Spine

SPINE Volume 47, Number 9, pp 680–690

DEFORMITY

Improvement and International Validation of the Predictive Probability of the Patient Demographics, Radiographic Index, and Surgical Invasiveness for Mechanical Failure (PRISM) Model for Preventive Procedures in Adult Spinal Deformity Surgery

Mitsuru Yagi, MD, PhD,[a] Naobumi Hosogane, MD, PhD,[b] Christopher P. Ames, MD,[c] Justin S. Smith, MD, PhD,[d] Christopher I. Shaffrey, MD,[e] Frank J. Schwab, MD,[f] Virginie Lafage, PhD,[f] Shay Bess, MD,[g] Satoshi Suzuki, MD, PhD,[a] Nori Satoshi, MD, PhD,[a] Yohei Takahashi, MD, PhD,[a] Osahiko Tsuji, MD, PhD,[a] Narihito Nagoshi, MD, PhD,[a] Masaya Nakamura, MD, PhD,[a] Morio Matsumoto, MD, PhD,[a] and Kota Watanabe, MD, PhD[a], International Spine Study Group (ISSG)

Study Design. This is an international multicenter retrospective review of 219 surgically treated consecutive adult spinal deformity (ASD) patients who had a minimum of five fused segments, completed a 2-year follow-up.

Objective. The purpose of this study was to add the indices of preventive procedures to improve and to validate the predictive probability of the PRISM (patient demographics, radiographic index, and surgical invasiveness for mechanical failure) for mechanical failure (MF) following ASD surgery.

Summary of Background Data. The PRISM was developed from the data of 321 ASD patients, which stratified the risk of MF from six types of risk.

Methods. Data from 136 Japanese ASD patients (age 49 ± 21 yr, 88% female) were used to develop PRISM2, and data from 83 US ASD patients (age 58 ± 12 yr, 86% female) were used for the external validation. We analyzed the associations between three preventive procedures (UIV+1 tethering [TH], teriparatide [TP], and multirod [MR]) and MF by multivariate logistic regression analysis (MRA). The values for the nearest integer of the β of the procedures were added to the six indices of the original PRISM to establish the PRISM2. The discriminative ability of the PRISM/PRISM2 for MF was evaluated using the area under the receiver operating characteristic curve (AUC) and the precision-recall (PR) curve. The Cochran-Armitage test was used to analyze the trend between PRISM/PRISM2 scores and MF.

Results. MF developed in 25% (34 cases). The β values for the preventive procedures calculated by MRA were TH: −2.5, TP: −3.0, and MR: −2.1. The Cochran-Armitage test showed an excellent trend between MF and PRISM/2. The diagnostic ability was superior for the PRISM2 compared with the PRISM (PRISM2; AUC = 0.94 [0.90–0.98], PRISM; AUC = 0.87 [0.81–0.93], difference = −0.07 [−0.11 to −0.03], $P < 0.01$). The AUC of the PRISM2 was 0.70 [0.59–0.81, $P < 0.01$] in the US patient cohort.

Conclusion. We refined the PRISM by adding preventive procedures to the risk indices. Further validation and adjustment in a large different patient cohorts may improve the predictive probability of PRISM2.

Key words: adult spinal deformity, complication, prevention, risk stratification.

Level of Evidence: 3
Spine 2022;47:680-690

From the [a]Department of Orthopaedic Surgery, Keio University School of Medicine, Tokyo, Japan; [b]Department of Orthopaedic Surgery, Kyorin University School of Medicine, Tokyo, Japan; [c]Department of Neurological Surgery, University of California, San Francisco, CA; [d]Department of Neurosurgery, University of Virginia Health System, Charlottesville, VA; [e]Department of Orthopedic Surgery, Duke University, NC; [f]Spine Service, Hospital for Special Surgery, New York, NY; and [g]Department of Orthopedic Surgery, Denver International Spine Center, Denver, CO.

Acknowledgment date: June 4, 2021. First revision date: October 28, 2021. Acceptance date: November 1, 2021.

The manuscript submitted does not contain information about medical device(s)/drug(s).

No funds were received in support of this work.

Relevant financial activities outside the submitted work: board membership, consultancy, grants, stocks, royalties.

Address correspondence and reprint requests to Kota Watanabe, MD, PhD, 35 Shinanomachi, Shinjyuku, Tokyo, Japan; E-mail: kw197251@keio.jp

DOI: 10.1097/BRS.0000000000004295

Mechanical failure (MF) following adult spinal deformity (ASD) surgery is a serious complication and often requires revision surgery.[1–11] The forms of MF include proximal junctional kyphosis/failure (PJK/PJF), distal junctional failure (DJK), rod fracture (RF), and other implant-related complications.[1–11] Previous studies have shown that these MFs develop overlappingly in many cases and that the risks of these MFs are identical or quite similar.[1,5,6,9,10] Several demographic, radiographic, and surgical potential risk factors for MF have been identified, including age, frailty, poor bone quality, large sagittal deformity, level of lower-most instrumented vertebra (LIV), and the application of pedicle subtraction osteotomy (PSO).[1–11] We have previously described six independent risk factors for developing MF (age, body mass index [BMI], bone mineral density [BMD], frailty, LIV, and pelvic tilt) from the data of 321 Japanese ASD patients.[12] We then designed a simplified, risk-stratification algorithm (PRISM; Patient demographics, Radiographic Index and Surgical invasiveness for Mechanical failure) to provide a score that could be used to stratify the risk of MF.[12] We included the abovementioned six variables in our risk-stratification model from the sum of the standardized regression coefficients (b) of each risk factor.[12] The discriminative ability of the PRISM algorithm for MF using the area under the receiver operating characteristic curve (AUC) showed excellent probability with an AUC of 0.86 (95% CI 0.77–0.95) in the external validation cohort. Recently, several procedures, including prophylactic teriparatide administration, UIV+1 or +1 and +2 tethering, cement augmentation of the UIV+1 vertebra, and multirod construction, have been described as preventive strategies for postoperative MF.[13–18] Safaee *et al*[13] showed that UIV+1 or +1 and +2 tethering significantly reduced the risk of PJF in ASD surgery. Yagi *et al*[14] reported that prophylactic administration of teriparatide significantly reduced the risk of PJF in ASD surgery. Merrill *et al*[15] revealed that a multirod construct significantly reduced the risk of RF in ASD surgery. Several risk stratification models for MF following ASD surgery have been described, including the global alignment and proportion (GAP) score and PRISM.[12,19] The incidence of postoperative MF in ASD surgery is quite high; therefore, preventive procedures should be used diligently in high-risk patients. However, given the costs and resources involved in preventive procedures, these promising preventive procedures should not necessarily be applied to all ASD surgeries. To date, there are no predictive models that include preventive procedures in the risk index for MF in ASD surgery. Our risk stratification model, PRISM, has also been developed based on the results of ASD patients who did not have any preventive procedures before or after surgery. Additionally, although the excellent predictive probability for the PRISM model has been described in Japanese ASD patients, the predictive probability has not been validated for different races or patients who live in a different geographical area.[12] The development of complications following spine surgery may be influenced by differences in race, lifestyle, and genetic background. Thus, the purpose of this study was to improve the predictive probability of PRISM for MF following ASD surgery by adding indices of the previously described preventive procedures and to validate the predictive probability in ASD patients from the United States (US).

HYPOTHESIS

Adding the indices of preventive procedures for MF to the existing indices of the risk stratification score (PRISM) may improve the predictive probability of PRISM in ASD surgery.

STUDY DESIGN

We conducted a retrospective analysis of data entered prospectively into a single center database to develop PRISM2 and to test the predictive probability of PRISM2 from data entered prospectively into the United States multicentered database.

MATERIALS AND METHODS

This study was approved by the institutional review board at our institutions, and all subjects consented and agreed with their inclusion. We attest that the oral and written informed consents were obtained from all included patients. All methods were performed in accordance with the relevant guidelines and regulations.

Patient Enrollment

We retrospectively reviewed charts and radiographs for 136 consecutive patients who underwent surgery for ASD by six board-certified spine surgeons in our institution between 2016 and 2018 and reached the 2-year postoperative followup point. For the external validation, we retrospectively reviewed data from 83 patients who underwent surgery for ASD by six board-certified spine surgeons in six institutions between 2016 and 2018 in a US multicenter database and reached the 2-year postoperative follow-up point.

Inclusion and Exclusion Criteria

Subjects were at least 20 years old at the index surgery and had a spinal deformity defined by a Cobb angle more than or equal to 20°, a C7 sagittal vertical axis (C7SVA) more than or equal to 5 cm, or pelvic tilt (PT) more than or equal to 25°. We included patients with at least five fused vertebrae, instrumentation, and fusion from the upper-most instrumented vertebra (UIV) to the LIV, and complete 2-year follow-up data. Patients were excluded if they lacked appropriate radiographs or had syndromic, neuromuscular, or other pathological conditions.

Data Collection and Radiographic Data Assessment

We collected demographic and clinical data for each patient, including age, sex, comorbidities, BMI, BMD, frailty, and history of spine surgery. Frailty was assessed using the modified frailty index (mFI).[20,21] We collected the following surgical data: the Schwab-SRS ASD classification, application of three-column osteotomies, UIV and LIV levels, and the number of fused vertebrae. Radiographic data were obtained at baseline and at the 6-week and 2-year postoperative followup. Each type of preventive procedure for MF was performed for patients who were considered to be at high risk of developing MF after surgery based on the findings of the previous literature.

Of 141 candidates, 136 subjects (96.4%) had complete demographic and radiographic data that sufficiently documented any instances of postoperative MF and were thus included in the study cohort. The remaining five candidates were lost during follow-up, including one candidate who died for reasons unrelated to the surgery and was excluded from the cohort. These five patients did not have MF prior to being lost to follow-up (5–10 months).

Inclusion of Mechanical Failures

Based on previous literature, we included all MFs found on the radiographs that developed within 2 years of the surgery (PJK/PJF, DJK, RF, and other implant-related complications).[15]

Data Preparation and Analysis of the Associations Between Three Preventive Procedures and Mechanical Failures Following ASD Surgery

Subjects were categorized into two groups: those who had any MF within 2 years of the surgery and those who did not have any MF. We investigated the associations between three preventive procedures (prophylactic teriparatide administration [TP], UIV+1 laminar tethering [Th], and multirod construct [MR]) and MF by a multivariate binary logistic regression model to evaluate the adjusted associations of each potential explanatory variable and to predict the likelihood of developing MF. TP was used for those who had pre-existing low BMD (T-score of femoral neck lower than −1.5).[14]

Building a Model to Predict Mechanical Failure

We first established values for each preventive procedure by rounding the β regression coefficient obtained in the regression analysis to the nearest whole integer (0–3). Next, the values for all preventive procedures were summed along with the six indices from the original PRISM score (age, BMI, BMD, frailty, baseline PT, and level of LIV [pelvis]) to establish the PRISM2 score.[12] The weights of the six indices identified in the original study were retained and used to build the PRISM2 algorithm. The PRISM2 score was determined as the sum of the values of the risk variables (Figure 1). If the total score was less than 0, a score of 0 was recorded. The PRISM2 score was used to stratify patient risk into low risk (total score 0–1), moderate risk (total score 2–4), high risk (total score 5–8), and very high risk (total score 9–12).

Internal Validation of PRISM and PRISM2 Algorithms for MF and Unplanned Surgery Related to MF

The Cochran-Armitage test for trends was performed to analyze whether there was a trend between the PRISM score, PRISM2 score, and the actual incidence of MF.[22] The discriminative ability of the PRISM and PRISM2 models for MF and unplanned surgery related to MF were evaluated using the AUC and PR curve with overall model quality.[23]

External Validation of PRISM2 Algorithms for MF

The Cochran–Armitage test for trends was performed to analyze whether there was a trend between the PRISM2 score and the actual incidence of MF. The discriminative ability of the PRISM2 models for MF was evaluated using the AUC. To adjust for the difference of BMI between the Japanese and US patients, the risk indicator of PRISM algorithm "BMI" was adjusted as follows: score 0 = BMI $<25\,g/m^2$, score 1 = BMI 25 to $30\,kg/m^2$, score 2 = BMI $>30\,kg/m^2$.

Statistical Analysis

Differences between the MF and MF-free groups were compared using an unpaired *t* test, a chi-square test, Tukey HSD test, and Fisher exact test where appropriate. A *P* value <0.05 with a CI of 95% was considered statistically significant.[24,25] For the discriminative ability of PRISM and PRISM2 models, an AUC of 0.5 was considered no discrimination, an AUC between 0.5 and 0.7 was considered poor discrimination, an AUC between 0.7 and 0.8 was considered acceptable, an AUC between 0.8 and 0.9 was considered moderate, and an AUC greater than 0.9 was considered excellent.[23] All analyses were performed using the Statistical Package for the Social Sciences (SPSS statistics version 27.0, SPSS modeler version 18, IBM Corp., Armonk, NY) and XLSTAT (version 2020.4, Addinsoft Inc., Paris).

RESULTS

Patient Characteristics

Patient demographic data and surgical descriptions are shown in Table 1. Among 136 patients, there were 39 MFs in 34 patients (25%) within 2 years of surgery. The most common form of MF was RF (n = 14,10.3%), and the second most common form of MF was PJF (n = 12, 8.8%). Among the patients who developed MF, five patients (15.1%) experienced more than two MFs, and 22 patients (64.7% of the patients who experienced MF and 16.2% of the overall patients) required unplanned additional surgeries to treat the MF. Among the patient cohort, 27 patients (19.9%) received prophylactic teriparatide administration for more than 3 months, 21 patients (15.4%) had UIV+1 laminar tethering, and four patients (2.9%) had a multirod construct. Twenty patients (14.7%) had two of the above-described preventive procedures, and two patients (1.5%) had all three procedures.

Refinement of the PRISM Model by Adding to the Value of Predictive Procedures

The result of univariate risk analysis for developing MF was seen in supplemental Table 1, http://links.lww.com/BRS/

	Score	Grade
Age		Total score: 0 - 1 Low risk
< 60yrs	0	
60yrs	1	
BMD (T-score)		
-1.5	0	Total score: 2 - 4 Moderate
< -1.5	2	
BMI		
< 18.5kg/m^2	0	
18.5 - 25kg/m^2	1	
> 25kg/m^2	2	Total score: 5 -8 high
Frailty		
Robust	0	
Prefrail	1	
Frail	2	Total score: 9 - 12 Very high
LIV		
Above L5	0	
Pelvis	2	
PT		
< 20°	0	
20 - 30°	2	
> 30°	3	
Teriparatide	-2	
Tethering	-3	
Multirod	-2	
If the total score s < 0, consider the score to be 0"		

Figure 1. The PRISM2 model: risk-grading system for mechanical failure following ASD surgery. The risk stratification score was used to stratify the risk into low-risk (risk score 0–1), moderate-risk (risk score 2–4), high-risk (risk score 5–8), and very high-risk (risk score 9–12) categories. ASD indicates adult spinal deformity.

B834. We first established values for each preventive procedure (Th, TP, and MR) by rounding the β regression coefficient obtained in the multivariate analysis to the nearest whole integer (0–3). The standardized regression coefficients (b) for Th, TP, and MR calculated by MRA were Th: −2.48, TP: −2.98, and MR: −2.06 (Table 2). The values for preventive procedures (Th: −2, TP: −3, and MR: −2) were summed along with the six indices of the original PRISM score to establish the PRISM2 score (0–12). In the surgical risk stratification classification using PRISM2, 13 (9.6%) patients were classified as very high-risk, 34 patients (25%) were classified as high risk, 30 patients (22.1%) were classified as moderate risk, and 59 patients (43.4%) were classified as low risk.

This proportion was significantly different when compared with the classification using the original PRISM score (29 patients [21.3%] were classified as having as very highrisk grade, 30 patients [22.1%] were classified as high risk, 22 patients [16.2%] were classified as moderate risk, and 55 patients [40.4%] were classified as low risk).

Comparisons of the Predictive Probability of the PRISM and PRISM2 Models for MF and Unplanned Surgery Related to MF

The Cochran-Armitage test showed an excellent trend with the incidence of MF and both the PRISM and PRISM2 scores ($P < 0.001$, Figure 2A, B). The discriminative ability of PRISM and PRISM2 for MF and for unplanned surgery related to MF were evaluated using the AUC and PR curve. The internal validation with the patient cohort for PRISM and PRISM2 showed good model fit for the prediction of MF with moderate accuracy for PRISM and excellent accuracy for PRISM2 (PRISM; AUC = 0.870 ± 0.030 [95% CI 0.812–0.929]; $P < 0.001$], PRISM2; AUC = 0.940 ± 0.030

TABLE 1. Demographic, Radiographic, and Surgical Description of the Patient Cohort

Variables	Total
No. of pts	136
Age, yrs	49 ± 21 [21, 80]
Sex (female)	120 (88%)
BMI, kg/m^2	20.7 ± 3.5 [11.4, 30.3]
BMD (T score)	−1.1 ± 0.9
Frailty (mFI)	0.7 ± 0.1 [0, 0.64]
Schwab-SRS classification (T: D: L: N)	38 (28%): 41 (30%): 35 (26%): 22 (16%)
LIV (pelvis)	48 (36%)
Level fused	10.1 ± 2.9 [5, 18]
3CO (PSO/VCR)	7 (5%)
Revision case	7 (5%)
Baseline sagittal spinal alignment	
C7SVA, mm	55.9 ± 48.1
PT, °	223 ± 13.8
PI-LL, °	22.1 ± 21.5

Mean and standard deviations. Range in brackets. Percentage in parentheses.
BMD indicates body mass index; BMI, bone mineral density; LIV, lower-most instrumented vertebra; PT, pelvic tilt.

[95% CI 0.899–0.979]; $P < 0.001$], Figure 3A, B). Moderate and excellent discriminating ability was also confirmed for unplanned surgeries related to MF (PRISM; AUC = 0.840 ± 0.037 [95% CI 0.767–0.914]; $P < 0.001$, PRISM2; AUC = 0.914 ± 0.030 [95% CI 0.856–0.973]; P < 0.001, Figure 3C, D). The comparisons of the predictive probability of the PRISM and PRISM2 models revealed statistically superior discriminative ability for PRISM2 for both MF development and for unplanned surgery related to MF (MF; difference in AUC: PRISM-PRISM2 = −0.069 [95% CI −0.106 to −0.032]; $P < 0.001$, unplanned surgery; difference in AUC: PRISM-PRISM2 = −0.074 [95% CI - 0.114 to −0.034]; $P < 0.001$). PR curve analyses also showed a good relationship between precision (positive predictive value) and recall (sensitivity) for the possible cutoff of the PRISM/PRISM2 score for MF and for unplanned surgery related to MF (overall model quality; PRISM: MF = 0.81, unplanned surgery = 0.77, PRISM2: MF = 0.90, unplanned surgery = 0.86, Figure 3A–D).

The Predictive Probability of the PRISM2 Models for MF in the US Patient Cohort

Among 83 US ASD patients, MF developed in 49% (41 cases, Table 3). The most common MF was PJK (n = 32, 39%). The result of univariate risk analysis for developing MF was seen in supplemental Table 2, http://links.lww.com/BRS/B834. The Cochran-Armitage test showed an excellent trend between MF and PRISM2 scores ($P < 0.01$, Figure 4). The AUC of the PRISM2 model for MF was 0.701 (0.585–0.808, $P = 0.034$), which was moderately accurate diagnostic ability.

DISCUSSION

Mechanical failure following ASD surgery is a serious concern because this complication can lead to not only devastating

TABLE 2. Multivariate Logistic Regression Analysis for Preventive Procedures for Mechanical Failure following ASD Surgery

	Variables	Regression Coefficient	*P* Value	OR
Preventive procedures	Tethering	−2.48 (1.02)	0.02*	0.08 [0.01, 0.62]
	TP	−2.98 (1.05)	<0.01*	0.05 [0.01, 0.40]
	Multirod	−2.06 (1.74)	0.24	0.13 [0.00. 3.82]
Original PRISM risk indices	Age	1.21 (0.88)	0.17	3.36 [0.60, 18.86]
	BMD	0.71 (0.41)	0.09	1.98 [0.90, 4.56]
	BMI	−0.40 (0.63)	0.53	0.68 [0.20, 2.31]
	Frailty	1.86 (0.69)	0.01*	6.46 [1.66, 24.90]
	LIV	1.50 (0.54)	0.01*	4.42 [1.56, 12.84]
	PT	0.21 (0.45)	0.64	1.18 [0.51, 2.96]

Standard error in parentheses. 95% confidence interval in brackets.
ASD indicates adult spinal deformity; BMD, body mass index; BMI, bone mineral density; LIV, lower-most instrumented vertebra; OR, odds ratio; PT, pelvic tilt.
*Statistically significant.

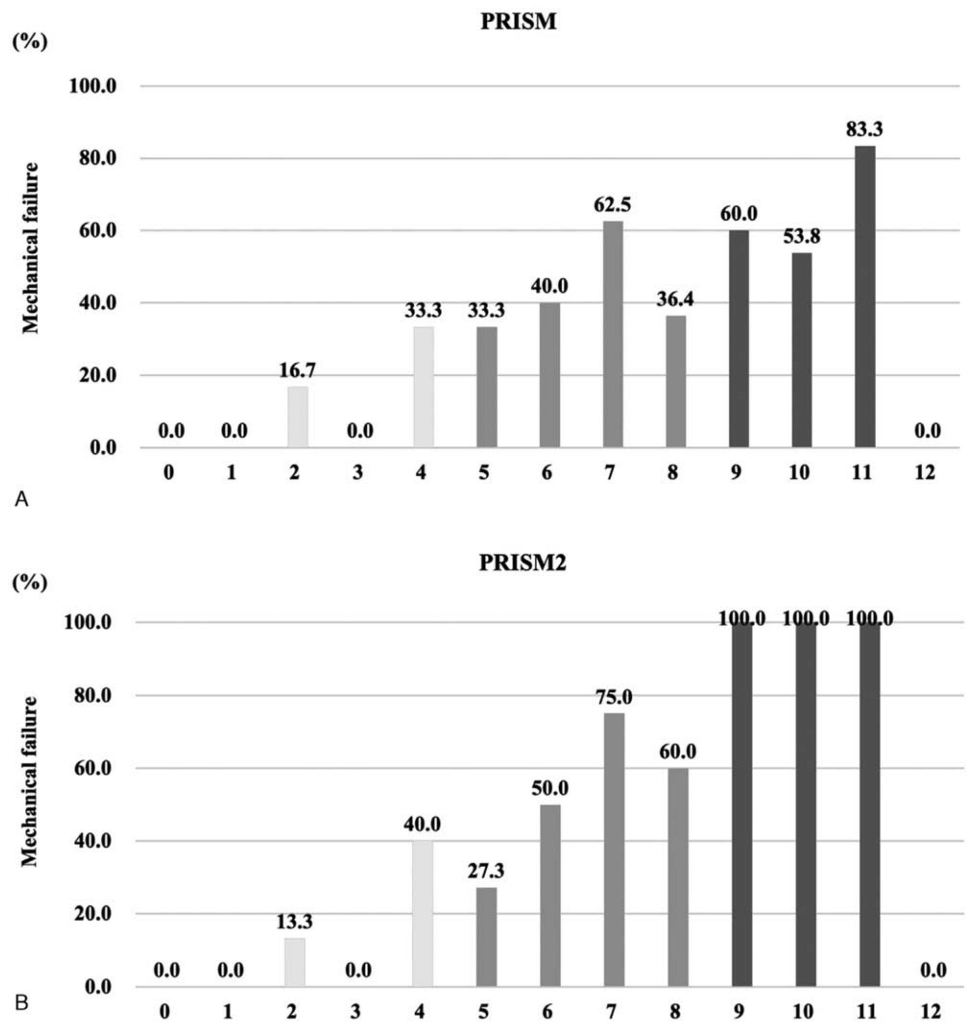

Figure 2. The distribution of mechanical failure in the patient cohort, stratified by the PRISM or PRISM2 score relative to the observed mechanical failure rate. **A**, The mechanical failure rate increased with the PRISM score. A statistically significant trend between the mechanical failure rate and the score was observed (P for trend <0.001, Cochran-Armitage test for trends). **B**, The mechanical failure rate increased with the PRISM2 score. A statistically significant trend between the mechanical failure rate and the score was observed *(P* for trend <0.001, Cochran-Armitage test for trends). PRISM indicates Patient Demographics, Radiographic Index, and Surgical Invasiveness for Mechanical Failure.

adverse events for the patient but also a significant increase in medical costs.[1–11,24–28] The serious adverse events that MF can cause include intolerable pain, paralysis, skin irritation, and deterioration of the quality of life associated with spinal malalignment.[1–11] The reported incidence of unplanned additional surgery for the treatment of MF ranges from 7% to 28%.[1–11] Crawford *et al*[1] described that 24% of patients required unplanned surgery following ASD surgery, and rod fracture occurred in 15% of patients and PJF occurred in 5.4% of patients. Even though previous literature suggests that revision surgery for MF is similarly effective when compared with patients who had primary surgery, the considerably high revision rate related to MF significantly increased the medical cost and decreased the cost-effectiveness of the surgical treatment of ASD.[24–28] Theologis *et al*[24] recently described that the cost for revision surgery for PJF, one of the most common forms of MF, reached $55,547 per case and increased the 2-year total cost by 12%. Raman *et al*[25] also described that the 2-year total medical cost for revision surgery for the treatment of MF was $115,509.

Recently, several prevention strategies for MF have been developed and used. Among them, tethering, teriparatide, and multirod constructs have been reported to predominantly decrease the risk of MF in ASD surgery. Safaee *et al*[13] described that tethering significantly decreased the development of PJF (OR 0.2). Guevara-Villazon *et al*[18] described that a multirod construct significantly reduced the incidence of RF and the number of revision surgeries related to MF in their matched ASD surgery cohort. Yagi *et al*[14] reported that prophylactic teriparatide administration improved the volumetric BMD and fine bone structure at UIV+1 and reduced the incidence of PJF in elderly ASD patients. Adding these preventive procedures actually decreased the development of MF in our patient cohort. The incidence of MF was 27.3% in patients with preventive procedures and 94.4% in patients without preventive procedures in a group of patients with an original PRISM grade greater than 8, which is considered a very high risk for MF development. The revision frequency also decreased from 66.7% to 13.6% among the patients with an original PRISM grade greater than eight when preventive procedures were applied. Therefore, we tried to refine the PRISM model by adding the indices of those preventive procedures. The Cochran-Armitage test showed an excellent trend with the incidence of MF and PRISM2 score, and the AUC and PR

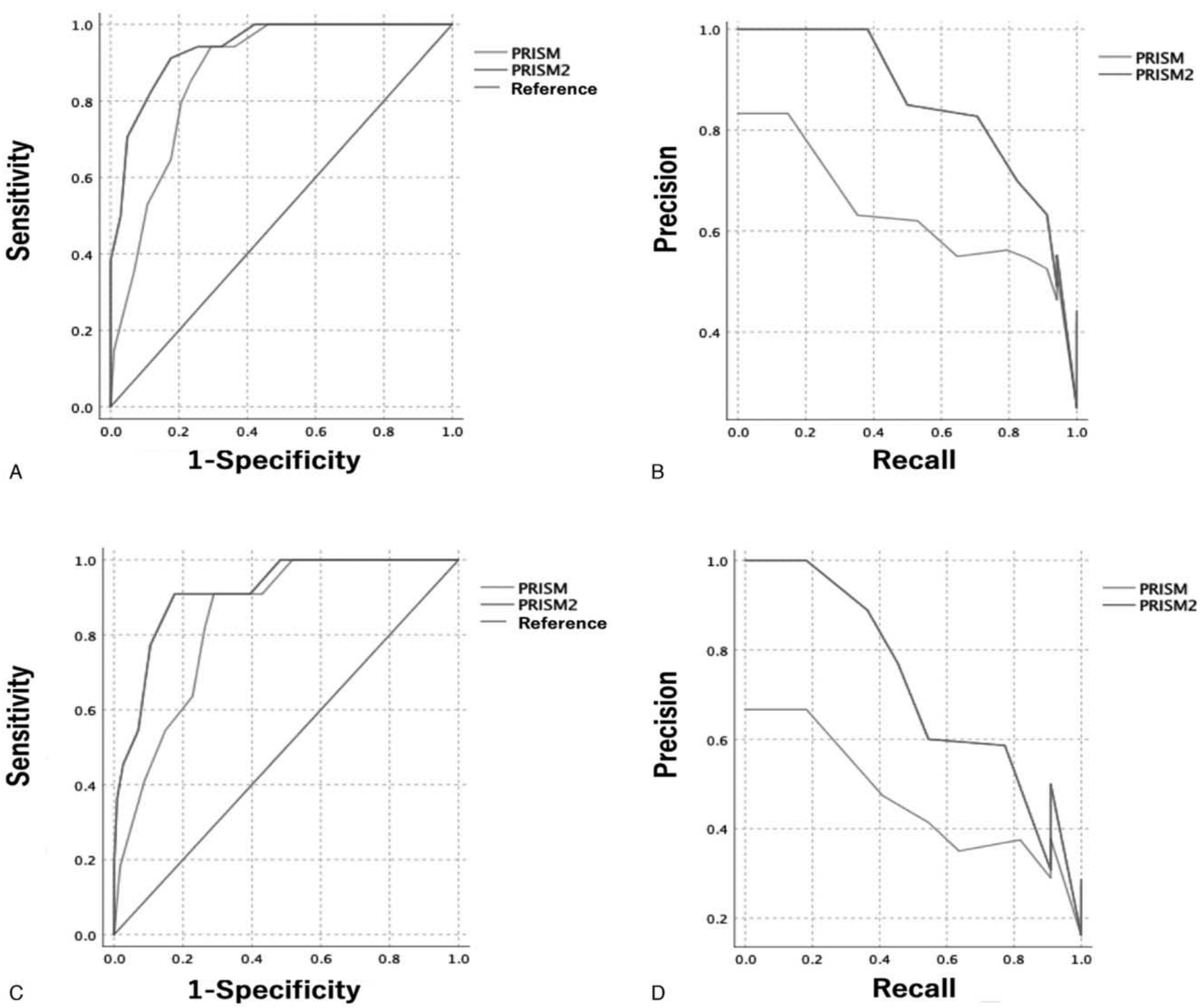

Figure 3. The distribution of the score, receiver operating characteristic (ROC) analysis and precision-recall (PR) analysis in the patient cohort relative to the observed mechanical failure rate and unplanned surgery related to MF for the PRISM/PRISM2 scores. **A**, ROC curve of the mechanical failure rate for the PRISM score (blue line) and PRISM2 score (red line) in the patient cohort. The area under the ROC curve (AUC) was 0.870, standard error = 0.030, $P<0.001$, 95% CI=.0.812–0.929 for the PRISM score. The AUC was 0.939, standard error = 0.020, $P<0.001$, 95% CI = 0.899–0.979 for the PRISM2 score. The difference in AUC was PRISM-PRISM2 = –0.069 [95% CI –0.106 to –0.032]; $P<0.001$, Z score = –3.619. **B**, PR curve of the mechanical failure rate for the PRISM score (blue line) and PRISM2 score (red line) in the patient cohort. The supplemental PR curve analyses also showed a good relationship between precision and recall for the possible cutoff PRISM/PRISM2 score for MF (overall model quality; PRISM = 0.81, PRISM2 = 0.90). **C**, ROC curve of the unplanned surgery rate for the PRISM score (blue line) and PRISM2 score (red line) in the patient cohort. The area under the ROC curve (AUC) was 0.840, standard error = 0.037, $P<0.001$, 95% CI = 0.767–0.914 for the PRISM score. The AUC was 0.914, standard error = 0.030, $P<0.001$, 95% CI = 0.856–0.973 for the PRISM2 score. The difference in AUC was PRISM-PRISM2 = –0.074 [95% CI –0.114 to –0.034]; $P<0.001$, Z score = –3.629. **D**, PR curve of the unplanned surgery rate for the PRISM score (blue line) and PRISM2 score (red line) in the patient cohort. The PR curve analyses also showed a good relationship between precision and recall for the possible cutoff of the PRISM/PRISM2 score for MF (overall model quality; PRISM: = 0.77, PRISM2: MF = 0.86).

analyses both indicated that newly the established PRISM2 model had good discriminative ability for MF and was better than that of the original PRISM model for our patient cohort (AUC: PRISM = 0.87, PRISM2 = 0.94, difference in AUC = –0.069, overall model quality; PRISM = 0.81, PRISM2 = 0.90). Moreover, the good discriminative ability of the PRISM2 model for unplanned surgery indicated that PRISM2 is useful not only to stratify the risk of MF but also to predict the risk of revision surgery needed after ASD surgery (AUC: PRISM = 0.84, PRISM2 = 0.91, difference in AUC = –0.074, overall model quality; PRISM = 0.77, PRISM2 = 0.90, Figure 5A–D).

In the US patient cohort, the Cochrane-Armitage test revealed statistically significant trend between the advance of PRISM2 score and the actual development of MF. Furthermore, the AUC analysis revealed moderate diagnostic ability of PRISM2 for the development of MF. The population of patients with ASD is a heterogeneous population in

TABLE 3. Demographic, Radiographic, and Surgical Description of the US Patient Cohort

Variables	Total
No. of pts	83
Age, yrs	59 ± 12 [24, 80]
Sex (female)	71 (86%)
BMI, kg/m^2	27.6 ± 5.4 [19.6, 41.9]
BMD (T score)	−1.0 ± .1.0
Frailty (mFI)	.6 ± .7 [0, .3]
Schwab-SRS classification (T: D: L: N)	5 (6%): 27 (33%): 29 (35%): 22 (27%)
LIV (pelvis)	56 (67%)
Level fused	11.4 ± 3.9 [5, 22]
3CO (PSO/VCR)	18 (22%)
Revision case	27 (33%)
Baseline sagittal spinal alignment	
C7SVA, mm	44.2 ± 61.5
PT, °	22.4 ± 10.3
PI-LL, °	11.4 ± 20.2

Mean and standard deviations. Range in brackets. Percentage in parentheses. BMD indicates body mass index; BMI, bone mineral density; LIV, lower-most instrumented vertebra; PT, pelvic tilt.

Figure 4. Representative case of surgically treated ASD patients stratified by PRISM and PRISM2 scores. The patient is a 62-year-old female with a Schwab-SRS type L curve and frailty; mFI 0.27 (prefrail), BMD; T-score −1.9, BMI; 22.9 kg/m^2, LIV of the pelvis. **A**, Preoperative posteroanterior (PA) view of a standing whole spine radiograph. The Cobb angle of the lumbar curve was 46°, the baseline SRS22 pain score was 2.2, the function score was 2.4, the self-image score was 2.8, and the mental health score was 3.2. **B**, Preoperative lateral view of a standing whole spine radiograph. The C7SVA was +11 cm, PT was 36°, and PI-LL was 62°. The PRISM score = age (1) + BMD (2) + BMI (1) + frailty (1) + LIV (2) + PT (3) = 10 (very high risk). The patient had a PSF T9-pelvis with L1/2 posterior column osteotomy, L2/3-L3/4 lateral lumbar interbody fusion, L5/S PLIF and T8 laminar tethering with a multirod construct. The patient also received prothoracic administration of teriparatide for 6months. The PRISM2 score = age (1) + BMD (2) + BMI (1) +frailty (1) + LIV (2) + PT (3) + teriparatide (−2) + tethering (−3) + multirod (2) = 3 (moderate risk). **C**, Two-years postoperative PA view of a standing whole spine radiograph. **D**, Two-year postoperative lateral view of a standing whole spine radiograph. Radiographs showed no MF at 2 years after surgery. The 2-year postoperative SRS22 pain score was 4.0, the function score was 3.8, the self-image score was 4.4, and the mental health score was 3.8. ASD indicates adult spinal deformity; BMD, body mass index; BMI, bone mineral density; LIV, lower-most instrumented vertebra.

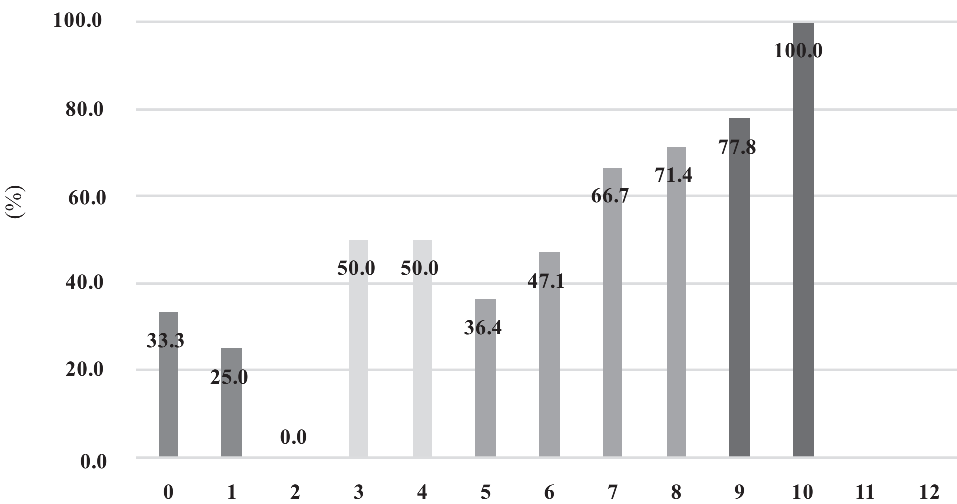

Figure 5. The distribution of mechanical failure in the US patient cohort, stratified by the mPRISM2 score relative to the observed mechanical failure rate. The mechanical failure rate increased with the PRISM2 score in the US patient cohort. A statistically significant trend between the mechanical failure rate and the score was observed (*P* for trend <0.001, Cochran-Armitage test for trends).

age, curve type, ADL, and associated comorbidities. Therefore, the clinical outcomes, deformity type, and incidence of surgical complications of ASD can differ based on lifestyle, genetics, and geographical area.

Yilgor *et al*[19] developed and described the GAP score which showed excellent diagnostic ability for developing MF following ASD surgery. Despite the excellent predictive power of the GAP score demonstrated in the original article, the predictive probability of the GAP score in different patient cohorts remains unclear. Bari *et al*[29] reported no significant association between the GAP score and MF or revision surgery following 149 consecutive ASD surgeries. Kwan *et al*[29] found that alignment targets based on the GAP score were not associated with increased risks of MF and MF related revision surgery in ASD surgery. Yagi *et al*[30] also described no correlation between the GAP score and actual development of MF following ASD surgery in Japanese ASD patients. In the present study, although the predictive probability is lower than that of the Japanese patient cohort, PRISM2 showed moderate diagnostic ability in the US ASD patients who had difference lifestyle and genetic background. Adjustment of risk variables using a larger sample cohort may further improve the predictive ability of PRISM in patients with different backgrounds.

The present study has several limitations. First, we did not incorporate all the preventive procedures in the present analysis. Terminal rod contour, administration of bisphosphonate, type of orthosis, and cement augmentation of UIV+1 vertebra may also prevent the development of MF.[17,21,31,32] The vast majority of our patient cohort used Jewett braces and did not have bisphosphonates before or after surgery. Additionally, in most of the procedures, the proximal tips of the rod were slightly over-bent (increased kyphosis) intraoperatively to decrease the chance of PJK, but we did not set a particular angle for the terminal contour of the rods. Similarly, the duration of the preoperative prophylactic administration of TP was different among the patients, ranging from 1 month to 6 months, whereas all the patients received prophylactic TP administration at least during the 3-month postoperative period, when the risk of PJF was considered high. Thus, as mentioned above, not all the intraoperative techniques and preoperative procedures could be included in the risk variables in the present study. These factors potentially affect the incidence of MF. However, the present study included patients treated by six different surgeons; therefore, we believe that the findings of the present study can be generally extrapolated to different patient populations, but further validation and adjustment are necessary.

Second, we included MF up to 2 years after surgery but did not include later MF. Previous studies have described that RF can develop more than 2 years after surgery.[2,5,11] Therefore, the predictive probability of PRISM2 may decrease when including all MFs after surgery. However, previous studies have indicated that not all RFs require revision surgery, especially those that develop in a late stage and can be managed conservatively. Thus, to focus more on symptomatic mechanical failure requiring revision surgery, we established the PRISM2 model using the data from MFs within 2 years of the index surgery.

Third, the relatively small sample size in the US patient cohort. In the US patient cohort, 186 patients were eligible, and 152 patients reached 2-year follow-up while 69 patients lacked the DXA score data and therefore, 83 patients were included. This relatively small sample size in the validation cohort is mainly due to the retrospective nature of the study. In Japan, DXA examination is the standard preoperative patient assessment, while this might not be the standard preoperative patient assessment in the United States. The small sample in the validation cohort most likely causes restricted clinical variation and truncation. Additionally, due to the limited sample size, the 95% CI of the PRISM2 risk variables in the US validation cohort was relatively large. However, the patient background and complication rate in the included patients were not different from the overall eligible patients and therefore, we believed that the included samples represented the overall patient cohort. Further validation and adjustment in a large different patient cohorts may improve the predictive probability of PRISM2.

CONCLUSIONS

We refined the PRISM model for predicting MF in ASD surgery by adding MF preventive procedures to the risk indices. The predictive probability and discriminative ability of the newly established PRISM2 model for the development of MF and for unplanned surgery related to MF were acceptable. Further validation and adjustment in a large different patient cohorts may improve the predictive probability of PRISM2.

➢ Key Points

- ❑ A risk stratification model (PRISM) for MF following ASD surgery was refined based on data from 136 surgically treated ASD patients.
- ❑ Standardized regression coefficients (b) for Th, TP, and MR were calculated by MRA and were added to the PRISM score to establish PRISM2 (*B;* Th: - 2.48, TP: −2.98, MR: −2.06).
- ❑ Diagnostic ability was superior for PRISM2 compared with PRISM (PRISM2; AUC = 0.94 [0.90–0.98], PRISM; AUC = 0.87 [0.81–0.93], difference = −0.07 [−0.11 to −0.03], $P < 0.01$).
- ❑ The AUC of the PRISM2 was 0.70 [0.59–0.81, $P < 0.01$] in the US patient cohort.
- ❑ Using PRISM2 measurements with their ability to accurately predict the risk of MF, surgeons can predict surgical risk for MF before surgery and optimize the procedure.

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's Web site (www.spinejournal.com).

References

1. Crawford CH 3rd, Glassman SD, Carreon LY, et al. Prevalence and indications for unplanned reoperations following index surgery in the adult symptomatic lumbar scoliosis NIH-sponsored clinical trial. *Spine Deform* 2018;6:741–4.
2. Lertudomphonwanit T, Kelly MP, Bridwell KH, et al. Rod fracture in adult spinal deformity surgery fused to the sacrum: prevalence, risk factors, and impact on health-related quality of life in 526 patients. *Spine J* 2018;18:1612–24.
3. Smith JS, Shaffrey E, Klineberg E, et al. Prospective multicenter assessment of risk factors for rod fracture following surgery for adult spinal deformity. *J Neurosurg Spine* 2014;21:994–1003.
4. Yasuda T, Hasegawa T, Yamato Y, et al. Lumbosacral junctional failures after long spinal fusion for adult spinal deformity-which vertebra is the preferred distal instrumented vertebra?. *Spine Deform* 2016;4:378–84.
5. Yagi M, Akilah KB, Boachie-Adjei O. Incidence, risk factors and classification of proximal junctional kyphosis: surgical outcomes review of adult idiopathic scoliosis. *Spine (Phila Pa 1976)* 2011;36:E60–8.
6. Yagi M, Fujita N, Tsuji O, et al. Low bone-mineral density is a significant risk for proximal junctional failure after surgical correction of adult spinal deformity: a propensity score-matched analysis. *Spine (Phila Pa 1976)* 2018;43:485–91.
7. Yagi M, Fujita N, Okada E, et al. Fine-tuning the predictive model for proximal junctional failure in surgically treated patients with adult spinal deformity. *Spine (Phila Pa 1976)* 2018;43:767–73.
8. Yagi M, Rahm M, Gaines R, et al. Characterization and surgical outcomes of proximal junctional failure in surgically treated patients with adult spinal deformity. *Spine (Phila Pa 1976)* 2014;39:E607–14.
9. Hallager DW, Karstensen S, Bukhari N, et al. Radiographic predictors for mechanical failure after adult spinal deformity surgery: a retrospective cohort study in 138 patients. *Spine (Phila Pa 1976)* 2017;42:E855–63.
10. Yagi M, Hosogane N, Fujita N, et al. Surgical risk stratification based on preoperative risk factors in adult spinal deformity. *Spine J* 2019;19:816–26.
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. Yagi M, Hosogane N, Fujita N, et al. The patient demographics, radiographic index and surgical invasiveness for mechanical failure (PRISM) model established for adult spinal deformity surgery. *Sci Rep* 2020;10:9341.
13. Safaee MM, Deviren V, Dalle Ore C, et al. Ligament augmentation for prevention of proximal junctional kyphosis and proximal junctional failure in adult spinal deformity. *J Neurosurg Spine* 2018;28:512–9.
14. Yagi M, Ohne H, Konomi T, et al. Teriparatide improves volumetric bone mineral density and fine bone structure in the UIVR1 vertebra, and reduces bone failure type PJK after surgery for adult spinal deformity. *Osteoporos Int* 2016;27:3495–502.
15. Merrill RK, Kim JS, Leven DM, et al. Multi-rod constructs can prevent rod breakage and pseudarthrosis at the lumbosacral junction in adult spinal deformity. *Global Spine J* 2017;7:514–20.
16. Viswanathan VK, Kukreja S, Minnema AJ, et al. Prospective assessment of the safety and early outcomes of sublaminar band placement for the prevention of proximal junctional kyphosis. *J Neurosurg Spine* 2018;28:520–31.
17. Ghobrial GM, Eichberg DG, Kolcun JPG, et al. Prophylactic vertebral cement augmentation at the uppermost instrumented vertebra and rostral adjacent vertebra for the prevention of proximal junctional kyphosis and failure following long-segment fusion for adult spinal deformity. *Spine J* 2017;17:1499–505.
18. Guevara-Villazon F, Boissiere L, Hayashi K, et al. Multiple-rod constructs in adult spinal deformity surgery for pelvic-fixated long instrumentations: an integral matched cohort analysis. *Eur Spine J* 2020;29:886–95.
19. 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.
20. Karam J, Tsiouris A, Shepard A, et al. Simplified frailty index to predict adverse outcomes and mortality in vascular surgery patients. *Ann Vasc Surg* 2013;27:904–8.
21. Yagi M, Fujita N, Okada E, et al. Impact of frailty and comorbidities on surgical outcomes and complications in adult spinal disorders. *Spine (Phila Pa 1976)* 2018;43:1259–67.
22. Cochran-Armitage J. A Wilcoxon-type test for trend. *Stat Med* 1985;4:87–9.
23. Powers D. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *J Mach Learn Technol* 2011;2:37–63.
24. Theologis AA, Miller L, Callahan M, et al. Economic impact of revision surgery for proximal junctional failure after adult spinal deformity surgery: a cost analysis of 57 operations in a 10-year

experience at a major deformity center. *Spine (Phila Pa 1976)* 2016;41:E964–72.
25. Raman T, Nayar SK, Liu S, et al. Cost-effectiveness of primary and revision surgery for adult spinal deformity. *Spine (Phila Pa 1976)* 2018;43:791–7.
26. Safaee MM, Dalle Ore CL, Zygourakis CC, et al. The unreimbursed costs of preventing revision surgery in adult spinal deformity: analysis of cost-effectiveness of proximal junctional failure prevention with ligament augmentation. *Neurosurg Focus* 2018;44:E13.
27. Yagi M, Ames CP, Keefe M, et al. A cost-effectiveness comparisons of adult spinal deformity surgery in the United States and Japan. *Eur Spine J* 2018;27:678–84.
28. McCarthy IM, Hostin RA, Ames CP, et al. Total hospital costs of surgical treatment for adult spinal deformity: an extended followup study. *Spine J* 2014;14:2326–33.
29. Bari TJ, Ohrt-Nissen S, Hansen LV, et al. Ability of the global alignment and proportion score to predict mechanical failure following adult spinal deformity surgery-validation in 149 patients with two-year follow-up. *Spine Deform* 2019;7:331–7.
30. 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.
31. Yagi M, Daimon K, Hosogane N, et al. Predictive probability of the global alignment and proportion score for the development of mechanical failure following adult spinal deformity surgery in Asian patients. *Spine (Phila Pa 1976)* 2021;46:E80–6.
32. Seki S, Hirano N, Kawaguchi Y, et al. Teriparatide versus low-dose bisphosphonates before and after surgery for adult spinal deformity in female Japanese patients with osteoporosis. *Eur Spine J* 2017;26:2121–7.
33. Lafage R, Line BG, Gupta S, et al. Orientation of the upper-most instrumented segment influences proximal junctional disease following adult spinal deformity surgery. *Spine (Phila Pa 1976)* 2017;42:1570–7.