# The Modified Fels and Abbreviated Modified Fels Knee Skeletal-Maturity Systems in the Prediction of Leg-Length Discrepancy

Dylan G. Kluck, MD, Marina R. Makarov, MD, David A. Podeszwa, MD, Ryan J. Furdock, MD, Raymond W. Liu, MD, Chan-Hee Jo, PhD, and John G. Birch, MD, FRCS(C)

*Investigation performed at the Texas Scottish Rite Hospital for Children, Dallas, Texas*

**Background:** The Modified Fels (mFels) and Abbreviated Modified Fels (abFels) knee systems have been recently developed as options for grading skeletal maturity without the need for a separate hand radiograph. We sought to determine the interobserver reliability of these systems and to compare their prediction accuracy with that of the Greulich and Pyle (G-P) atlas in a cohort managed with epiphysiodesis for leg-length discrepancy (LLD).

**Methods:** Three reviewers scored 20 knee radiographs using the mFels system, which includes 5 qualitative and 2 quantitative measures as well as a quantitative output. Short leg length (SL), long leg length (LL), and LLD prediction errors at maturity using the White-Menelaus (W-M) method and G-P, mFels, or abFels skeletal age were compared in a cohort of 60 patients managed with epiphysiodesis for LLD.

**Results:** Intraclass correlation coefficients for the 2 quantitative variables and the quantitative output of the mFels system using 20 knee radiographs ranged from 0.55 to 0.98, and kappa coefficients for the 5 qualitative variables ranged from 0.56 to 1, indicating a reliability range from moderate to excellent. In the epiphysiodesis cohort, G-P skeletal age was on average 0.25 year older than mFels and abFels skeletal ages, most notably in females. The majority of average prediction errors between G-P, mFels, and abFels were <0.5 cm, with the greatest error being for the SL prediction in females, which approached 1 cm. Skeletal-age estimates with the mFels and abFels systems were statistically comparable.

**Conclusions:** The mFels skeletal-age system is a reproducible method of determining skeletal age. Prediction errors in mFels and abFels skeletal ages were clinically comparable with those in G-P skeletal ages in this epiphysiodesis cohort. Further work is warranted to optimize and validate the accuracy of mFels and abFels skeletal ages to predict LLD and the impact of epiphysiodesis, particularly in females. Both the mFels and abFels systems are promising means of estimating skeletal age, avoiding additional radiation and health-care expenditure.

**Level of Evidence:** Prognostic Level II. See Instructions for Authors for a complete description of levels of evidence.

Estimation of skeletal maturity is an essential component of the evaluation of leg-length discrepancy (LLD) in the growing child. Estimations of remaining skeletal growth and expected LLD at skeletal maturity are key to determining appropriate timing or impact of epiphysiodesis in the management of LLD. Among the available methods to estimate skeletal age, the Greulich and Pyle (G-P) bone-age atlas[1] remains the gold standard, despite concerns as to interobserver reliability and generalizability[2-5]. A previous study demonstrated that the combination of the White-Menelaus (W-M) formulae with G-P skeletal age was superior to both the W-M method with chronological age as well as other prediction methods in epiphysiodesis-aged cohorts[6]. Obtaining skeletal age via the G-P atlas requires a left hand-and-wrist radiograph, resulting in increased radiation exposure for the patient and increased health-care expenditure. There is therefore interest in a skeletal-maturity assessment system based on lower-extremity radiographs.

The Fels Longitudinal Study represents an immense collection of radiographs and patient data collected over most of the last century, providing longitudinal datasets for the assessment of skeletal maturity[7]. From these datasets, the Roche-Wainer-Thissen (RWT) method of knee-based skeletal age was

**Disclosure:** The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJS/H760).

developed utilizing 36 parameters from an anteroposterior knee radiograph[8]. Understandably, the measurement of 36 parameters makes this method impractical for daily clinical use. Other methods, including the Pyle and Hoerr bone-age atlas[9], remain limited by the same subjectivity inherent to the G-P hand atlas.

The recently described Modified Fels (mFels) knee skeletal-maturity system represents a promising method for estimation of skeletal age using an anteroposterior knee radiograph. As described by Benedick et al., the mFels system represents a refinement of the RWT method to 7 (5 qualitative and 2 quantitative) radiographic parameters; the system also includes chronological age and sex, and then outputs from these 9 parameters[10]. The mFels system has been shown to be effective in the prediction of leg length at skeletal maturity, is resistant to rotational variation, and outperforms leg-length predictions based on chronological age[11,12]. In an effort to further streamline this method, the creators of the mFels system also developed an Abbreviated Modified Fels (abFels) system using only 2 or 3 radiographic knee parameters, with potentially similarly effective skeletal maturity estimation[13].

There are limited data establishing the interobserver reliability of the mFels system[5]. In addition, to our knowledge neither the mFels nor the abFels system has been evaluated with respect to predicting LLD at skeletal maturity relative to the G-P atlas. The present study was performed to determine the interobserver reliability of the mFels system. We then sought to compare the accuracy of mFels, abFels, and G-P estimates of skeletal age in the prediction of short leg length (SL), long leg length (LL), and residual LLD at skeletal maturity in a previously reported cohort of patients managed with epiphysiodesis for LLD[6]. We hypothesized that skeletal-age estimates via the mFels and abFels systems would provide prediction accuracy comparable with estimates via the G-P atlas.

## Materials and Methods

To assess the interobserver reliability of the mFels system, a sample cohort with 20 radiographs furnished and scored by the originators of the mFels system was reviewed by 3 independent and blinded interpreters. The intraclass correlation coefficients (ICCs) and kappa coefficients were obtained for both the individual radiographic parameters of the mFels system and the overall estimates of skeletal age.

After receiving institutional review board approval, we used the radiographs and leg-length data of a previously reported cohort of 77 patients managed with epiphysiodesis for LLD[6]. This earlier study established that the W-M formulae combined with skeletal age as determined from the G-P atlas performed better than the W-M method with chronological age and was comparable with or superior to the Green-Anderson growth-remaining[14,15], Rotterdam straight-line graph[16], and Paley multiplier[17] methods in the prediction of SL, LL, and residual LLD at skeletal maturity.

From the original cohort of 77 epiphysiodesis patients, 60 were determined to have adequate perioperative knee radiographs (from either the perioperative scanogram or a dedicated anteroposterior radiograph of the contralateral leg) to allow estimation of mFels and abFels skeletal ages. G-P skeletal ages for the patients in this cohort were previously determined by the senior author using the G-P atlas[1], recorded to the nearest corresponding 6 months. These skeletal-age estimates were subsequently re-reviewed by the same author for the current study to confirm accuracy, with no differences found.

SL, LL, and LLD prediction errors were determined via the W-M method using the G-P, mFels, and abFels skeletal-age estimates[6]. We used the W-M method as previously described, including assuming skeletal maturity at age 16 years in males and 14 years in females and converting annual distal femoral and proximal tibial growth of 3/8 inch and 1/4

**TABLE I Interobserver Reliability Between mFels Creators and 3 Raters Using 20 Knee Radiographs***

| | Rater 1 | Rater 2 | Rater 3 | Overall (95% CI) |
|---|---|---|---|---|
| Kappa coefficient | | | | |
| Femoral capping (FEM-K) | 0.70 | 0.77 | 0.62 | 0.80 (0.59-0.95) |
| Femoral fusion (FEM-F) | 0.88 | 0.77 | 0.56 | 0.78 (0.61-1.00) |
| Tibial lateral capping (TIB-LK) | 0.63 | 0.89 | 0.66 | 0.74 (0.35-1.00) |
| Tibial medial capping (TIB-MK) | 1.00 | 0.79 | 0.69 | 0.78 (0.47-1.00) |
| Tibial fusion (TIB-F) | 0.72 | 0.90 | 0.90 | 0.74 (0.49-0.92) |
| Intraclass correlation coefficient | | | | |
| Proximal tibial epiphyseal-metaphyseal ratio (Tib E:M) | 0.78 | 0.85 | 0.84 | 0.73 (0.52-0.87) |
| Proximal fibular epiphyseal-metaphyseal ratio (Fib E:M) | 0.55 | 0.77 | 0.81 | 0.72 (0.51-0.86) |
| Estimated skeletal age | 0.96 | 0.97 | 0.98 | 0.97 (0.93-0.99) |

*mFels = modified Fels. Kappa coefficients were used for the 5 categorical variables, and intraclass correlation coefficients were used for the 3 continuous variables, to assess the interobserver agreement between the mFels creators and each of the 3 raters. The overall reliability across the 3 raters was then reported. Skeletal age was estimated using the 7 qualitative and quantitative radiographic parameters, excluding chronological age and gender for this portion of the study to isolate the overall reliability of the measured parameters specifically.

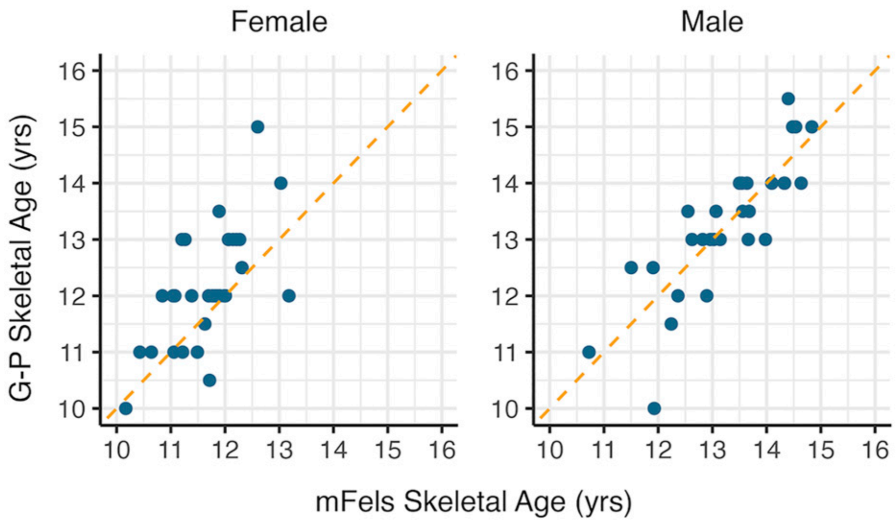


Fig. 1
Scatterplot comparison of skeletal age as determined with the Greulich and Pyle (G-P) atlas and the modified Fels (mFels) system. While skeletal age in males was comparable, skeletal age in females tended to be somewhat younger with the mFels system.

inch to 0.95 cm and 0.64 cm, respectively[18-21]. We used the estimated 71% and 57% contributions from the distal aspect of the femur and proximal aspect of the tibia to calculate expected growth of the entire segment[14] (1.34 cm/year for the femur, 1.12 cm/year for the tibia, and 2.46 cm/year for the entire lower limb). Predicted SL, LL (including a calculated epiphysiodesis effect), and residual LLD were then compared with the actual outcome at maturity to determine the prediction error. Prediction errors were analyzed as an absolute value. We categorized LLD prediction errors of >1 cm as clinically relevant.

Statistical analysis was performed using SAS version 9.4 (SAS Institute) and SPSS version 27 (IBM). Interobserver reliability was assessed using ICCs for continuous variables and kappa coefficients for categorical variables. Continuous variables were first examined for normality with the Shapiro-Wilk test, and a nonparametric test such as the Wilcoxon signed rank test was executed for the comparison. Repeated-measures analysis of variance (ANOVA) was used for the comparison of G-P, mFels, and abFels skeletal-age estimates followed by a paired t test comparing G-P with mFels and abFels-based skeletal-age estimates for all leg-length prediction errors. Significance was defined as a p value of <0.05.

*Source of Funding*
There was no external source of funding for this study.

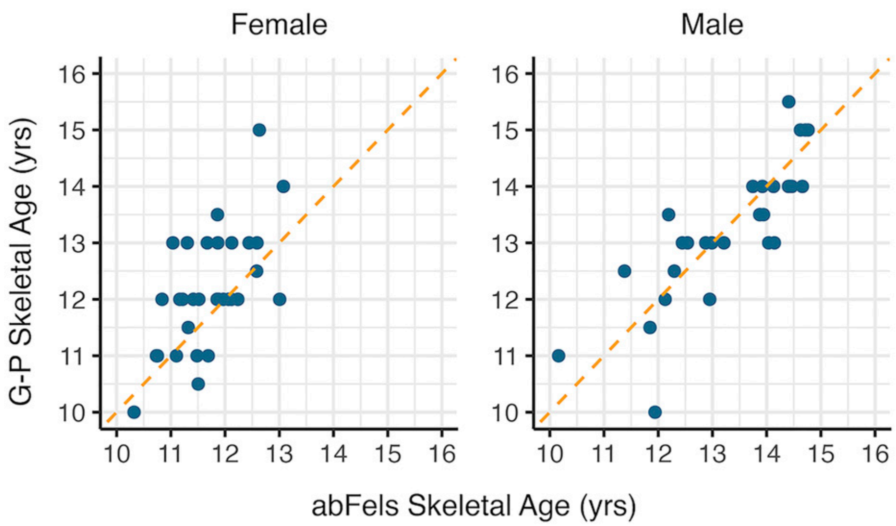


Fig. 2
Scatterplot comparison of skeletal age as determined with the Greulich and Pyle (G-P) atlas and the abbreviated Fels (abFels) system. Comparative abFels skeletal ages for males and females were similar to those noted with the mFels system.

**TABLE II G-P, mFels, and abFels Skeletal-Age Estimates in the 60-Patient Cohort***

| | Skeletal-Age Estimate *(yr)* | | | Mean Difference Between Methods *(yr)* (P Value)† | | |
|---|---|---|---|---|---|---|
| Group (Size) | G-P | mFels | abFels | G-P Versus mFels | G-P Versus abFels | mFels Versus abFels |
| Entire cohort (60) | 12.7 ± 1.3 | 12.4 ± 1.2 | 12.5 ± 1.5 | 0.3 ± 0.8 (**0.01**) | 0.2 ± 0.8 (0.07) | −0.1 ± 0.3 (0.09) |
| Females (31) | 12.2 ± 1.1 | 11.7 ± 0.7 | 11.7 ± 0.7 | 0.5 ± 0.8 (**<0.01**) | 0.4 ± 0.8 (**<0.01**) | 0 ± 0.2 (0.28) |
| Males (29) | 13.3 ± 1.2 | 13.3 ± 1.0 | 13.3 ± 1.2 | 0 ± 0.7 (0.98) | −0.1 ± 0.7 (0.59) | −0.1 ± 0.3 (0.19) |

*G-P = Greulich and Pyle atlas, mFels = modified Fels, and abFels = abbreviated modified Fels. †The mean difference ± SD are reported as skeletal age (G-P) – skeletal age (mFels), skeletal age (G-P) – skeletal age (abFels), and skeletal age (mFels) – skeletal age (abFels). P values are derived from repeated-measures ANOVA.

## Results

Overall ICC values across 3 raters for the 2 quantitative components of the mFels system (the proximal tibial and proximal fibular epiphyseal-metaphyseal ratios) measured 0.73 and 0.72 respectively, indicating moderate interobserver reliability (Table I). For the 5 qualitative components of the mFels system, overall kappa coefficients across 3 raters ranged from 0.74 to 0.80, indicating moderate to good agreement between raters. Skeletal-age estimates demonstrated excellent reliability, with the overall ICC measuring 0.97 (95% confidence interval [CI], 0.93 to 0.99).

The epiphysiodesis cohort consisted of 31 females and 29 males with an average chronological age (and standard deviation [SD]) at the time of surgery of 11.9 ± 1.0 years (range, 10.5 to 14.3 years) and 13.0 ± 1.3 years (range, 9.6 to 15.6 years), respectively. The average LLD at time of the intervention was 3.7 ± 1.3 cm, and the average residual LLD was 2.1 ± 1.1 cm.

In the 60-subject epiphysiodesis cohort, mFels and abFels skeletal ages in males were virtually identical to G-P skeletal age. In females, mFels and abFels skeletal ages averaged 0.5 year less than G-P skeletal age; these differences were significant (Figs. 1 and 2). For the entire cohort, mFels and abFels ages were on average 0.25 year less than ages determined with the G-P atlas, with significance for the mFels system. There were no significant differences between the mFels and abFels skeletal ages in this cohort (Table II).

The comparison of SL, post-epiphysiodesis LL, and residual LLD prediction errors using the W-M formulae and G-P, mFels, or abFels skeletal age in the 60-patient epiphysiodesis cohort is summarized in Table III. The average SL prediction error with the mFels or abFels skeletal age was 0.9 cm more than with the G-P atlas in females, 0.3 to 0.7 cm less in males, and 0.1 to 0.3 cm more in the entire cohort. The average post-epiphysiodesis LL prediction error was 0.4 cm with the mFels or abFels skeletal age compared with the G-P age in females, 0.3 to 0.5 cm less in males, and virtually identical (1.4 to 1.5 cm) in the entire cohort. The average LLD prediction error with the mFels or abFels was 0.2 to 0.3 cm more compared with G-P in females, virtually identical in males, and 0.2 to 0.3 cm more in the entire cohort.

Clinically relevant differences (which we defined as >1.0 cm between predicted and observed values) were noted throughout

**TABLE III Leg-Length Prediction Errors at Maturity Using W-M Formulae and G-P, mFels, and abFels Skeletal-Age Estimates in the 60-Patient Cohort***

| | | Prediction Error *(cm)* | | | Mean Difference Between Methods *(cm)* (P Value)† | | |
|---|---|---|---|---|---|---|---|
| Length Prediction | Group (Size) | G-P | mFels | abFels | G-P Versus mFels* | G-P Versus abFels | mFels Versus abFels |
| Short leg | Entire cohort (60) | 2.0 ± 1.5 | 2.3 ± 1.4 | 2.1 ± 1.3 | −0.3 ± 1.6 (0.19) | −0.1 ± 1.7 (0.68) | 0.2 ± 0.6 (**0.03**) |
| | Females (31) | 1.4 ± 1.0 | 2.3 ± 1.3 | 2.3 ± 1.3 | −0.9 ± 1.5 (**<0.01**) | −0.9 ± 1.5 (**<0.01**) | 0 ± 0.5 (0.81) |
| | Males (29) | 2.6 ± 1.7 | 2.3 ± 1.5 | 1.9 ± 1.2 | 0.4 ± 1.5 (0.21) | 0.8 ± 1.4 (**<0.01**) | 0.4 ± 0.6 (**<0.01**) |
| Long leg‡ | Entire cohort (60) | 1.4 ± 1.4 | 1.5 ± 1.2 | 1.4 ± 1.1 | −0.1 ± 1.0 (0.50) | 0 ± 1.1 (0.99) | 0.1 ± 0.4 (0.07) |
| | Females (31) | 0.9 ± 0.7 | 1.3 ± 1.0 | 1.3 ± 1.1 | −0.4 ± 1.0 (**0.03**) | −0.4 ± 1.1 (**0.03**) | 0 ± 0.3 (0.72) |
| | Males (29) | 1.9 ± 1.7 | 1.6 ± 1.4 | 1.4 ± 1.2 | 0.3 ± 1.0 (0.17) | 0.5 ± 1 (**0.01**) | 0.2 ± 0.4 (**0.01**) |
| LLD§ | Entire cohort (60) | 0.8 ± 0.6 | 1.1 ± 0.7 | 1.0 ± 0.7 | −0.2 ± 0.7 (**0.02**) | −0.1 ± 0.7 (0.15) | 0.1 ± 0.4 (0.14) |
| | Females (31) | 0.8 ± 0.5 | 1.0 ± 0.7 | 1.1 ± 0.7 | −0.3 ± 0.7 (**0.04**) | −0.3 ± 0.8 (**0.03**) | −0.1 ± 0.3 (0.37) |
| | Males (29) | 1.0 ± 0.7 | 1.1 ± 0.8 | 0.9 ± 0.6 | −0.2 ± 0.7 (0.23) | 0.1 ± 0.6 (0.59) | 0.2 ± 0.5 (**0.02**) |

*W-M = White-Menelaus formulae, G-P = Greulich and Pyle atlas, mFels = modified Fels, abFels = abbreviated modified Fels. †The mean difference ± SD are reported as prediction error (G-P) – prediction error (mFels), prediction error (G-P) – prediction error (abFels), and prediction error (mFels) – prediction error (abFels). P values are derived from paired t tests. ‡At maturity, with the calculated epiphysiodesis effect. §LLD = residual leg-length discrepancy at maturity.

the 60-subject cohort for SL, LL, and LLD. SL differences of >1.0 cm were noted in 38 patients using G-P skeletal age (range, 1.1 to 7.1 cm) compared with 45 (range, 1.1 to 6.4 cm) using mFels. LL differences of >1.0 cm were noted in 25 patients using G-P (range, 1.2 to 6.9 cm) compared with 30 (range, 1.1 to 6.4 cm) using mFels. LLD predictions >1.0 cm from observed values were noted for 18 patients using G-P (range, 1.1 to 2.8 cm) compared with 26 (range, 1.1 to 2.5) using mFels. G-P skeletal age was >1.0 cm more accurate than mFels skeletal age in SL prediction for 19 patients (12 girls), in LL prediction for 7 patients (all girls), and in LLD prediction for 7 patients (4 girls). Modified Fels skeletal age was >1.0 cm more accurate than G-P skeletal age in SL prediction for 9 patients (2 girls), in LL prediction for 6 patients (1 girl), and in LLD prediction for 1 patient (a boy). Overall, prediction errors with the abFels system were equal to or slightly less than those with the mFels system, apart from LLD in females. For LLD in females, the error with abFels was slightly greater than with mFels; however, this difference was not significant.

## Discussion

Accurate and reproducible estimation of skeletal age remains a mainstay for the prediction of LLD and the epiphysiodesis effect in epiphysiodesis-aged patients. It has been reported previously that the W-M formulae combined with skeletal age as determined by the G-P atlas are superior to other prediction methods and to chronological age[6]. In the current study, we sought to determine whether the mFels system had acceptable interobserver reliability, and whether skeletal age as determined by the mFels or abFels system could supplant G-P skeletal age in prediction of leg lengths in these epiphysiodesis-aged patients.

We found that skeletal-age estimates with mFels and abFels were clinically comparable with G-P estimates in our cohort. More specifically, estimates for the entire cohort and for males in particular tended to be almost identical. For females, however, average skeletal-age estimates with mFels and abFels ranged from 2 to 6 months younger than with G-P. This may be secondary to the inclusion of chronological age in the mFels and abFels formulae and may reflect differences in modern children versus the historical cohort used to develop the systems[5,22].

When incorporated into the W-M method, both the mFels and abFels skeletal-age estimates provided similar leg-length and segment-length prediction accuracy, with nearly all prediction errors being ≤0.5 cm from G-P-derived estimates. Despite some prediction errors reaching significance, the minimal overall difference across the 3 different methods likely represents clinically comparable results. The notable exception remains SL prediction in females, where the mFels and abFels-derived prediction errors were 0.9 cm greater relative to G-P, nearly reaching our arbitrary 1-cm threshold for clinical importance. While average differences would appear clinically inconsequential, it is important to recognize that individual differences can be substantial. The astute surgeon must remain conscious of this and carefully consider as much growth information as possible when deciding when to proceed with epiphysiodesis in any individual. There were certainly female patients who would have been inadequately managed if we had relied on the mFels or abFels-derived SL predictions in decision-making regarding the proper timing of epiphysiodesis. There is therefore likely a role for optimization of the mFels and abFels systems for females in particular. Nonetheless, both the mFels and abFels systems were otherwise able to provide clinically comparable LLD predictions. The current study therefore suggests that these systems, with some optimization, may have the potential to obviate the need for a separate hand-and-wrist radiograph. In addition to avoiding additional patient radiation and decreasing health-care expenditure, this would also allow timing of surgical intervention based on the segment in question rather than on a separate anatomical region that may theoretically mature at a different rate.

The comparable results obtained by the abFels system relative to the mFels system also represent an encouraging finding, as the abFels is estimated to take approximately 30 seconds for analysis[13]. Based on the sex and femoral-capping (FemK)[11] score for any given patient, abFels determines which additional 1 or 2 parameters are highest-yield and excludes the remaining parameters that do not add value at any given skeletal-maturity range. Relative to G-P, which has been shown in 1 study to require an average of 1.4 minutes for assessment[23], and the mFels system, which has been estimated to require 1 to 2 minutes for assessment[13], this offers potentially substantial time savings. Combined with the availability and rapidity of an application-based skeletal-age assessment tool, the abFels system would therefore likely be attractive to the busy clinician, with the knowledge that more widespread implementation of artificial-intelligence assessment of skeletal maturity may help to alleviate concerns as to time for interpretation.

There are important limitations to our study. The most important is the relatively small cohort of study subjects, and our findings may not be appropriately extrapolated to a larger, diverse population. Before their widespread implementation, there may be a role for optimization of the mFels and abFels systems to capture the skeletal maturity of female patients more accurately. Another limitation is that, although the G-P bone-age atlas is considered the gold standard for skeletal-age assessment, this atlas as well as the mFels and abFels systems are all based on the same Bolton-Brush dataset[24], which was compiled from a cohort of healthy, predominantly Caucasian children of an elevated socioeconomic standing from 1931 to 1942. With increasing evidence that modern adolescents have more advanced skeletal age and are developing secondary sexual characteristics earlier than prior generations[5,22,25-27], the relevance of this dataset has been called into question[2] and there may be a role for optimization of the mFels and abFels systems in a more contemporary patient population, particularly in females.

Inherent to the inexact science that prediction of leg length represents, another limitation is that there is certainly a degree of error associated with the W-M method. As this method has been shown to be simple and accurate relative to other prediction methods[6], we feel that it is reasonable to use this technique to answer the current clinical question regarding reliability and accuracy. It is important to note that individual case comparisons identify both statistically significant and clinically important differences in leg-length predictions by any method. The treating surgeon must always use caution and as many growth data as are available in determining optimum timing of epiphysiodesis to manage LLD.

The results of this study provide a step toward validation of the mFels and abFels systems relative to the current gold-standard G-P atlas both in estimation of skeletal maturity and in prediction of leg length at skeletal maturity. Although there is certainly a role for further study and optimization, these systems appear to provide clinically comparable SL, LL, and residual LLD predictions and may eventually obviate the need for a separate left hand-and-wrist radiograph and the associated increase in patient radiation and health-care cost. Via a user-friendly and free application, the mFels and abFels systems appear to provide reliable, accurate, and rapid estimations of skeletal maturity and therefore appear to be potentially useful tools in the management of LLD. ■

Dylan G. Kluck, MD[1]
Marina R. Makarov, MD[2]
David A. Podeszwa, MD[2]
Ryan J. Furdock, MD[3]
Raymond W. Liu, MD[4]
Chan-Hee Jo, PhD[2]
John G. Birch, MD, FRCS(C)[2]

[1]Department of Orthopaedic Surgery and Sports Medicine, Shriners Children's Lexington, University of Kentucky, Lexington, Kentucky

[2]Texas Scottish Rite Hospital for Children, Dallas, Texas

[3]Department of Orthopedics, University Hospital Cleveland Medical Center, Cleveland, Ohio

[4]Rainbow Babies and Children's Hospital, Cleveland, Ohio

Email for corresponding author: dylan.kluck@shrinenet.org

## References

**1.** Greulich WW, Pyle SI. Radiographic Atlas of Skeletal Development of the Hand and Wrist. 2nd ed. Palo Alto: Stanford University Press; 1959.

**2.** Alshamrani K, Messina F, Offiah AC. Is the Greulich and Pyle atlas applicable to all ethnicities? A systematic review and meta-analysis. Eur Radiol. 2019 Jun;29(6):2910-23.

**3.** Satoh M. Bone age: assessment methods and clinical applications. Clin Pediatr Endocrinol. 2015 Oct;24(4):143-52.

**4.** Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. Bone age assessment: a large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods. Arch Dis Child. 1999 Aug;81(2):172-3.

**5.** Furdock RJ, Kuo A, Chen KJ, Benedick AJ, Liu RW. Applicability of the Modified Fels and Optimized Oxford Skeletal Maturity Estimation Systems to the Modern Pediatric Population. J Pediatr Orthop. 2023 Mar 1;43(3):e254-9.

**6.** Makarov MR, Jackson TJ, Smith CM, Jo CH, Birch JG. Timing of Epiphysiodesis to Correct Leg-Length Discrepancy: A Comparison of Prediction Methods. J Bone Joint Surg Am. 2018 Jul 18;100(14):1217-22.

**7.** Roche AF. Growth, Maturation, and Body Composition: The Fels Longitudinal Study 1929–1991. Cambridge: Cambridge University Press; 1992.

**8.** Roche AF, Wainer H, Thissen D. Skeletal maturity: the knee joint as a biological indicator. New York: Plenum Medical Book; 1975.

**9.** Pyle SI, Hoerr NL. Radiographic Atlas of Skeletal Development of the Knee: A Standard of Reference. Springfield: Charles C Thomas; 1955.

**10.** Benedick A, Knapik DM, Duren DL, Sanders JO, Cooperman DR, Lin FC, Liu RW. Systematic isolation of key parameters for estimating skeletal maturity on knee radiographs. J Bone Joint Surg Am. 2021 May 5;103(9):795-802.

**11.** Furdock RJ, Cho E, Benedick AJ, Yu J, Sattar A, Liu RW. The utility of the modified Fels knee skeletal maturity system in limb length prediction. J Pediatr Orthop. 2022 Jul 1;42(6):327-34.

**12.** Castillo Tafur JC, Benedick A, Knapik DM, Janes JL, Delozier SJ, Liu RW. Skeletal Maturity Using Knee X-rays: Understanding the Resilience of 7 Radiographic Parameters to Rotational Position. J Pediatr Orthop. 2021 Oct 1;41(9):e733-8.

**13.** Yuan JT, Furdock RJ, Benedick A, Liu RW. Estimating Skeletal Maturity by Segmented Linear Modeling of Key AP Knee Radiographic Parameters. J Pediatr Orthop. 2022 Mar 1;42(3):169-73.

**14.** Anderson M, Green WT, Messner MB. Growth and predictions of growth in the lower extremities. J Bone Joint Surg Am. 1963 Jan;45-A:1-14.

**15.** Anderson M, Messner MB, Green WT. Distribution of lengths of the normal femur and tibia in children from one to eighteen years of age. J Bone Joint Surg Am. 1964 Sep;46:1197-202.

**16.** Beumer A, Lampe HI, Swierstra BA, Diepstraten AF, Mulder PG. The straight line graph in limb length inequality. A new design based on 182 Dutch children. Acta Orthop Scand. 1997 Aug;68(4):355-60.

**17.** Paley D, Bhave A, Herzenberg JE, Bowen JR. Multiplier method for predicting limb-length discrepancy. J Bone Joint Surg Am. 2000 Oct;82(10):1432-46.

**18.** White JW, Stubbins SG. Growth arrest for equalizing leg lengths. JAMA. 1944;126(18):1146-9.

**19.** Menelaus MB. Correction of leg length discrepancy by epiphysial arrest. J Bone Joint Surg Br. 1966 May;48(2):336-9.

**20.** Westh RN, Menelaus MB. A simple calculation for the timing of epiphysial arrest: a further report. J Bone Joint Surg Br. 1981 Feb;63-B(1):117-9.

**21.** Kluck DG, Makarov MR, Kanaan Y, Jo CH, Birch JG. Comparison of "Human" and Artificial Intelligence Hand-and-Wrist Skeletal Age Estimation in an Epiphysiodesis Cohort. J Bone Joint Surg Am. 2023 Feb 1;105(3):202-6.

**22.** Boeyer ME, Sherwood RJ, Deroche CB, Duren DL. Early Maturity as the New Normal: A Century-long Study of Bone Age. Clin Orthop Relat Res. 2018 Nov;476(11):2112-22.

**23.** King DG, Steventon DM, O'Sullivan MP, Cook AM, Hornsby VP, Jefferson IG, King PR. Reproducibility of bone ages when performed by radiology registrars: an audit of Tanner and Whitehouse II versus Greulich and Pyle methods. Br J Radiol. 1994 Sep;67(801):848-51.

**24.** Sanders JO, Qiu X, Lu X, Duren DL, Liu RW, Dang D, Menendez ME, Hans SD, Weber DR, Cooperman DR. The Uniform Pattern of Growth and Skeletal Maturation during the Human Adolescent Growth Spurt. Sci Rep. 2017 Dec 1;7(1):16705.

**25.** Herman-Giddens ME, Slora EJ, Wasserman RC, Bourdony CJ, Bhapkar MV, Koch GG, Hasemeier CM. Secondary sexual characteristics and menses in young girls seen in office practice: a study from the Pediatric Research in Office Settings network. Pediatrics. 1997 Apr;99(4):505-12.

**26.** Herman-Giddens ME, Steffes J, Harris D, Slora E, Hussey M, Dowshen SA, Wasserman R, Serwint JR, Smitherman L, Reiter EO. Secondary sexual characteristics in boys: data from the Pediatric Research in Office Settings Network. Pediatrics. 2012 Nov;130(5):e1058-68.

**27.** Kaplowitz PB, Slora EJ, Wasserman RC, Pedlow SE, Herman-Giddens ME. Earlier onset of puberty in girls: relation to increased body mass index and race. Pediatrics. 2001 Aug;108(2):347-53