Spine

SPINE Volume 48, Number 6, pp 414–420

DEFORMITY

Optimizing the Definition of Proximal Junctional Kyphosis: A Sensitivity Analysis

Francis Lovecchio, MD,[a] Renaud Lafage, MS,[b] Breton Line, BS,[c] Shay Bess, MD,[c] Christopher Shaffrey, MD,[d] Han Jo Kim, MD,[a] Christopher Ames, MD,[e] Douglas Burton, MD,[f] Munish Gupta, MD,[g] Justin S. Smith, MD, PhD,[h] Robert Eastlack, MD,[i] Eric Klineberg, MD,[j] Gregory Mundis, MD,[i] Frank Schwab, MD,[b] Virginie Lafage, PhD,[b] and on behalf of the International Spine Study Group ISSG

Study Design. Diagnostic binary threshold analysis.

Objective. (1) Perform a sensitivity analysis demonstrating the test performance metrics for any combination of proximal junctional angle (PJA) magnitude and change; (2) Propose a new proximal junctional kyphosis (PJK) criteria.

Summary of Background Data. Previous definitions of PJK have been arbitrarily selected and then tested through retrospective case series, often showing little correlation with clinical outcomes.

Materials and Methods. Surgically treated adult spinal deformity patients (≥ 4 levels fused) enrolled into a prospective, multicenter database were evaluated at a minimum 2-year follow-up for proximal junctional failure (PJF). Using PJF as the outcome of interest, test performance metrics including sensitivity, positive predictive value, and F1 metrics (harmonic mean of precision and recall) were calculated for all combinations of PJA magnitude and change using different combinations of perijunctional vertebrae. The combination with the highest F1 score was selected as the new PJK criteria. Performance metrics of previous PJK definitions and the new PJK definition were compared.

Results. Of the total, 669 patients were reviewed. PJF rate was 10%. Overall, the highest F1 scores were achieved when the upper instrumented vertebrae −1 (UIV−1)/UIV+2 angle was measured. For lower thoracic cases, out of all the PJA and magnitude/change combinations tested, a UIV−1/UIV+2 magnitude of −28° and a change of −20° was associated with the highest F1 score. For upper thoracic cases, a UIV−1/UIV+2 magnitude of −30° and a change of −24° were associated with the highest F1 score. Using PJF as the outcome, patients meeting this new criterion (11.5%) at 6 weeks had the lowest survival rate (74.7%) at 2 years postoperative, compared with Glattes (84.4%) and Bridwell (77.4%).

Conclusions. Out of all possible PJA magnitude and change combinations, without stratifying by upper thoracic versus lower thoracic fusions, a magnitude of $\leq -28°$ and a change of $\leq -22°$ provide the best test performance metrics for predicting PJF.

Key words: proximal junctional kyphosis, proximal junctional failure, PJK, PJF, postoperative outcomes, adult spinal deformity

Spine 2023;48:414–420

Proximal junctional kyphosis (PJK), the most common complication after adult spinal deformity surgery, is often defined per the criteria proposed by Glattes *et al*[1] [a kyphotic change in the proximal junctional angle (PJA) of at least 10°]. However, this definition has been criticized for its lack of correlation with clinical outcomes.[1–4] Large series have demonstrated similar Scoliosis Research Society (SRS)-24 scores in patients with and without PJK.[2,3] Furthermore, radiographic PJK does not herald a need for revision surgery: rates of revision for PJK are as low as one-tenth the incidence of the radiographic finding.[4,5] To rectify this issue, investigators have proposed increasing the PJA threshold to 20° or using classification schemes that take into account factors other than the PJA.[6–8] Despite this controversy, the Glattes criteria remains the most popular definition.[9]

From the [a]Department of Orthopedic Surgery, Hospital for Special Surgery, New York, NY; [b]Department of Orthopedic Surgery, Northwell Health, Lenox Hill Hospital, New York, NY; [c]Denver International Spine Center, Presbyterian St. Luke's/Rocky Mountain Hospital for Children, Denver, CO; [d]Department of Neurosurgery, Duke University Medical Center, Durham, NC; [e]Department of Neurosurgery, University of California School of Medicine, San Francisco, CA; [f]Department of Orthopedic Surgery, University of Kansas Medical Center, Kansas City, KS; [g]Department of Orthopedic Surgery, Washington University, St Louis, MO; [h]Department of Neurosurgery, University of Virginia Medical Center, Charlottesville, VA; [i]Department of Orthopedic Surgery, Scripps Clinic Torrey Pines, La Jolla, CA; and [j]Department of Orthopedic Surgery, University of California, Davis, Sacramento, CA.

Acknowledgment date: October 3, 2022. Acceptance date: November 27, 2022.

The International Spine Study Group (ISSG) is funded through research grants from DePuy Synthes (current), Nuvasive (current), K2M (current), Innovasis (past), Biomet (past), and individual donations.

The authors report no conflicts of interest.

Address correspondence and reprint requests to Virginie Lafage, PhD, Lenox Hill | Northwell Health, 110 E 77th Street, New York, NY 10075; E-mail: virginie.lafage@gmail.com

DOI: 10.1097/BRS.0000000000004564

From a pathophysiologic standpoint, proximal junctional pathology likely occurs along a spectrum of severity. However, defining PJK for clinical and research purposes requires setting a threshold along this spectrum. Because the definition will be used to both identify the incidence of PJK and inform clinical decision-making, the chosen threshold will inevitably fall short in one of the two areas. For example, a low threshold will likely capture most proximal junctional pathology but have little to no clinical prognostic value. On the other end, a high threshold may be useful for decision-making, but fail to capture some patients who are early on in the disease process.

Any clinical decision threshold should ideally be chosen based on test performance criteria [sensitivity, positive predictive value (PPV), and accuracy], as the clinical scenario or need dictates the criteria to be used. Rather than set an arbitrary threshold and test its performance within a given cohort (as has been done in the past),[1–4] a more nuanced approach to defining PJK could be to analyze the test properties of various PJK criteria. Recent literature on proximal junctional failure (PJF) prevention strategies and pathophysiology suggest that the vertebrae involved in the PJK definition [upper instrumented vertebrae (UIV)-UIV–1, UIV+1-UIV-1, *etc.*] must also be considered.[10]

In the setting of PJK, the caveat to this approach is that test performance characteristics are based on a set outcome (*e.g.,* the condition to be tested for). In the absence of patient-reported outcomes (PROs), we argue that the ability of a PJK threshold to predict the need for revision is the most clinically significant consequence.[8] Thus, using the need for revision surgery as the outcome of interest, the purposes of the present study are: (1) To perform a sensitivity analysis demonstrating the test performance metrics for any combination of PJA magnitude and change and (2) to propose novel PJK criteria based on the combination with the best performance metrics. Finally, a secondary purpose of this analysis is to compare the test performance characteristics of this novel criteria with historical criteria (specifically, those of Glattes *et al*[1] and Bridwell *et al*[6]).

MATERIALS AND METHODS

Patient Sample

A retrospective review was performed of a prospectively collected database of adult patients ($\geq$18 yr old) who underwent spine deformity surgery at one of 11 participating centers. Institutional Review Board approval was obtained at each site. The data set only includes patients who met the following preoperative radiographic inclusion criteria: (1) Coronal Cobb angle $>$20°, (2) sagittal vertical axis (SVA) $>$5 cm, (3) pelvic tilt $>$25°, or (4) thoracic kyphosis (TK) $>$60°.

Data Collection

Demographics and surgical information contained in the data set were collected through chart review at each center. Using previously validated software (Spineview, ENSAM, Paris),[11] independent research technicians measured full-length, standing lateral radiographs for the following spinopelvic parameters: (1) Pelvic incidence (PI), (2) pelvic tilt, (3) C7 SVA, (4) lumbar lordosis [LL (L1-S1)], (5) PI-LL mismatch (PI-LL), (6) T1-pelvic angle, (7) TK (T1-T12), and (8) thoracolumbar alignment (T10-L2). PJAs were measured at the following levels: (1) UIV to UIV+2, (2) UIV–1 to UIV+1, and (3) UIV–1 to UIV+2 (Figure 1). Radiographs were measured preoperatively, and at 6 weeks, 1 year, and 2 years postoperatively. For patients who underwent revision of their index procedure, imaging on the date closest to the revision surgery was used for "prerevision" measurements.

Determining Proximal Junctional Angles Associated With Proximal Junctional Failure

The primary outcome of interest, PJF, was defined as a surgical revision performed for symptomatic PJK. All PJF cases in the database with a UIV between T1 and T12 were identified and then confirmed by a board-certified orthopedic spine surgeon specializing in adult deformity correction (S. B.). To determine the PJA associated with PJF, preoperative (*i.e.,* before index surgery) and prerevision (*i.e.,* before revision surgery) spinopelvic parameters, and

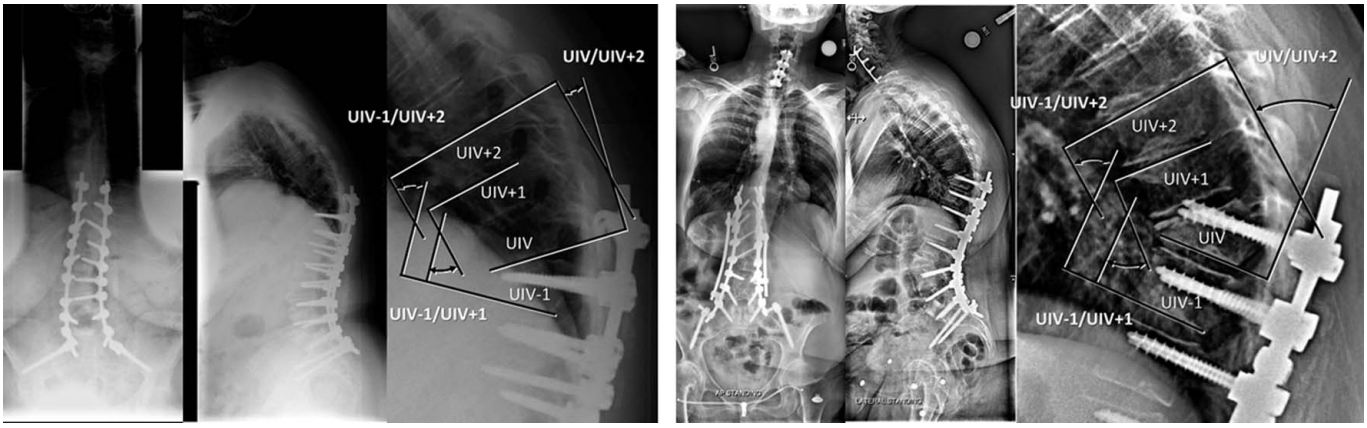

Figure 1. Proximal junctional angles were measured between UIV and UIV+2, UIV–1 and UIV+1, and finally UIV–1 and UIV+2. UIV indicates upper instrumented vertebrae.

PJAs were compared in cases of PJF. Furthermore, to determine the influence of UIV on the PJA associated with PJF, PJA comparisons were repeated after stratifying the PJF cohort by UIV [lower thoracic (LT), UIV T7-T12; upper thoracic (UT), UIV T1-T6].

Test Properties of Proximal Junctional Angle Magnitude/Change Combinations in Predicting Proximal Junctional Failure

Using a cohort of all patients in the database with a UIV between T1 and T12 and ≥2-year follow-up, various combinations of PJA magnitude and change were tested for their ability to predict PJF. In PJF patients, PJA change was calculated between the preoperative and prerevision imaging, and in patients without PJF, preoperative to 2-year postoperative imaging was used. The sensitivity, PPV, and F1 scores of the various combinations were calculated and graphically represented on heat maps. The F1 score is the mathematical harmonic mean between the sensitivity and PPV, and functions as a substitute for the accuracy of a test when modeling imbalanced samples (*i.e.,* samples, in which the outcome of interest occurs in a ratio other than 1:1).[12] Out of all the PJA measured (UIV/UIV+2, UIV−1 /UIV+1, and UIV−1/UIV+2), the magnitude and change combination that led to the highest F1 score (*e.g.*, criteria with the best balance between sensitivity and PPV) was chosen as the new radiographic PJK criteria. For practical reasons, thresholds were rounded to the nearest integer. To facilitate comparisons to historical criteria, the sensitivity, PPV, and F1 scores of the Glattes and Bridwell criteria were also calculated and depicted on the UIV/UIV+2 heat maps.

Survival Analysis of Early Radiographic Proximal Junctional Kyphosis

With PJF as the outcome of interest, revision, as well as 1 and 2-year survival rates, were calculated for patients who met the various PJK criteria (our new criteria, Bridwell, and Glattes) at 6 weeks.

Statistical Analyses

All statistical analyses were performed using IBM SPSS Version 25.0 (Armonk, NY). Categorical variables were compared using χ^2 tests. Continuous variables were assessed for normality using the Kolmogorov-Smirnof test and compared using student *t* tests or Wilcoxon rank-sum as appropriate. A *P* value of <0.05 was set as the type I error rate for statistical significance.

RESULTS

Determining Proximal Junctional Angles Associated With Proximal Junctional Failure

Eighty-five revision cases for PJF were identified from the data set (73 unique patients; 62 with one revision, 10 with 2 revisions, and 1 with 3 revisions). Five were excluded due to missing prerevision imaging, leaving 80 cases available for analysis. The mean age was 65.6 ± 9.2 years, with 76% of females with a mean body mass index of 28.2 ± 5.5 kg/m^2. The majority of cases were fused to the pelvis (90%) with an LT UIV (65%). Mean fusion length was 10.7 ± 3.6 levels, 72.5% underwent a Schwab Grade 2 osteotomy, 25.4% Schwab Grade 3 or higher, and 82.1% received interbody fusion (mean of 2.2 ± 1.3 interbodies). All spinopelvic radiographic parameters (with the exceptions of PI and T10-L2 sagittal angle) improved between baseline and prerevision imaging (Table 1).

The magnitude of prerevision PJA varied depending on the number of vertebrae measured, with an increased average magnitude as a larger number of vertebrae were included (UIV/UIV+2 −24.4°, UIV−1/UIV+1 −32.0°, UIV−1/UIV+2 −34.8°; $P < 0.001$). A similar trend was seen with regards to the change in PJA (UIV/UIV+2 −20.2°, UIV−1/UIV+1 −26.3°, UIV−1/UIV+2 −27.5°; $P < 0.001$). When stratified by UIV, the PJA magnitude and change were significantly different only when UIV−1/UIV+1 and UIV−1/UIV+2 were used ($P < 0.05$) (Table 2).

Rate of Proximal Junctional Failure in Test Cohort

A total of 669 patients met the criteria for inclusion in the test cohort. The mean age of the cohort was 59.3 ± 14.3 years, 80% females, with an average body mass index of 27.6 ± 6.1 kg/m^2. Mean number of levels fused was 11.5 ± 3.4 with a UIV located mainly at T3/T4 (34.1%) or T10/T11 (39.8%). Most (72.3%) underwent a Schwab Grade 2 osteotomy, 20.2% Schwab Grade 3 or higher, and 59.7% received interbody fusion (mean of 2.4 ± 1.6 interbodies). Among these patients, 67 (10.0%) experienced PJF (whereas 73 unique patients had PJF in the entire data set and 6 were excluded from the test cohort as they did not meet 2-year follow-up). The median time to PJF was

TABLE 1. Changes in Preoperative Parameters in Patients With PJF

	N	Mean	SD	*P*
Preoperative PI	80	53.09	11.54	0.627
Prerevision PI	80	53.33	11.97	—
Preoperative PT	80	25.33	10.33	0.001
Prerevision PT	80	22.62	9.30	—
Preoperative PI-LL	80	17.25	22.10	0.000
Prerevision PI-LL	80	−0.01	14.14	—
Preoperative T10-L2	80	−18.40	18.07	0.901
Prerevision T10-L2	80	−18.72	17.45	—
Preoperative T1-T12 (TK)	80	−40.30	21.22	0.000
Prerevision T1-T12 (TK)	80	−64.69	20.04	—
Preoperative TPA	80	24.98	12.88	0.000
Prerevision TPA	80	18.57	10.33	—
Preoperative SVA	80	80.24	71.60	0.000
Prerevision SVA	80	35.20	58.39	—

LL indicates lumbar lordosis; PI, pelvic incidence; PT, pelvic tilt; PJF, proximal junctional failure; SVA, sagittal vertical axis; TK, thoracic kyphosis; TPA, T1-pelvic angle.

TABLE 2. PJAs Before Revision for PJF, Measured Across Different Levels About the UIV and Stratified by Location of the UIV

	Prerevision PJA	*P* UT *vs.* LT	PJA Change Basline-To-Prerevision	*P* UT *vs.* LT
UIV/UIV+2	−24.4° ± 14.3	—	−20.2° ± 14.3	—
UT	−28.5° ± 15.7	0.219	−24.2° ± 17.3	0.174
LT	−24.4° ± 12.0	—	−19.4° ± 12.3	—
UIV−1/UIV+1	−32.0° ± 14.3	—	−26.3° ± 13.5	—
UT	−40.3° ± 17.3	0.010	−32.6° ± 16.3	0.037
LT	−29.8° ± 10.2	—	−24.6° ± 10.6	—
UIV−1/UIV+2	−34.8° ± 14.9	—	−27.5° ± 13.8	—
UT	−42.0° ± 16.8	0.032	−33.5° ± 16.1	0.046
LT	−33.6° ± 10.9	—	−26.0° ± 11.3	—

LT indicates lower thoracic; PJA, proximal junctional angle; PJF, proximal junctional failure; UIV, upper instrumented vertebrae; UT, upper thoracic.

12 months (interquartile range, 16 months), with 80.6% (54) occurring before 2 years, 47.6% (32) before 1 year, and 7.5% (5) before 6 weeks.

Test Properties of Proximal Junctional Angle Magnitude/Change Combinations in Predicting Proximal Junctional Failure

With PJF as the outcome of interest, the PPV, sensitivity, and F1 scores of various PJA magnitude/change combinations are graphically depicted on heat maps (Figures 2–4). The highest PPV correlated with the largest magnitude and change (left graph in Figures 2–4), whereas the highest sensitivity was achieved by minimizing these parameters (middle graph in Figures 2–4). In contrast, the F1 score showed a more complex relationship to magnitude and change due to the opposing progression of sensitivity and PPV (right graph in Figures 2–4). Overall, the highest F1 scores were achieved when the UIV–1/UIV+2 angle was measured.

Stratification by UIV revealed similar patterns in the LT cohort, but more complex shapes in the UT cohort. For LT cases, out of all the PJA and magnitude/change combinations tested, a UIV–1/UIV+2 magnitude of –28° and a change of –20° were associated with the highest F1. For UT cases, a UIV–1/UIV+2 magnitude of –30° and a change of –24° were associated with the highest F1 score. For simplicity and ease of clinical use, a magnitude of –28° and a change of –22° for both UT and LT UIV were proposed as the novel PJK criteria. This definition represents the best compromise to have a single definition for the entire thoracic spine and the highest F1 score possible.

Performance of Proximal Junctional Kyphosis Criteria in Predicting Proximal Junctional Failure

The novel and the historic PJK criteria (Glattes and Bridwell) were plotted on the corresponding heat maps (Figures 2–4). When the Glattes criteria were applied to the sample, the rate of radiographic PJK was found to be 48.2%. With PJF as the outcome of interest, the PPV was 16.2%, sensitivity 77.6%, and F1 score 26.7%. When the Bridwell criteria were applied, the rate of radiographic PJK was 34.3%, with a PPV of 19.7%, sensitivity of 67.2%, and an F1 score of 30.5%. Finally, the application of the new criteria showed a PPV of 33.8%, sensitivity of 34.3%, and F1 score of 34.1%.

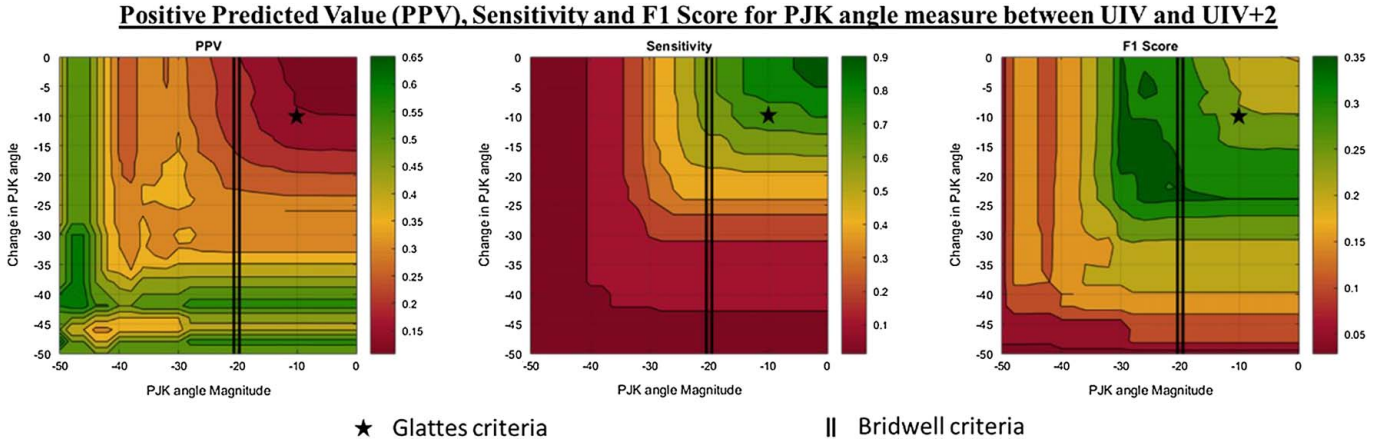

Figure 2. Heat map displaying PPV, Sensitivity, and F1 score for PJK angle measure between UIV and UIV+2. Historical thresholds (Glattes and Bridwell criteria) display on each graph. PJK indicates proximal junctional kyphosis; PPV, positive predictive value; UIV, upper instrumented vertebrae. full color online

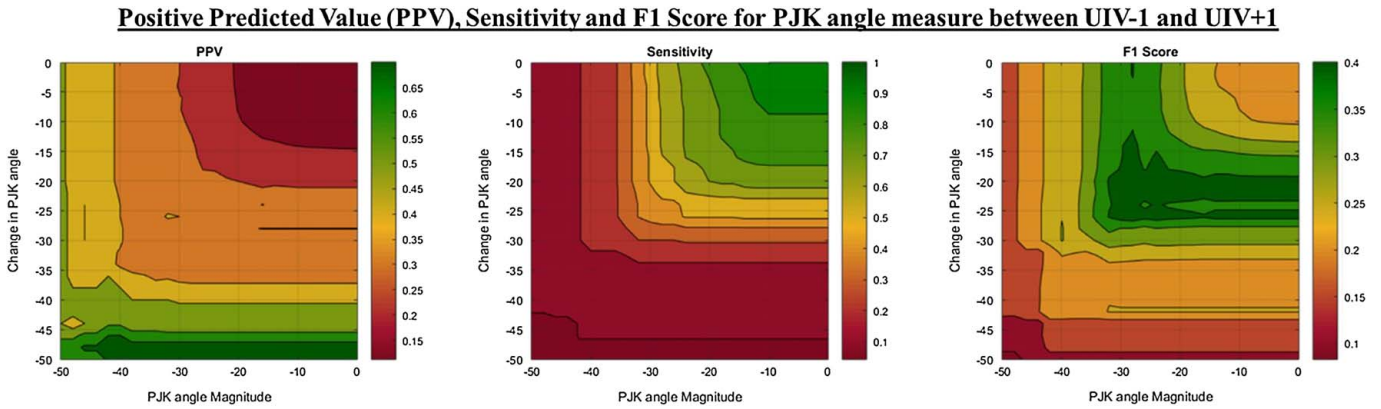

Figure 3. Heat map displaying PPV, Sensitivity, and F1 score for PJK angle measure between UIV−1 and UIV+1. PJK indicates proximal junctional kyphosis; PPV, positive predictive value; UIV, upper instrumented vertebrae. full color online

Survival Analysis of Early Radiographic Proximal Junctional Kyphosis

Among the 225 patients who met Glattes criteria at 6 weeks (33.6%), 47 (20.9%) underwent revision surgery for PJK, with 23 (48.9%) occurring before 1 year and 35 (85.4%) before 2 years, corresponding to survival rates of 89.8% and 84.4% at 1 and 2 years, respectively. For the 156 patients who met Bridwell criteria (23.3%), similar rates of revision were seen (N = 17, 47.2% before 1 year and N = 24, 77.4% before 2 years), corresponding to survival rates of 89.1% and 84.6% at 1 and 2 years postoperatively. Finally, patients meeting our criteria (11.5%) showed the highest conversion rate with 26/75 patients revised for PJK (61.5% and 82.6% before 1 and 2 years) and the lowest survival rate (78.7% and 74.7%) at 1 and 2 years postoperatively.

DISCUSSION

In our multicenter cohort of 669 patients, we found a 10% rate of revision for PJK, consistent with contemporary and past series.[13–15] In the absence of PROs associated with PJK, the most important feature of a PJK definition is its ability to accurately identify patients who will have a clinically significant consequence—revision surgery, which we defined as PJF. In our analyses, the PJA magnitude and change in PJA associated with PJF differed based on the vertebrae that were measured, with the UIV–1/UIV+2 angle associated with the largest values. Furthermore, we used test performance metrics (sensitivity, PPV, and F1 score) to evaluate the performance of various PJA magnitude and change thresholds in their ability to predict PJF. When comparing the performance of various PJA angles (*i.e.*, UIV/UIV+2 *vs.* UIV–1/UIV+1, *etc.*), the use of UIV–1/UIV+2 was associated with the highest F1 scores. For LT cases, out of all the PJA and magnitude/change combinations tested, a UIV–1/UIV+2 magnitude of –28° and a change of –20° were associated with the highest F1. For UT cases, a UIV–1/UIV+2 magnitude of –30° and change of –24° was associated with the highest F1 score. Thus, for ease of clinical applications, we propose that a PJA magnitude of ≤–28° and a change of ≤–22° are used as novel criteria to define PJK.

In past studies, PJK definitions were evaluated through retrospective cohort comparisons between arbitrarily selected cutoffs. The most common PJK criteria were proposed by Glattes *et al*[1] in a series of 81 adult deformity patients followed for a minimum of 2 years. In their study, they found a PJK rate of 26%, noting that SRS-24 scores were similar between PJK and non-PJK patients. Given the failure of a 10° and even higher threshold (15°) to show

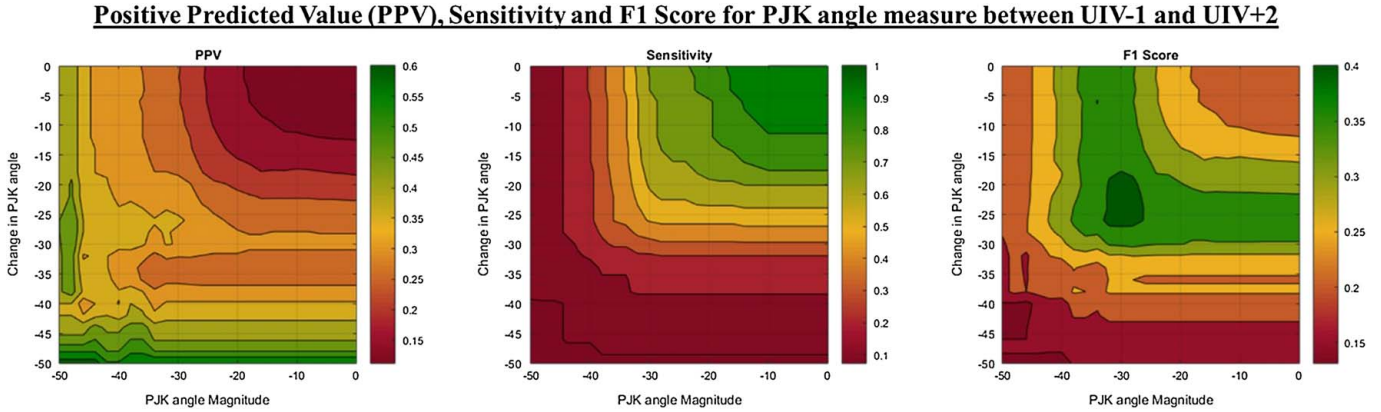

Figure 4. Heat map displaying PPV, Sensitivity, and F1 score for PJK angle measure between UIV−1 and UIV+2. PJK indicates proximal junctional kyphosis; PPV, positive predictive value; UIV, upper instrumented vertebrae. full color online

differences in PROs,[2,16] Bridwell *et al*[6] proposed a cutoff of 20°. In a series of 90 primary adult deformity surgeries, the authors reported a 27.8% PJK rate, with no differences in the SRS-22 or Oswestry Disability Index between the cohorts. In contrast, a recent study of patients fused to the thoracolumbar junction found that patients with PJF (PJA of 20°, fracture, or hardware pullout) did report lower SRS-22 and higher Oswestry Disability Index scores.[17] However, 14/23 (61%) of the PJF cohort had a fracture or hardware complications, weakening comparisons to other research. Notably, none of the aforementioned studies reported the test performance characteristics of their chosen definitions, key limitations to the evaluation, and comparison of the various cutoffs.

The performance of a diagnostic binary threshold can be evaluated by using 2 metrics (*i.e.,* sensitivity and PPV) or a single "composite" metric (*e.g.,* F1 score or accuracy). Composite metrics allow easy comparisons between various thresholds, making them particularly suitable for sensitivity analyses. In our investigation, we used the F1 score as a composite measure of test performance. Although relatively new to medicine, the metric has been used to evaluate the predictive utility of machine learning algorithms.[18–20] The F1 score substitutes for the test's accuracy (*i.e.,* the rate, at which a test identifies true positives and true negatives). However, the disadvantage to using accuracy is that it is unreliable when the outcome of interest occurs rarely (*i.e.,* 10% or less),[12] which mandated the use of the F1 score. Our chosen PJK criteria were associated with the highest F1 score achieved in the analyses (34.1%). Clinically, this suggests that approximately a third of the time the test will correctly identify a patient that will undergo revision for PJK. These findings were supported by our survival analysis, which showed that among those patients who met our criteria at 6 weeks, 25.3% underwent revision by 2 years.

Although our criteria performed better than those of Glattes and Bridwell (F1 scores of 26.7% and 30.5%, respectively), our analyses support the claim that a truly comprehensive, accurate prediction of PJF requires the integration of patient- and surgery-related factors outside of PJA.[2,6,7,17,21] In other words, any PJA-only-based definition of PJK is destined to perform poorly, as it does not account for other factors that may predict junctional failure.[8] Investigators have repeatedly shown that poor patient-reported outcomes or the need for revision surgery involve many other factors other than the PJA.[2,6,7,21] Thus, although we believe that our new criteria represent the best possible solution to the problem of defining PJK, we recognize that further work is needed to elucidate the many mechanical and biological aspects that influence this condition, ultimately arriving at a comprehensive definition based on more than just PJA.

In addition to sensitivity analyses on change and magnitude, we also examined which vertebrae should be used to define PJK, finding that definitions, which included a larger number of perijunctional vertebrae (*i.e.,* UIV–1/UIV+2) were superior performers. This finding is consistent with literature showing that PJK can take the form of various pathophysiologies, which may occur at different levels (*i.e.,* instrumentation pull out at the UIV *vs.* a fracture of the UIV+1).[7] In addition, the effectiveness of PJK prevention strategies may vary based on the vertebrae that are modified,[8,18,19] bolstering our conclusion that using vertebrae above and below the UIV to define PJK is the optimal strategy.

There were several limitations to the current study. The first is that the definition of PJF (our outcome of interest) was based on whether the patient underwent revision. We could not account for patients who were clinically symptomatic but were unfit for surgery or refused a revision operation. Therefore, the rate of PJF in our study is likely underestimating the actual rate of clinically relevant PJK. Second, as previously mentioned, our PJK definition is based only on the PJA and changes in PJA. The creation of a comprehensive definition is currently limited by the lack of PROs differentiating clinically relevant PJK and the use of retrospective data sets with incomplete data capture. In other words, if feasible, a prospective study with accurate, long-term data capture on hundreds of variables in clinically complex adult deformity patients could best answer the question of how to define PJK. Finally, we purposefully limited our sensitivity analyses to 2 classically used parameters to define PJK. The use of other parameters (*i.e.,* TK, SVA, *etc.*) lead to higher PPVs, but the representations were too complex, often leading to conclusions that did not make clinical sense.

CONCLUSION

In summary, out of all possible PJA magnitude and change combinations, a magnitude of $\leq -28°$, and a change of $\leq -22°$ provide the best test performance metrics for the prediction of revision for PJK. Until clinically significant PJK can be defined using PROs, we propose that these cutoffs are utilized to define PJK in future research given their association with a clinically significant consequence. Further research will be needed to elucidate the myriad of factors that predict the need for revision other than the PJA alone.

➢ Key Points

- ❑ Different PJK definitions in the literature have been criticized for their lack of correlation with clinical outcomes.
- ❑ PJA of $\leq -28°$ and change of $\leq -22°$ provide the best test performance metrics for the prediction of PJF.
- ❑ We propose utilizing these cutoffs to define PJK in future research given their association with a clinically significant consequence.

References

1. Glattes RC, Bridwell KH, Lenke LG, et al. Proximal junctional kyphosis in adult spinal deformity following long instrumented posterior spinal fusion: Incidence, outcomes, and risk factor analysis. *Spine (Phila Pa 1976)*. 2005;30:1643–9.
2. Kim YJ, Bridwell KH, Lenke LG, et al. Proximal Junctional Kyphosis in Adult Spinal Deformity After Segmental Posterior Spinal Instrumentation and Fusion. *Spine (Phila Pa 1976)*. 2008;33:2179–84.
3. Yagi M, King AB, Boachie-Adjei O. Incidence, risk factors, and natural course of proximal junctional kyphosis: surgical outcomes review of adult idiopathic scoliosis. Minimum 5 years of follow-up. *Spine (Phila Pa 1976)*. 2012;37:1479–89.
4. Sebaaly A, Sylvestre C, El Quehtani Y, et al. Incidence and risk factors for proximal junctional kyphosis. *Clin Spine Surg A Spine Publ*. 2018;31:E178–83.
5. Park S-J, Park J-S, Nam Y-J, et al. The long-term fate of asymptomatic proximal junctional kyphosis following long instrumented fusion in elderly patients with sagittal imbalance. *Spine (Phila Pa 1976)*. 2021;46:E1097–104.
6. Bridwell KH, Lenke LG, Cho SK, et al. Proximal junctional kyphosis in primary adult deformity surgery: evaluation of 20 degrees as a critical angle. *Neurosurgery*. 2013;72:899–906.
7. 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.
8. Park S-J, Park J-S, Nam Y, et al. Who will require revision surgery among neurologically intact patients with proximal junctional failure after surgical correction of adult spinal deformity? *Spine (Phila Pa 1976)*. 2020;46:520–9.
9. Mika A, Mesfin A, Rubery P, et al. Proximal junctional kyphosis: a pediatric and adult spinal deformity surgery dilemma. *JBJS Reviews*. 2019;7:e4.
10. Buell TJ, Bess S, Xu M, et al. Optimal tether configurations and preload tensioning to prevent proximal junctional kyphosis: A finite element analysis. *J Neurosurg Spine*. 2019;30:574–84.
11. Champain S, Benchikh K, Nogier A, et al. Validation of new clinical quantitative analysis software applicable in spine orthopaedic studies. *Eur Spine J*. 2006;15:982–91.
12. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag*. 2009;45:427–37.
13. Kelly MP, Lurie JD, Yanik EL, et al. Operative versus nonoperative treatment for adult symptomatic lumbar scoliosis. *J Bone Jt Surg - Am Vol*. 2019;101:338–52.
14. Line BG, Bess S, Lafage R, et al. Effective Prevention of proximal junctional failure in adult spinal deformity surgery requires a combination of surgical implant prophylaxis and avoidance of sagittal alignment overcorrection. *Spine (Phila Pa 1976)*. 2020;45:258–67.
15. Maruo K, Ha Y, Inoue S, et al. Predictive factors for proximal junctional kyphosis in long fusions to the sacrum in adult spinal deformity. *Spine (Phila Pa 1976)*. 2013;38:E1469–76.
16. Helgeson MD, Shah SA, Newton PO, et al. Evaluation of proximal junctional kyphosis in adolescent idiopathic scoliosis following pedicle screw, hook, or hybrid instrumentation. *Spine (Phila Pa 1976)*. 2010;35:177–81.
17. Park SJ, Lee CS, Park JS, et al. Should thoracolumbar junction be always avoided as upper instrumented vertebra in long instrumented fusion for adult spinal deformity?: risk factor analysis for proximal junctional failure. *Spine (Phila Pa 1976)*. 2020;45:686–93.
18. Wu JT, Wong KCL, Gur Y, et al. Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents + supplemental content. *JAMA Netw Open*. 2020;3:2022779.
19. Kilic A, Goyal A, Miller JK, et al. Predictive utility of a machine learning algorithm in estimating mortality risk in cardiac surgery. *Ann Thorac Surg*. 2020;109:1811–9.
20. Crowson MG, Hartnick CJ, Diercks GR, et al. Machine learning for accurate intraoperative pediatric middle ear effusion diagnosis. *Pediatrics*. 2021;147:e2020034546.
21. Cerpa M, Sardar Z, Lenke L. Revision surgery in proximal junctional kyphosis. *Eur Spine J*. 2020;29:78–85.