

# Reliability Analysis of the Distal Radius and Ulna Classification for Assessing Skeletal Maturity for Patients with Adolescent Idiopathic Scoliosis

Jason Pui Yin Cheung<sup>1</sup> Dino Samartzis<sup>1</sup> Prudence Wing Hang Cheung<sup>1</sup> Kenneth M. C. Cheung<sup>1</sup>  
Keith D. K. Luk<sup>1</sup>

<sup>1</sup>Department of Orthopaedics and Traumatology, The University of Hong Kong, Pokfulam, Hong Kong, SAR, China

Global Spine J 2016;6:164–168.

Address for correspondence Keith D. K. Luk, MCh (Orth), Department of Orthopaedics and Traumatology, Queen Mary Hospital, 5th Floor, Professorial Block, 102 Pokfulam Road, Pokfulam, Hong Kong, SAR, China (e-mail: hrmoldk@hku.hk).

## Abstract

**Study Design** Prospective radiographic study.

**Objective** To test the reliability of the Distal Radius and Ulna Classification (DRU).

**Methods** This single-center study included prospectively recruited subjects with adolescent idiopathic scoliosis managed with bracing. The left-hand radiographs were measured using the DRU classification by two examiners. Intra- and interobserver reliability analysis were performed using intraclass correlation (ICC) analysis.

**Results** From these clinics, 161 patients (124 females and 37 males) with left-hand radiographs were included in the study. The mean age was 13.3 years (standard deviation: 1.5). There was excellent intra- (ICC: 0.93 to 0.95) and interobserver (ICC: 0.97) reliability.

**Conclusions** The DRU classification scheme has been shown to be accurate in determining the peak growth phase and growth cessation. It has now been confirmed to be a reliable tool. Future prospective studies should be performed to investigate its application in deciding when to apply bracing or operative treatment.

## Keywords

- ▶ scoliosis
- ▶ spine
- ▶ classification
- ▶ skeletal maturity
- ▶ radius
- ▶ ulna

## Introduction

The prediction of a patient's growth potential is of clinical significance because it determines the optimal timing of starting or stopping bracing and the timing of growth-guidance surgery with an expandable prosthesis.<sup>1–4</sup> Many clinical and radiologic methods are available to predict a patient's growth potential including the age at menarche, difference in body height growth and arm span, Risser sign, and Tanner and Whitehouse (TW3) digital skeletal age.<sup>5–8</sup> However, most of these measures have disadvantages in the current clinical practice.<sup>9</sup> Clinical measurements of standing height and arm span require serial measurements to determine growth trends, and its retrospective nature thus likely misses the period of peak growth spurt. The menarche age is not consistent in determining the termination of peak growth as some patients develop delayed menarche and its presence marks a stage too late for any meaningful

bracing.<sup>10</sup> Similar to menarche, the Risser sign has been shown by several studies to be a poor indicator of skeletal maturity, cessation of growth, and more importantly the risk of curve progression.<sup>11–15</sup> Up to 75.2% of patients may have persistent growth despite capping of the iliac apophysis.<sup>13</sup> In addition, the progression of iliac apophysis has been shown to proceed in a reverse direction (posteromedially to anterolaterally) or in fragments.<sup>15</sup> The digital skeletal age assessment using the TW3<sup>6–8</sup> or the Greulich and Pyle<sup>16</sup> methods are more accurate in determining different growth phases, but these techniques are complex and far too difficult to apply in the clinical setting. Sanders et al showed that according to the TW3, the finger epiphysis maturation is closely related to the peak height velocity.<sup>17</sup> The peak height velocity is noted to occur during early adolescence, and menarche and the Risser sign appear after the peak growth spurt.<sup>18</sup>

received  
February 19, 2015  
accepted after revision  
May 22, 2015  
published online  
July 9, 2015

DOI <http://dx.doi.org/10.1055/s-0035-1557142>.  
ISSN 2192-5682.

© 2016 Georg Thieme Verlag KG  
Stuttgart · New York

License terms



In response to the limitations possessed by the previous methods in assessing skeletal maturity, the Distal Radius and Ulna (DRU) classification was created and reported by Luk et al.<sup>19</sup> This classification, which included 11 radius grades (R1 to R11) and 9 ulna grades (U1 to U9), was found to accurately determine the peak height velocity (R7 and U5) and cessation of growth (R10 and U9). Menarche occurs one to two stages after the peak height velocity. However, the reliability of the classification has not been assessed. As such, the following study aimed to test the classification's reliability.

## Materials and Methods

A prospective radiographic study assessing patients with adolescent idiopathic scoliosis (AIS) undergoing bracing during the month of July 2014 was performed to assess the reliability of the DRU classification.<sup>19</sup> Ethics approval was obtained from our institutional review board. The study was performed at the Duchess of Kent Children's Hospital, Pokfulam, Hong Kong, a tertiary referral center for spinal disorders. For the DRU classification to be practical clinically, the reliability testing should be based on a wide range of patient ages, as commonly seen in our clinics. This group of early adolescents was chosen to reflect this range of ages. In addition, the aim of this classification was to correctly identify patients in the peak growth spurt so that timely interventions can be applied. Thus, choosing braced patients would likely encompass a larger range of grades and avoid reliability assessment of only a certain age group or DRU grade. This group of subjects would not be specific to gender either.

The follow-up left-hand radiographs during the study month were obtained in all subjects, and data was extracted regarding the age at which the radiograph was taken and the sex of the patient. These anteroposterior radiographs of the hand and wrist were taken with 42 peak kilovoltage and 1.6 mA-seconds of X-ray energy. All patient information was blinded to the two examiners (J.P.Y.C. and K.D.-K.L.), one junior and one senior consultant spine surgeon, for grading. There was no discussion between the two examiners about how to classify each grade prior to the initiation or during the study. Radiographs were accessed by a DICOM-based Radworks 5.1 (Applicare Medical Imaging BV, Zeist, The Netherlands) computer software program. Intra- and inter-observer reliability was assessed for all subjects in this group. All intraobserver reliability assessments were performed at least 1 month apart.

## Statistical Analysis

All data was coded and entered on a spreadsheet and was blinded to the examiners until the end of reliability measurements. SPSS version 20 (Chicago, Illinois, United States) was utilized to perform the statistical analysis. Descriptive and frequency statistics were performed with the data. The reliability assessment was based on intraclass correlation (ICC), which was an appropriate statistical tool for this analysis.<sup>20</sup> ICC could be interpreted based on the following

$\alpha$  values: 0 to 0.29 indicated poor agreement; 0.30 to 0.49 indicated fair agreement; 0.50 to 0.69 indicated moderate agreement; 0.70 to 0.80 indicated strong agreement; and  $>0.80$  indicated almost perfect agreement.<sup>21,22</sup> The 95% confidence interval (CI) bounds were assessed for precision. A  $p$  value of  $<0.05$  was considered significant.

## Results

One hundred sixty-one subjects (124 females and 37 males) satisfied our inclusion criteria and were recruited for this study. The mean age was 13.3 (standard deviation: 1.5; range: 8 to 18). The spread of grades measured were from R4 to R11 and U1 to U9. Excellent reliability was found through our analysis. The interobserver reliability was ICC = 0.97 (95% CI: 0.96 to 0.98) and ICC = 0.97 (95% CI: 0.96 to 0.98) for the radius and ulna, respectively. The intraobserver reliability for J.P.Y.C. was ICC = 0.93 (95% CI: 0.91 to 0.95) for the radius and ICC = 0.95 (95% CI: 0.93–0.96) for the ulna; for K.D.-K.L., ICC = 0.94 (95% CI: 0.91 to 0.95) for the radius and ICC = 0.95 (95% CI: 0.93 to 0.96) for the ulna. All  $p$  values were  $<0.001$ . The frequency of disagreement was also assessed.

## Interobserver Disagreement

For the radius, there were 18 one-grade disagreements (11.1%) and 1 two-grade disagreement (0.6%). For the ulna, there were 39 one-grade disagreements (24.1%) and 1 two-grade disagreement (0.6%). The intraobserver disagreement for J.P.Y.C. was as follows: For the radius, there were 28 one-grade disagreements (17.3%) and 5 two-grade disagreements (3.1%). For the ulna, there were 42 one-grade disagreements (25.9%) and 2 two-grade disagreements (1.2%). The intraobserver disagreement for K.D.-K.L. was as follows: For the radius, there were 26 one-grade disagreements (16.0%) and 4 two-grade disagreements (2.5%). For the ulna, there were 40 one-grade disagreements (24.7%) and 5 two-grade disagreements (3.1%). There were no disagreements more than two grades. All disagreements of two grades were discussed until a consensus was reached.

## Discussion

From this study, we found that the DRU classification had excellent reliability among a wide range of radius and ulna grades.<sup>19</sup> Although we did not have patients young enough to present with radius stages R1 to R4, interventions for AIS such as bracing are only applicable during the peak growth spurt. The R6 to R9 and U5 to U8 were important grades because they represented the period of peak growth spurt, and these grades were well represented in our testing.

After all measurements were made through this vigorous process of testing, the two examiners thoroughly discussed each grade description and were able to form a consensus about which aspects of the classification were more representative and easier to apply during reading. The difficulties encountered by both readers during testing mainly arose from the grading of R7 to R9 and U5 to U8. Both reviewers agreed that the grading of R7 to R9 and U5 to U8 were the



**Fig. 1** Radiograph of the distal radius and ulna showing the R7 grade. Examiners should focus on the medial side capping and the absence of lateral side capping (red arrow). Note the hooklike structure/sharp outgrowth at the medial physis facing the metaphysis that deviates from the physal line. This radiograph also represents the U6 grade with the appearance of an epiphysis and metaphysis of the same width but no narrowing of the medial physal plate (white arrow).

most difficult because they were based on descriptions of multiple parameters. Each clinician might focus on something different, such as the appearance of the physis narrowing, the definition of capping, and when the epiphysis acquired the same width as the metaphysis, among others. Disagreements in the ulna grades appear larger than the radius grades, likely due to a larger range of ulna grades measured as compared with the radius.

For the radius, the main identifying factor between R7 and R8 was the appearance of capping over the medial side in R7 (→**Fig. 1**) and over the lateral side in R8 (→**Fig. 2**). Both readers agreed that capping was defined as a hornlike structure at the ends of the epiphysis like a sharp bony outgrowth. A simple step to confirm this observation on the radius was to trace the proximal border of the epiphysis to find any dipping of the line toward the metaphysis at the medial or lateral ends. For R9 (→**Fig. 3**), strong capping was not always a consistent finding, thus narrowing of the physis was the main determinant description for this grade.

Generally, grading of U1 to U4 was not difficult due to some distinct features such as the size of the epiphysis for U3 and the appearance of the styloid for U4. The difficulties were noted for descriptions in U5 to U8. For U5 (→**Fig. 4**), the appearance of a denser ulna head was not easily observed and so flattening of the radial half of the ulnar epiphysis was the main determinant. For U6 (→**Fig. 1**), we found that the overlapping of the metaphysis with the epiphysis at the center third was inconsistent in most radiographs, and



**Fig. 2** Radiograph of the distal radius and ulna showing the R8 grade. Examiners should focus on the lateral side capping (red arrow). Note the hooklike structure/sharp outgrowth at the lateral physis facing the metaphysis that deviates from the physal line.



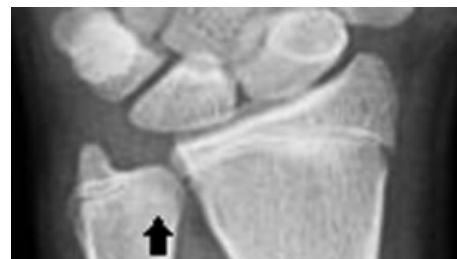
**Fig. 3** Radiograph of the distal radius and ulna showing the R9 grade. Examiners should focus on the narrowing of the physis. In this radiograph, the ulna is graded U7 as there is narrowing of the medial physal plate (black arrow) but incomplete fusion.



**Fig. 4** Radiograph of the distal radius and ulna showing the U5 grade. Examiners should focus on the flattening of the radial epiphysis (white arrow) and should note that the epiphysis is not as wide as the metaphysis, indicated by the vertical white line touching the metaphysis but not the epiphysis.

thus we often relied on the appearance of the epiphysis being as wide as the metaphysis. To differentiate U7 (→**Fig. 3**) from U8 (→**Fig. 5**), we usually used narrowing of the medial ulnar physis (U7) and fusion of the medial ulnar physis (U8) as the determining factor. The descriptions of the smooth curve line articulation in U7 might not be very easily understood.

Interestingly, the interobserver reliability was somewhat better than the intraobserver reliability, which suggests that the readers may become more competent in the measuring skills with increased grading experience from measuring these 161 radiographs, demonstrating a learning curve effect. Regardless of the issues raised previously, excellent reliability was obtained. Hence, this classification scheme should be useful as a guide for decision making and as a communication tool between clinicians. Despite not having any prior



**Fig. 5** Radiograph of the distal radius and ulna showing the U8 grade. Examiners should focus on fusion of the medial physis (black arrow).

discussions or training into how each grade should be correctly assessed, the results showed that both examiners used similar grading methods in approaching each radiograph. Thus, the difficulties and suggestions raised during the consensus meeting were important to help less experienced examiners with this classification scheme.

The DRU has been shown to be strongly associated with the peak growth phase and growth cessation,<sup>19</sup> which can help clinicians decide between bracing and surgery. The results from this study suggested that the DRU classification was a reliable method to assess skeletal growth. Although the most appropriate time to get these radiographs still requires further investigation, the authors suggest a yearly radiograph with consideration of radiographs every 6 to 9 months during the period of peak growth spurt. In the future, prospective studies are being developed to apply this classification scheme on patients with AIS to determine its sensitivity and specificity in identifying the peak growth spurt, predicting curve deterioration, as well as detecting skeletal maturity. In addition, it can be used to help decide the appropriate timing for initiation and cessation of bracing, when to use growth guidance surgery, and when to recommend fusion. By more accurately predicting a child's remaining growth potential, we can potentially narrow the duration of effective bracing without jeopardizing the deformity control, thereby reducing any unnecessary psychological burden.<sup>23</sup> This knowledge could also guide the duration and interval of follow-up and the number of X-rays needed, thus reducing radiation exposure. Additional studies are needed to validate this DRU classification in other ethnicities and populations and to further directly compare it with other new and established skeletal maturity schemes.<sup>24-27</sup>

This study has shown that the DRU classification is a reliable tool for measuring skeletal growth. Further prospective studies are required to test its predictability of peak growth spurt and growth cessation and to compare it with other skeletal growth measures such as the Risser sign, TW3 grading, and age of menarche.

#### Disclosures

Jason Pui Yin Cheung, none

Dino Samartzis, none

Prudence Wing Hang Cheung, none

Kenneth M. C. Cheung, none

Keith D. K. Luk, none

#### References

- Akbarnia BA, Mundis GM Jr, Salari P, Yaszay B, Pawelek JB. Innovation in growing rod technique: a study of safety and efficacy of a magnetically controlled growing rod in a porcine model. *Spine (Phila Pa 1976)* 2012;37(13):1109-1114
- Akbarnia BA, Cheung K, Noordeen H, et al. Next generation of growth-sparing techniques: preliminary clinical results of a magnetically controlled growing rod in 14 patients with early-onset scoliosis. *Spine (Phila Pa 1976)* 2013;38(8):665-670
- Cheung JP, Samartzis D, Cheung KM. A novel approach to gradual correction of severe spinal deformity in a pediatric patient using the magnetically-controlled growing rod. *Spine J* 2014;14(7):e7-e13
- Cheung KM, Cheung JP, Samartzis D, et al. Magnetically controlled growing rods for severe spinal curvature in young children: a prospective case series. *Lancet* 2012;379(9830):1967-1974
- Risser JC. The classic: the iliac apophysis: an invaluable sign in the management of scoliosis. 1958. *Clin Orthop Relat Res* 2010;468(3):643-653
- Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Arch Dis Child* 1976;51(3):170-179
- Tanner JM, Whitehouse RH, Hughes PC, Carter BS. Relative importance of growth hormone and sex steroids for the growth at puberty of trunk length, limb length, and muscle width in growth hormone-deficient children. *J Pediatr* 1976;89(6):1000-1008
- Tanner JM, Whitehouse RH, Marubini E, Resele LF. The adolescent growth spurt of boys and girls of the Harpenden growth study. *Ann Hum Biol* 1976;3(2):109-126
- Little DG, Song KM, Katz D, Herring JA. Relationship of peak height velocity to other maturity indicators in idiopathic scoliosis in girls. *J Bone Joint Surg Am* 2000;82(5):685-693
- Chang SH, Tzeng SJ, Cheng JY, Chie WC. Height and weight change across menarche of schoolgirls with early menarche. *Arch Pediatr Adolesc Med* 2000;154(9):880-884
- Biondi J, Weiner DS, Bethem D, Reed JF III. Correlation of Risser sign and bone age determination in adolescent idiopathic scoliosis. *J Pediatr Orthop* 1985;5(6):697-701
- Dhar S, Dangerfield PH, Dorgan JC, Klenerman L. Correlation between bone age and Risser's sign in adolescent idiopathic scoliosis. *Spine (Phila Pa 1976)* 1993;18(1):14-19
- Hoppenfeld S, Lonner B, Murthy V, Gu Y. The rib epiphysis and other growth centers as indicators of the end of spinal growth. *Spine (Phila Pa 1976)* 2004;29(1):47-50
- Little DG, Sussman MD. The Risser sign: a critical analysis. *J Pediatr Orthop* 1994;14(5):569-575
- Shuren N, Kasser JR, Emans JB, Rand F. Reevaluation of the use of the Risser sign in idiopathic scoliosis. *Spine (Phila Pa 1976)* 1992;17(3):359-361
- Greulich WW. *Radiographic Atlas of Skeletal Development of the Hand and Wrist*. Stanford, CA: Stanford University Press; 1959
- Sanders JO, Browne RH, McConnell SJ, Margraf SA, Cooney TE, Finegold DN. Maturity assessment and curve progression in girls with idiopathic scoliosis. *J Bone Joint Surg Am* 2007;89(1):64-73
- Risser JC. The Iliac apophysis; an invaluable sign in the management of scoliosis. *Clin Orthop* 1958;11(11):111-119
- Luk KD, Saw LB, Grozman S, Cheung KM, Samartzis D. Assessment of skeletal maturity in scoliosis patients to determine clinical management: a new classification scheme using distal radius and ulna radiographs. *Spine J* 2014;14(2):315-325
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307-310
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174
- Vangeneugden T, Laenen A, Geys H, Renard D, Molenberghs G. Applying concepts of generalizability theory on clinical trial data to investigate sources of variation and their impact on reliability. *Biometrics* 2005;61(1):295-304
- Chan SL, Cheung KM, Luk KD, Wong KW, Wong MS. A correlation study between in-brace correction, compliance to spinal orthosis and health-related quality of life of patients with adolescent idiopathic scoliosis. *Scoliosis* 2014;9(1):1

- 24 Diméglio A, Charles YP, Daures JP, de Rosa V, Kaboré B. Accuracy of the Sauvegrain method in determining skeletal age during puberty. *J Bone Joint Surg Am* 2005;87(8):1689–1696
- 25 Nicholson AD, Liu RW, Sanders JO, Cooperman DR. Relationship of calcaneal and iliac apophyseal ossification to peak height velocity timing in children. *J Bone Joint Surg Am* 2015;97(2):147–154
- 26 Sauvegrain J, Nahum H, Carle F. [Bone maturation. Importance of the determination of the bone age. Methods of evaluation (general review)]. *Ann Radiol (Paris)* 1962;5:535–541
- 27 Verma K, Sitoula P, Gabos P, et al. Simplified skeletal maturity scoring system: learning curve and methods to improve reliability. *Spine (Phila Pa 1976)* 2014;39(26):E1592–E1598