< 0.001*

Values are presented as mean ± standard deviation or number.

GAP, global alignment and proportion; BMI, body mass index; BMD, bone mineral density.

\*p < 0.05, statistically significant differences.

with R, we compared logistic regression, which is widely used conventionally, gradient boosting, which is a representative boosting method, random forest, which is a representative bagging method, and deep neural network, which has recently become an issue. We performed 4 analyses to classify the occurrence of complications. First, univariable and multivariable logistic regressions were used. Variables with  $p < 0.05$  in the univariable analysis were entered in the multivariable analysis. The final multivariable model was determined using a stepwise variable selection method. Second, the gradient boosting model was created with the R package “xgboost,” and variable importance was visualized. For this analysis, a maximum tree depth of 2, learning rate of 0.3, and number of boosting of 20 were considered. Third, random forest classification was performed using the R package “random forest.” For this analysis, the number of

trees was set to 500, and the number of variables used in each 3 was set to 5, which had the largest Kappa value. Fourth, a deep neural network was used via the R package “nnet.” For this analysis, a hidden layer of 10 was employed.

Diagnostic performance was evaluated using the area under receiver operating characteristic (AUROC), area under precise recall curve (AUPRC), accuracy, sensitivity, and specificity for each dataset. To calculate the accuracy, sensitivity, and specificity, the optimal cutoff points were computed using Youden index. Comparisons of AUROC, AUPRC, accuracy, sensitivity, and specificity were performed using generalized estimating equations.

#### 4. Statistical Analysis

Descriptive statistics are presented as frequencies and percent-

**Table 2.** Logistic regression

Variable	Univariable model		Multivariable model 3		Multivariable model 4	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Age	1.03 (0.98–1.07)	0.215	1.284 (1.073–1.536)	0.006*	1.284 (1.073–1.536)	0.006*
Sex						
Male	Reference					
Female	0.98 (0.39–2.48)	0.973				
BMI (kg/m <sup>2</sup> )	1.14(1.03–1.27)	0.015*	1.13 (0.99–1.30)	0.080		
BMD (T-score)	0.26 (0.16–0.41)	< 0.001*	0.28 (0.16–0.46)	< 0.001*	0.28 (0.17–0.47)	< 0.001*
Relative pelvic version score						
0	Reference		Reference			
1	1.45 (0.45–4.61)	0.532	1.42 (0.35–5.73)	0.620		
2	1.72 (0.80–3.70)	0.163	0.87 (0.29–2.58)	0.798		
3	5.55 (2.18–14.10)	< 0.001*	0.22 (0.04–1.26)	0.090		
Relative lumbar lordosis score						
0	Reference		Reference		Reference	
2	1.51 (0.67–3.39)	0.320	1.71 (0.54–5.38)	0.360	1.43 (0.57–3.55)	0.444
3	13.28 (5.10–34.62)	< 0.001*	28.81 (4.77–174.05)	< 0.001*	11.02 (3.80–31.98)	< 0.001*
Lordosis distribution index						
0	Reference					
1	1.09 (0.36–3.34)	0.879				
2	1.75 (0.76–3.99)	0.187				
3	1.45 (0.65–3.24)	0.359				
Relative sagittal alignment score						
0	Reference		Reference			
1	1.42 (0.67–3.01)	0.363	0.94 (0.33–2.65)	0.906		
3	5.83 (2.25–15.08)	< 0.001*	1.22 (0.24–6.11)	0.813		

OR, odds ratio; CI, confidence interval; GAP, global alignment and proportion; BMI, body mass index; BMD, bone mineral density.

\* $p < 0.05$ , statistically significant differences.

ages for categorical variables and as means and standard deviations for continuous variables. To compare the characteristics of patients in the complication and no complication groups, the chi-square test (or Fisher exact test) was used for categorical variables and an independent 2-sample t-test was used for continuous variables. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA). Statistical significance was set at  $p < 0.05$ .

## RESULTS

### 1. Patient Demographics

Two hundred thirty-eight patients underwent ASD surgery (204 females [86%], 34 males [14%]); their demographic data are shown in Table 1. Of those patients, 167 (70.2%) were assigned to the training set and 71 (29.8%) to the test set. The patients' average age and follow-up period were  $67.1 \pm 6.17$  years and  $28.54 \pm 4.25$  months, respectively. The mean ages of patients in the training and test sets were  $67.80 \pm 7.49$  years and  $66.94 \pm 6.98$  years, respectively. When comparing the groups with and without complications in the training set, BMI, BMD, the relative pelvic version score, the relative lumbar lordosis score, and the relative sagittal alignment score were statistically significant. When comparing the groups with and without complications in the test set, BMI, BMD, the relative pelvic version score, the relative lumbar lordosis score, and the relative sagittal alignment score were statistically significant. When comparing the group

with and without complications in the test set, BMI, BMD, the relative pelvic version score, the relative lumbar lordosis score, the lordosis distribution index score, and the relative sagittal alignment score were statistically significant.

### 2. Logistic Regression Model

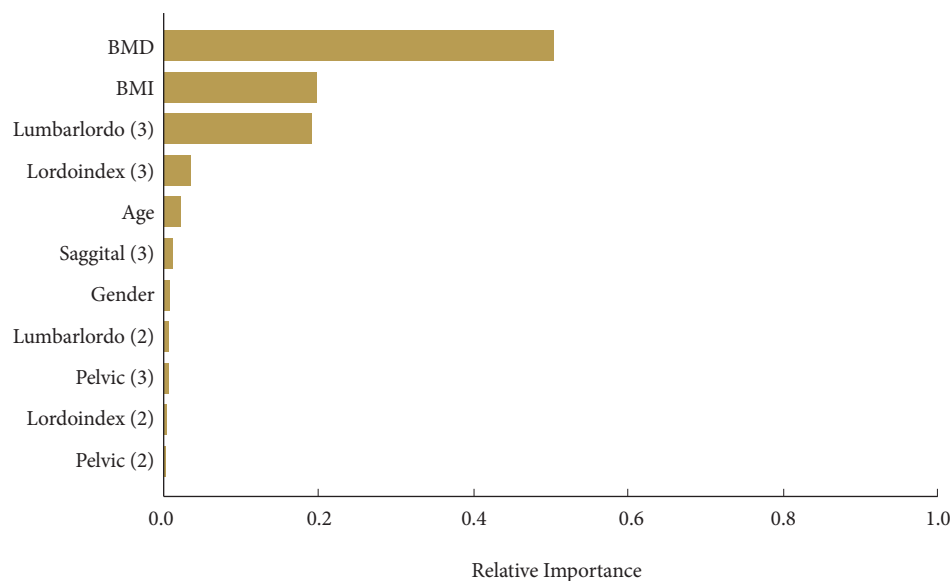
The results of the univariate and multivariate logistic regression analyses are presented in Table 2. The following variables were significantly related to mechanical complications of ASD surgery in univariate logistic regression: BMI, BMD, relative pelvic version score, relative lumbar lordosis score, and relative sagittal alignment score. In the multivariate logistic regression, BMD and relative lumbar lordosis score were significantly related to mechanical complications of ASD surgery.

### 3. Gradient Boosting Model

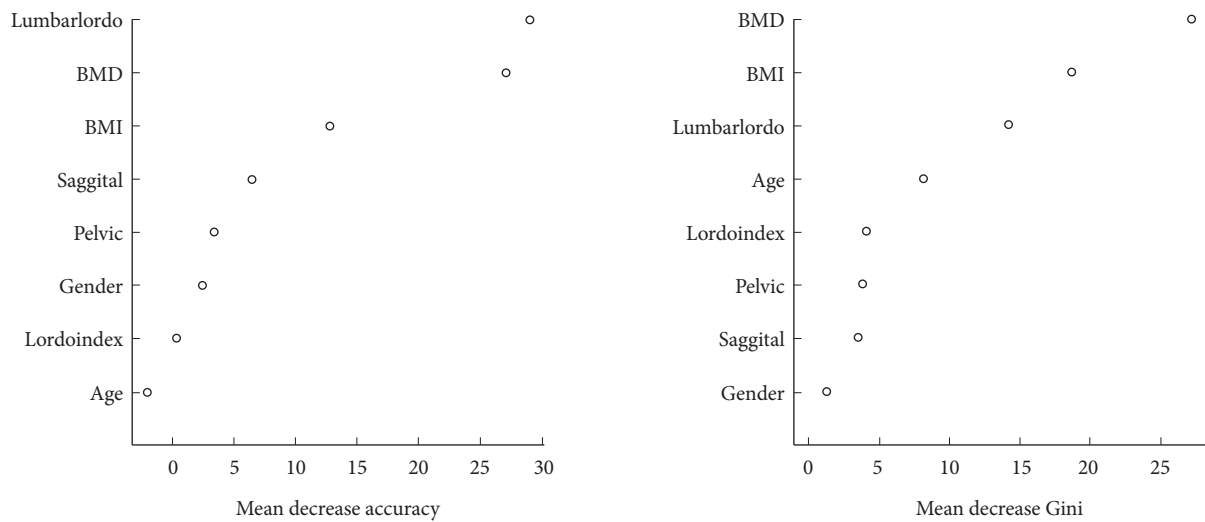
The results of the gradient boosting analysis are shown in Fig. 2. BMI, BMD, and relative lumbar lordosis score were the most important variables in the gradient boosting model.

### 4. Random Forest Model

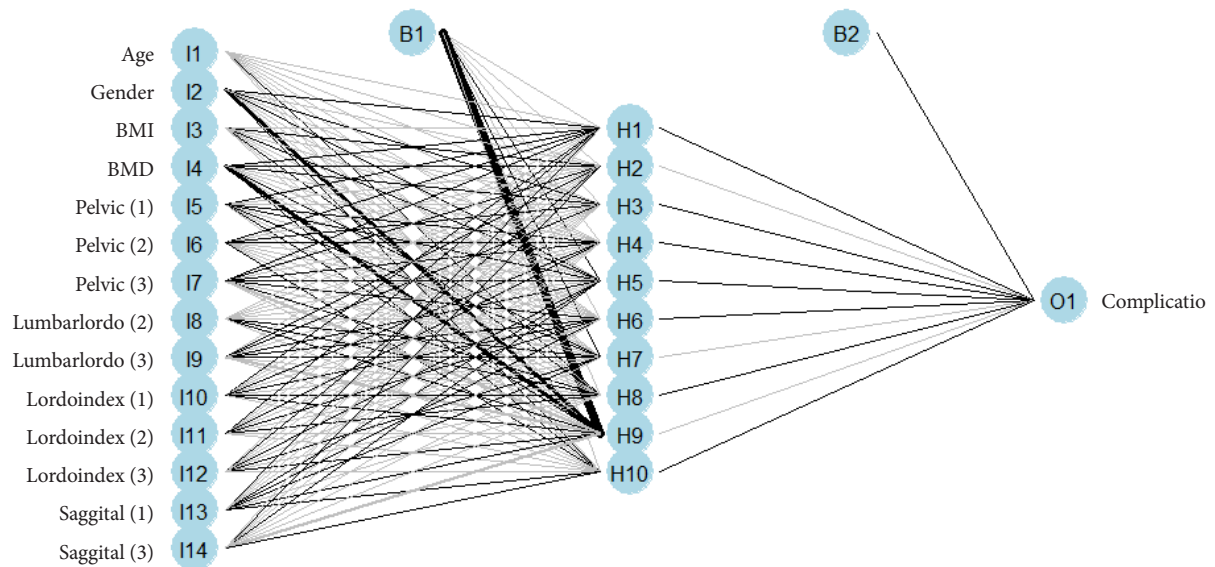
The results of the random forest analyses are shown in Fig. 3. BMI, BMD, and relative lumbar lordosis score were the most important variables in the random forest model. Since random forest has the possibility of overfitting in the training set, it must be interpreted carefully considering the validation result.



**Fig. 2.** Results of the gradient boosting model. The most important variables in the model were BMI, BMD, and relative lumbar lordosis score. BMD, bone mineral density; BMI, body mass index; Lumbarlordo, relative lumbar lordosis score; Lordoindex, lordosis distribution index score; Saggital, relative spinopelvic alignment score; Pelvic, relative pelvic version score.



**Fig. 3.** Results of the random forest model. The most important variables in this model were BMI, BMD, and relative lumbar lordosis score. BMD, bone mineral density; BMI, body mass index; Lumbarlordo, relative lumbar lordosis score; Lordoindex, lordosis distribution index score; Saggital, relative spinopelvic alignment score; Pelvic, relative pelvic version score.



**Fig. 4.** Results of the gradient boosting model. The most important variables in this model were lordosis distribution index score and relative sagittal alignment score. BMD, bone mineral density; BMI, body mass index; Lumbarlordo, relative lumbar lordosis score; Lordoindex, lordosis distribution index score; Saggital, relative spinopelvic alignment score; Pelvic, relative pelvic version score.

**5. Deep Neural Network Model**

The results of the deep neural network analyses are shown in Fig. 4. The most important variables in this model were the lordosis distribution index score and relative sagittal alignment score.

**6. Diagnostic Performance of the Machine Learning Models**

The AUROCs and AUPRCs for the 4 machine learning models are presented in Table 3. In the training set, the AUROCs for

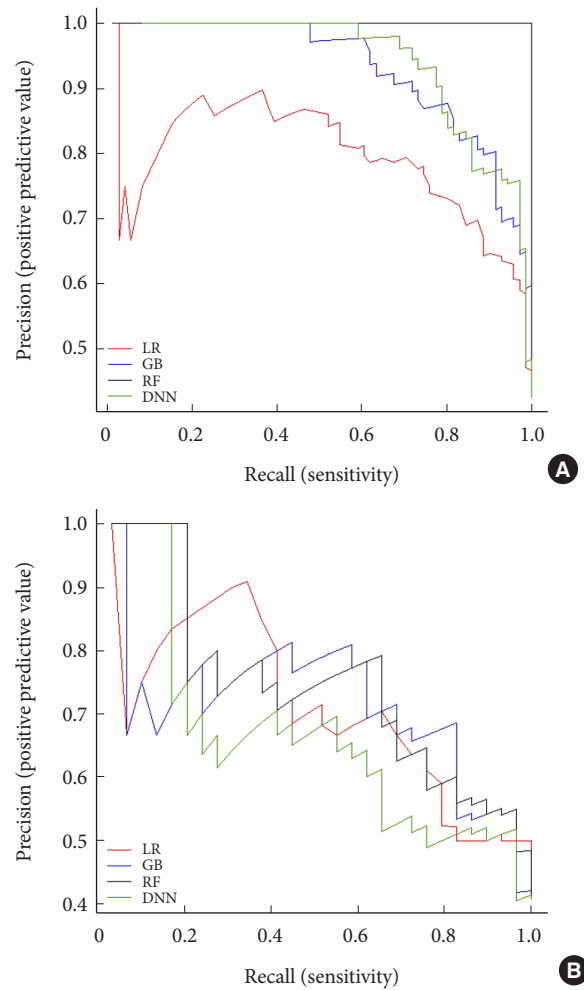
logistic regression, gradient boosting, random forest, and deep neural network model were 0.871 (0.817–0.925), 0.942 (0.911–0.974), 1.000 (1.000–1.000), and 0.947 (0.915–0.980), respectively, the AUPRCs for logistic regression, gradient boosting, random forest, and deep neural network model were 0.793 (0.677–0.895), 0.93 (0.878–0.965), 1.000 (1.000–1.000), and 0.942 (0.898–0.972), respectively, and the accuracies were 0.784 (0.722–0.847), 0.868 (0.817–0.920), 1.000 (1.000–1.000), and 0.856 (0.803–0.909),



**Table 3.** Diagnostic performance between machine learning model

Model	Training set					Test set						
	Cutoff point	AUROC (95% CI)	AUPRC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cutoff point	AUROC (95% CI)	AUPRC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Logistic regression	>0.311	0.871 (0.817–0.925)	0.793 (0.677–0.895)	0.784 (0.722–0.847)	0.873 (0.796–0.951)	0.719 (0.629–0.809)	>0.311	0.785 (0.678–0.893)	0.711 (0.523–0.87)	0.732 (0.629–0.835)	0.759 (0.603–0.914)	0.714 (0.578–0.851)
Gradient boosting	>0.405	0.942 (0.911–0.974)	0.93 (0.878–0.965)	0.868 (0.817–0.920)	0.915 (0.851–0.980)	0.833 (0.759–0.908)	>0.405	0.808 (0.702–0.914)	0.717 (0.529–0.89)	0.718 (0.614–0.823)	0.828 (0.690–0.965)	0.643 (0.498–0.788)
Random forest	>0.483	1.000 (1.000–1.000)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	1.000 (1.000–1.000)	>0.483	0.810 (0.710–0.910)	0.748 (0.554–0.882)	0.732 (0.629–0.835)	0.759 (0.603–0.914)	0.714 (0.578–0.851)
Deep neural network	>0.406	0.947 (0.915–0.980)	0.942 (0.898–0.972)	0.856 (0.803–0.909)	0.972 (0.933–1.000)	0.771 (0.687–0.855)	>0.406	0.730 (0.610–0.850)	0.667 (0.475–0.818)	0.620 (0.507–0.733)	0.897 (0.786–1.000)	0.429 (0.279–0.578)

AUROC, area under receiver operating characteristic; AUPRC, area under precise recall curve; CI, confidence interval.



**Fig. 5.** Area under precise recall curve of each model in the training set (A) and test set (B). LR, logistic regression; GB, gradient boosting; RF, random forest; DNN, deep neural network.

respectively. In the test set, the AUROCs for the same models were 0.785 (0.678–0.893), 0.808 (0.702–0.914), 0.810 (0.710–0.910), and 0.730 (0.610–0.850), respectively, the AUPRCs for logistic regression, gradient boosting, random forest, and deep neural network model were 0.711 (0.523–0.87), 0.717 (0.529–0.89), 0.748 (0.554–0.882), and 0.667 (0.475–0.818), respectively, and the accuracies were 0.732 (0.629–0.835), 0.718 (0.614–0.823), 0.732 (0.629–0.835), and 0.620 (0.507–0.733), respectively. The random forest achieved the best predictive performance on the training and test dataset. Fig. 5 shows the AUPRCs of each model in the training and test sets.

## DISCUSSION

The prevalence of mechanical complications, with radiologic

and clinical manifestations, after surgery for adult spinal deformities is reported to be 30%, and more than 50% of these patients undergo revision surgery for treatment.<sup>10</sup> Soroceanu et al.<sup>11</sup> reported that radiographic and implant-related complications accounted for 31.7%, and in 52.6% of these complications, reoperation for mechanical correction was required. There are many aspects of ASD surgery with notable variability, including the occurrence of complications and outcomes.<sup>12</sup> GAPB is a system that is used to predict mechanical complications that occur after ASD surgery, including both patient-specific and radiological factors.<sup>6</sup> In this study, we constructed a model to predict mechanical complications after ASD surgery using GAPB factors. The GAPB system, including BMI and BMD, showed improved predictability of predicting mechanical complications compared to the GAP scoring system.<sup>8</sup> In particular, Noh et al.<sup>9</sup> reported that GAPB better predicted mechanical complications in the moderately disproportioned and severely disproportioned groups in GAP. Park et al.<sup>13</sup> reported that osteoporosis and obesity are important risk factors for proximal junctional kyphosis, proximal junctional failure and other mechanical complications. Since most elderly patients in ASD surgery have low muscle mass and severe osteoporosis, BMI and osteoporosis are essential when discussing mechanical complications. Recently, several studies using deep learning algorithms, such as random forest, gradient boosting, and neural networks, have been conducted for the spine.<sup>14</sup> Yagi et al.<sup>15</sup> created a postsurgical complication prediction model for ASD surgery in adults using spinal alignment, demographic data, and surgical invasiveness; 170 participants were enrolled in this study. A decision tree for 2-year postoperative complications was constructed and confirmed by splitting data in a 7:3 ratio for training and testing, with the external validation of 25 ASD patients who underwent surgery at different hospitals.<sup>15</sup> For the test sample, the predictive model was 92% accurate, the AUC was 0.963, and the external validation was 84% accurate. Lafage et al.<sup>16</sup> created a machine learning model to determine the upper vertebra in ASD surgery. The samples were stratified into 3 groups: 70% for training, 15% for validation, and 15% for performance testing. A neural network model was used, and the results showed an accuracy of 81.0%, precision of 87.5%, and recall of 87.5%. Pelisé et al.<sup>17</sup> created a model to predict the incidence of adverse events after ASD surgery using a random forest model. The model was trained using 80% of the data for the training set and 20% for the test set and showed adequate predictive accuracy, with AUCs ranging from 0.67 to 0.92.<sup>17</sup> Durand et al.<sup>18</sup> created a model for predicting blood transfusion following sur-

gery for adult spinal deformities. A total of 1,029 patients were analyzed and divided into datasets for training ( $n = 824$ ) and validation ( $n = 205$ ). The random forest model showed an AUC of 0.85 (95% confidence interval, 0.80–0.90) and was reported to show better predictive ability than single-decision tree models.<sup>18</sup> Ames et al. created a model to predict the cost of surgery for ASD. The regression tree and random forest models were used to predict the occurrence of treatment costs exceeding \$100,000.<sup>19</sup> The results of the regression tree analysis using CTREE resulted in an adjusted  $R^2$  value of 56% at 90 days and 35.6% at 2 years of direct cost forecasting. Random C-forest regression analysis showed an adjusted  $R^2$  value of 57.4% at 90 days and 28.8% at 2 years of direct cost forecasts. Peng et al.<sup>20</sup> created a model to predict proximal junctional kyphosis after surgery in adolescent patients with idiopathic scoliosis. The random forest has great value for predicting the individual risk of developing proximal junctional kyphosis after long instrumentation and fusion surgery in patients with Lenke 5 adolescent idiopathic scoliosis. Jain created a model to predict discharge delay, medical complications, and readmission within 90 days after long-segment posterior lumbar spine fusion surgery<sup>21</sup> using logistic regression, random forest, and elastic net. In our study, we created a model to predict the mechanical complications that occur after ASD surgery. We used logistic regression, gradient boosting, random forest, and deep neural networks. Important factors were BMD, BMI, relative lumbar lordosis score, lordosis distribution index score, and relative sagittal alignment score. The patients were randomly divided into training (70%) and test (30%) datasets. In the training set, the AUROC for random forest was 1.000 and the accuracy was 1.000. In the test set, the AUROC for random forest was 0.81 and the accuracy was 0.732. Random forest achieved the best predictive performance on the training and test dataset.

This study has several limitations. Because our models were built using retrospective data, future efforts to update these models are required. Additionally, the reasons for mechanical complications after ASD correction are multifactorial. Many factors affect the outcome of surgery, including the surgical method, upper level instrumentation, muscle mass, and various underlying conditions. These factors were excluded when the model was created.

However, the GAPB system is helpful in predicting mechanical complications after ASD surgery.<sup>6</sup> Noh et al.<sup>9</sup> reported that the GAPB system was more meaningful in the moderately disproportioned and severely disproportioned GAP groups. We believe that it will be helpful to develop models that predict me-



mechanical complications through machine learning. And the overfitting problem caused by using small data samples is a limitation of this study. We will study with more data samples later.

## CONCLUSION

This study created a comprehensive model to predict mechanical complications after ASD surgery. The best prediction accuracy was 73.2% for predicting mechanical complications after ASD surgery. This information can be used to prevent mechanical complications during ASD surgery.

## NOTES

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Author Contribution:** Conceptualization: SHN, SHK, KHK; Data curation: SHN, HSL, GEP, SHK, KHK; Formal analysis: SHN, SHK, KHK; Methodology: SHN, HSL, GEP, SHK, KHK; Project administration: SHN, HSL, GEP, SHK, KHK; Visualization: SHN, YH, JYP, SUK, DKC, KSK, YEC, SHK, KHK; Writing - original draft: SHN; Writing - review & editing: SHN.

## ORCID

Sung Hyun Noh: 0000-0003-2732-0031

Hye Sun Lee: 0000-0001-6328-6948

Go Eun Park: 0000-0002-6670-5500

Yoon Ha: 0000-0002-3775-2324

Jeong Yoon Park: 0000-0002-3728-7784

Sung Uk Kuh: 0000-0003-2566-3209

Dong Kyu Chin: 0000-0002-9835-9294

Keun Su Kim: 0000-0002-3384-5638

Yong Eun Cho: 0000-0001-9815-2720

Sang Hyun Kim: 0000-0003-2643-7035

Kyung Hyun Kim: 0000-0002-1338-5523

## REFERENCES

1. Grabel ZJ, Hart RA, Clark AJ, et al. Adult spinal deformity knowledge in orthopedic spine surgeons: impact of fellowship training, experience, and practice characteristics. *Spine Deform* 2018;6:60-6.
2. Soroceanu A, Burton DC, Oren JH, et al. Medical complications after adult spinal deformity surgery: incidence, risk factors, and clinical impact. *Spine (Phila Pa 1976)* 2016;41:1718-23.
3. Lee NJ, Kothari P, Kim JS, et al. Early complications and outcomes in adult spinal deformity surgery: an NSQIP study based on 5803 patients. *Global Spine J* 2017;7:432-40.
4. Schwab F, Patel A, Ungar B, et al. Adult spinal deformity-post-operative standing imbalance: how much can you tolerate? An overview of key parameters in assessing alignment and planning corrective surgery. *Spine (Phila Pa 1976)* 2010;35:2224-31.
5. Lafage R, Schwab F, Challier V, et al. Defining spino-pelvic alignment thresholds: should operative goals in adult spinal deformity surgery account for age? *Spine (Phila Pa 1976)* 2016;41:62-8.
6. Noh SH, Ha Y, Obeid I, et al. Modified global alignment and proportion scoring with body mass index and bone mineral density (GAPB) for improving predictions of mechanical complications after adult spinal deformity surgery. *Spine J* 2020;20:776-84.
7. Yilgor C, Sogunmez N, Boissiere L, et al. Global alignment and proportion (GAP) score: development and validation of a new method of analyzing spinopelvic alignment to predict mechanical complications after adult spinal deformity surgery. *J Bone Joint Surg Am* 2017;99:1661-72.
8. Noh SH, Ha Y, Obeid I, et al. Modified global alignment and proportion scoring with body mass index and bone mineral density (GAPB) for improving predictions of mechanical complications after adult spinal deformity surgery. *Spine J* 2020;20:776-84.
9. Noh SH, Ha Y, Park JY, et al. Modified global alignment and proportion scoring with body mass index and bone mineral density analysis in global alignment and proportion score of each 3 categories for predicting mechanical complications after adult spinal deformity surgery. *Neurospine* 2021;18:484-91.
10. Kwan KYH, Lenke LG, Shaffrey CI, et al. Are higher global alignment and proportion scores associated with increased risks of mechanical complications after adult spinal deformity surgery? An external validation. *Clin Orthop Relat Res* 2021;479:312-20.
11. Soroceanu A, Diebo BG, Burton D, et al. Radiographical and implant-related complications in adult spinal deformity surgery: incidence, patient risk factors, and impact on health-related quality of life. *Spine (Phila Pa 1976)* 2015;40:1414-21.
12. Ames CP, Smith JS, Pellisé F, et al. Artificial intelligence based hierarchical clustering of patient types and intervention cat-

- egories in adult spinal deformity surgery: towards a new classification scheme that predicts quality and value. *Spine (Phila Pa 1976)* 2019;44:915-26.
13. Park SJ, Lee CS, Chung SS, et al. Different risk factors of proximal junctional kyphosis and proximal junctional failure following long instrumented fusion to the sacrum for adult spinal deformity: survivorship analysis of 160 patients. *Neurosurgery* 2017;80:279-86.
  14. Joshi RS, Haddad AF, Lau D, et al. Artificial intelligence for adult spinal deformity. *Neurospine* 2019;16:686-94.
  15. Yagi M, Hosogane N, Fujita N, et al. Predictive model for major complications 2 years after corrective spine surgery for adult spinal deformity. *Eur Spine J* 2019;28:180-7.
  16. Lafage R, Ang B, Alshabab BS, et al. Predictive model for selection of upper treated vertebra using a machine learning approach. *World Neurosurg* 2021;146:e225-32.
  17. Pellisé F, Serra-Burriel M, Smith JS, et al. Development and validation of risk stratification models for adult spinal deformity surgery. *J Neurosurg Spine* 2019 Jun 28:1-13. doi: 10.3171/2019.3.SPINE181452. [Epub].
  18. Durand WM, DePasse JM, Daniels AH. Predictive modeling for blood transfusion after adult spinal deformity surgery: a tree-based machine learning approach. *Spine (Phila Pa 1976)* 2018;43:1058-66.
  19. Ames CP, Smith JS, Gum JL, et al. Utilization of predictive modeling to determine episode of care costs and to accurately identify catastrophic cost nonwarranty outlier patients in adult spinal deformity surgery: a step toward bundled payments and risk sharing. *Spine (Phila Pa 1976)* 2020;45:E252-65.
  20. Peng L, Lan L, Xiu P, et al. Prediction of proximal junctional kyphosis after posterior scoliosis surgery with machine learning in the lenke 5 adolescent idiopathic scoliosis patient. *Front Bioeng Biotechnol* 2020;8:559387.
  21. Jain D, Durand W, Burch S, et al. Machine learning for predictive modeling of 90-day readmission, major medical complication, and discharge to a facility in patients undergoing long segment posterior lumbar spine fusion. *Spine (Phila Pa 1976)* 2020;45:1151-60.