

Reproducibility and reliability analysis of the Luk Distal Radius and Ulna Classification for European patients with adolescent idiopathic scoliosis

James Houston¹
Amy Chiang²
Shahnawaz Haleem³
Jason Bernard¹
Timothy Bishop¹
Darren F. Lui¹

Abstract

Purpose Current clinical and radiological methods of predicting a patient's growth potential are limited in terms of practicality, accuracy and known to differ in different races. This information influences optimal timing of bracing and surgical intervention in adolescent idiopathic scoliosis (AIS). The Luk classification was developed to mitigate limitations of existing tools. Few reliability studies are available and are limited to certain geographical regions with varying results. This study was performed to analyze reproducibility and reliability of the Luk Distal Radius and Ulna Classification in European patients.

Methods This is a radiological study of 50 randomly selected left hand and wrist radiographs of patients with AIS referred to a tertiary referral centre. They were assessed for bone maturity using the Luk Distal Radius and Ulna Classification. Assessment was performed twice by four examiners at an interval of one month. Statistical analysis was performed using the intraclass correlation (ICC) method to determine the reliabilities within and between the examiners.

Results In total, 50 radiographs (M:F = 13:37) with a mean age of 13.7 years (10 to 18) were assessed for reliability. The inter-rater ICC value was 0.918 for radius assessment and 0.939 for ulna assessment. The intra-rater ICC values for radius assessment ranged between 0.897 and 0.769 and be-

tween 0.948 and 0.786 for ulna assessment. There was near perfect correlation for both assessments.

Conclusion This study provides independent evidence that the Luk Distal Radius and Ulna Classification is a reliable tool for assessment of skeletal maturity for European patients. Minimal clinical experience is required to reliably utilize it.

Level of evidence: IV

Cite this article: Houston J, Chiang A, Haleem S, Bernard J, Bishop T, Lui DF. Reproducibility and reliability analysis of the Luk Distal Radius and Ulna Classification for European patients with adolescent idiopathic scoliosis. *J Child Orthop* 2021;15:166-170. DOI: 10.1302/1863-2548.15.200251

Keywords: adolescent idiopathic scoliosis; Luk classification; growth velocity; bone age; reliability

Introduction

Bone maturity assessment underpins the management of all types of scoliosis but is especially useful during the peak growth spurt just prior to cessation of growth.¹ Understanding the intricacies of bone maturation in this adolescent growth spurt period is invaluable for decision-making with regards to the use of a brace, employing growing rod techniques or proceeding to fusion.²⁻⁵ Correct use of a brace can mitigate the need for surgery and the development of accurate, practical and reliable tools is paramount to optimizing patient experience and outcomes.²

Multiple methods have been used to determine skeletal maturity that include clinical history, physical examination and imaging. Clinical measurements include recording height, spinal length, arm span and foot size.⁶⁻⁸ While simple to perform, these rely on differences between serial recordings to gauge growth spurt. Multiple recordings may not be efficient, their entry not sensitive, may differ in different races and by the time the change has been noted, it may already be too late to intervene.⁹⁻¹¹ Secondary sexual characteristics, such as the timing of menarche can also be used but this is only applicable to female patients, and the peak growth spurt may have already passed by its occurrence.¹²

Given that there is considerable variety in patient size, even within a particular stage of skeletal maturity, various

¹ St Georges University Hospitals NHS Foundation Trust, London, UK

² St George's Hospital Medical School, University of London, London, UK

³ Royal Orthopaedic Hospital, The Woodlands, Bristol Rd South, Birmingham, UK

Correspondence should be sent to Shahnawaz Haleem, Royal Orthopaedic Hospital, The Woodlands, Bristol Road South, Northfield, Birmingham, B31 2AP, UK.
E-mail: shahnawaz.haleem@gmail.com

radiographic assessment tools have been developed. The Risser grading system is a simple and commonly used system that measures the iliac crest apophysis, however, it has been noted to grade peak growth poorly.¹³⁻¹⁶ Other areas of the body used in bone maturity assessments include the proximal humerus, elbow and the fusion of the rib heads and ring epiphysis.¹⁷⁻¹⁹

Radiographs of the hand and wrist are also commonly used to assess age or grade maturity. These include the Tanner and Whitehouse (TW), Greulich and Pyle (GP) and Sanders classifications.^{12,20-23} The Tanner-Whitehouse III (TW3) classification established a relationship between bone maturity and bone age and scores each bone based on maturity with the total weighted score converted to a bone age.²⁴ The GP classification is based upon 'Atlas-Matching' and is less objective than the scoring method of TW.²³ Moreover, the GP, while reliable in a developed Caucasian population was found to be unreliable in other ethnic populations.²⁵⁻³¹ Studies comparing both methods considered TW to be more reliable but TW takes longer to grade, taking on average almost 8 minutes.^{32,33} While both scoring systems are accurate, they are complex and time consuming to use.³² This practically limits their use in a busy clinical setting. The Sanders or simplified TW3 classification (sTW3) removed the radius and ulna assessment and evaluated only Digital Skeletal Age.¹² The authors argued that even with removal of the radius and ulna assessment, curve progression could still be predicted accurately, however, given the complexity of the assessment pre-clinical training was still required. A shorthand bone age assessment study attempted to shorten and simplify assessment but is valid only for an age range of 12.5 to 16 years for boys and ten to 14 years for girls, hence limiting its use in the overall scoliosis setting.^{34,35}

Luk et al³⁶ developed the Luk Distal Radial and Ulna Classification (LRDU), due to the limitations of the other skeletal maturity classifications in adolescent idiopathic scoliosis (AIS).³⁶ In contrast to the Sanders classification, the LRDU looks solely at the distal radius and ulna. The original LRDU was simplified following feedback from early adopters removing descriptions of multiple parameters at different stages. The current classification is based on examination of a plain film radiograph of the left wrist. It has 11 radius grades (R1 to R11) and nine ulna grades (U1 to U9). Growth was found to have peak spurt at levels R7 and U5 and halt at levels R10 and U9. Table 1 characterizes the LRDU.

Our study aims to perform a reliability analysis on the Luk Distal Radius and Ulna Classification system at a different geographical location. By performing the study on a set of patients based in the United Kingdom it attempts to ascertain reliability analysis on a different geographical set of patients as previous studies were limited to Chinese and Japanese cohorts.³⁷ The study also uses assessors at different grades of seniority to assess whether there is a differ-

ence in reliability due to experience. By broadly repeating the methodology of the original LRDU study it aims to provide independent verification or rebuttal of reliability of the original classification.

Materials and methods

A radiographic study was performed on 50 randomly selected patients from a prospectively collected database who had presented with AIS to our tertiary referral centre for children with spinal conditions between 07 February 2017 and 29 August 2018. The study was registered within our hospital, but ethics committee approval was not required as the study did not affect patient management.

Left hand radiographs (as per standard agreement) were obtained in all patients.^{23,38} Four examiners were used (DFL, SH, JH, AC), one consultant spine surgeon, one senior fellow, one registrar and one medical student. There was no discussion between the examiners during the study or about the classification. Radiographs were accessed using a picture archiving and communication system (PACS v4.4; Intellispace PACS Enterprise, Philips, Foster City, California). Intraobserver reliability assessments were performed one month apart with data being stored securely on a spreadsheet by AC until the end of reliability measurements.

Statistical analysis

Data was evaluated in descriptive and frequency terms and analyzed using SPSS version 25 (SPSS Inc, Chicago, Illinois). Inter- and intraobserver reliability was evaluated using intraclass correlation coefficients (ICC) within its 95% confidence intervals, as per the original reliability analysis. The ICC was further categorized into standard groups based on alpha values; poor agreement (0 to 0.29), fair agreement (0.30 to 0.49), moderate agreement (0.5 to 0.69), strong agreement (0.7 to 0.8) and near perfect agreement (> 0.8). Further analysis was performed for interobserver disagreement.

Results

In total, 50 patients (M:F = 13:37) with a mean age of 13.7 years (10 to 18) were assessed in the study. The inter-rater ICC values for radius (0.918) and ulna (0.939) assessment showed near perfect agreement as shown in Table 2. The mean intra-rater ICC values for radius was 0.822 (0.769 to 0.897) and ulna was 0.847 (0.786 to 0.948) assessment also showed near perfect agreement as shown in Table 3.

Intraobserver disagreement

For the radius (n = 200) there were a total of 84 (42%) one-grade disagreements and 18 (9%) two-grade dis-

Table 1 Characterization of the Luk Distal Radius and Ulna Classification

Stage	Description
Radius	
R1	Epiphysis appears as single or multiple spots
R2	Distinct and oval-shaped epiphysis
R3	Maximal diameter is more than half the width of the metaphysis
R4	Double line at the distal border of the epiphysis represents the palmar and dorsal surface
R5	Width of epiphysis not as wide as the metaphysis
R6	Epiphysis is as wide as metaphysis. No capping or narrowing of the physis is seen
R7	Epiphysis capping on medial side, but not on the lateral side
R8	Epiphysis capping on both medial and lateral sides. The medial and lateral ends of the physis are wider than the centre
R9	Ossification has begun with blurring of the central physis
R10	The physal line is closed, forming a sclerotic line. A notch is still visible at the medial or the lateral end of the growth plate
R11	Complete fusion of the physis with no notch
Ulna	
U1	The epiphysis appears at single/multiple spots
U2	A round shaped epiphysis
U3	The epiphysis is at least half the width of the metaphysis
U4	The styloid is visible on the medial end of the epiphysis
U5	Epiphysis width up to the metaphyseal width
U6	Medial epiphysis as wide as the metaphysis
U7	Medial physal plate narrowing or fusion
U8	More than half the medial growth plate fused with the unfused part just proximal to the styloid process
U9	Complete fusion of the physis

Table 2 Inter-rater intraclass correlation coefficient (ICC) values for radius and ulna assessment

	Inter-rater ICC value	95% confidence interval
Radius assessment	0.918	0.878 to 0.948
Ulna assessment	0.939	0.908 to 0.962

Table 3 Intra-rater intraclass correlation coefficient values for radius and ulna assessment according to each examiner

	AC	DL	JH	SH	Mean
Radius assessment	0.897	0.809	0.769	0.814	0.822
Ulna assessment	0.948	0.843	0.786	0.810	0.847

agreements. For the ulna (n = 200) there were at total of 62 (31%) one-grade disagreements, 13 (6.5%) two-grade disagreements and two (1%) three-grade disagreements. No discussions were held to discuss disagreements as the intention of the study was to assess reliability and reproducibility at different knowledge and experience levels.

Discussion

This study finds that the LDRU classification had near perfect agreement for interobserver and intraobserver analysis. Our study is broadly representative of the original reliability analysis with similar average age and range of scores.

The authors of the LDRU suggest that the system is simple, practical and provides accurate prediction of a patient's peak growth spurt. One study has suggested that the LDRU is more reliable than the sTW3, and may

also provide more practical indications of when to start weaning off the brace.^{39,40} Bracing was suggested in the initial study to begin at R6 and R7 although Li et al⁴¹ suggested earlier adoption at R5 and more frequent observations around R7 to R9 to begin brace weaning at R10.³⁶ If the benefits of the LDRU are true, it could provide better guidance on whether bracing is appropriate, and when to begin and end bracing for any given patient with AIS.^{2,41,42}

The evidence for reliability in the LDRU is limited, and the key to cementing evidence is reproducibility in research.⁴³ Few reliability analyses have been performed so far, and they are all confined to a particular geographical continent.^{39,41,44} The larger reliability analysis was performed by the same unit which designed the classification, and reported both inter- and intraobserver reliability to be near perfect, leaving the study open to potential criticism for not being independent.^{36,44} The second study gave data on interobserver disagreement which was lower than our findings. None of the independent analyses included boys and results varied between them.

Okuda et al³⁹ found significantly lower reliability in the LDRU in their cohort of Japanese patients compared with ours and the initial study.³⁶ This may be due to the use of a different statistical tool; Kappa values instead of ICC. They also found in contrast to our study that the radius was the more reliable tool for assessment, although they noted that the low reliability in their ulna classification may be due to different amounts of pronation on their radiographs. Li et al⁴¹ have suggested that either the radius or the ulna could be assessed in isolation given there is good correlation between radial and ulna progression, however this was only applicable to older children and knowing that variations in radius and ulna maturity might exist. Their study of 40 physically immature girls in a Chinese population found near perfect reliability.

Our study used four observers, which is a greater number than most of the other reliability analyses. This study also had a greater number of patients evaluated than the other independent analyses. Our study finds that interobserver reliability was better than intraobserver reliability. Similar to the findings of Cheung et al,⁹ this may demonstrate a possible learning curve effect for examiners, although this was short and may not be clinically significant. Our interobserver reliability was high without significant prior experience of using the LDRU. In our study three of the four of the examiners were novices at using the system. There was no apparent increase in reliability based upon the radiograph experience of the examiner. Of interest, the least experienced examiner, a medical student, had the highest intra-rater reliability. This shows that contrary to the sTW3, little to no experience is required of the system and this is also not dependant on experience of the observer.

Our study has some limitations. Similar to the other studies, there were no patients with R1 to R4 grades. This could be due to our referral system which does not involve a school screening programme which is the usual mode of referral for the original authors. Our patients may, therefore, be referred at a later stage in their growing phase. However, bracing decisions are only applicable to patients around their peak growth spurt, which is typically around R7 and U5, and these scores were well represented in our study. The majority of our patients are Caucasian in origin (both in the outpatient and the theatre settings), we have however not specifically analyzed the ethnic origins of our study patients.

Conclusions

This study confirms that the Luk Distal Radius and Ulna Classification system is a reliable system for assessing skeletal maturity. The LDRU system can be used reliably in patients from the United Kingdom as well as in Hong Kong. Little to no experience is required to use the LDRU, which is also not dependant on the clinical experience of the observer. Further research is required to evaluate its reliability in younger children, its sensitivity and specificity, as well as how different grades should influence management decisions.

Received 30 November 2020; accepted after revision 03 February 2021.

COMPLIANCE WITH ETHICAL STANDARDS

FUNDING STATEMENT

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

OA LICENCE TEXT

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) licence (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed.

ETHICAL STATEMENT

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Not required.

ICMJE CONFLICT OF INTEREST STATEMENT

DFL reports personal fees from Stryker and Zimmer Biomet, outside the submitted work.

The other authors declare no conflict of interest relevant to this work.

AUTHOR CONTRIBUTIONS

JH: Discussed framework, Analyzed data, Wrote and critically reviewed article for final acceptance.

AC: Analyzed data, Wrote and critically reviewed article for final acceptance.

SH: Original concept, Discussed framework, Analyzed data, Wrote and critically reviewed article for final acceptance.

JB: Discussed framework, Critically reviewed article for final acceptance.

TB: Discussed framework, Critically reviewed article for final acceptance.

DFL: Original concept, Discussed framework, Analyzed data, Wrote and critically reviewed article for final acceptance.

REFERENCES

1. **Noshchenko A, Hoffecker L, Lindley EM, et al.** Predictors of spine deformity progression in adolescent idiopathic scoliosis: A systematic review with meta-analysis. *World J Orthop* 2015;6:537-558.
2. **Weinstein SL, Dolan LA, Wright JG, Dobbs MB.** Effects of bracing in adolescents with idiopathic scoliosis. *N Engl J Med* 2013;369:1512-1521.
3. **Rahman T, Bowen JR, Takemitsu M, Scott C.** The association between brace compliance and outcome for patients with idiopathic scoliosis. *J Pediatr Orthop* 2005;25:420-422.
4. **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* 2013;38:665-670.
5. **Betz RR, Kim J, D'Andrea LP, et al.** An innovative technique of vertebral body stapling for the treatment of patients with adolescent idiopathic scoliosis: a feasibility, safety, and utility study. *Spine* 2003;28 (suppl):S255-S265.
6. **Dimeglio A.** Growth in pediatric orthopaedics. *J Pediatr Orthop* 2001;21:549-555.
7. **Jarzem PF, Gledhill RB.** Predicting height from arm measurements. *J Pediatr Orthop* 1993;13:761-765.
8. **Busscher I, Kingma I, Wapstra FH, et al.** The value of shoe size for prediction of the timing of the pubertal growth spurt. *Scoliosis* 2011;6:1-7.
9. **Cheung JPY, Samartzis D, Cheung PWH, et al.** The distal radius and ulna classification in assessing skeletal maturity: a simplified scheme and reliability analysis. *J Pediatr Orthop B* 2015;24:546-551.
10. **Zhang A, Sayre JW, Vachon L, Liu BJ, Huang HK.** Racial differences in growth patterns of children assessed on the basis of bone age. *Radiology* 2009;250:228-235.
11. **Ontell FK, Ivanovic M, Ablin DS, Barlow TW.** Bone age in children of diverse ethnicity. *AJR Am J Roentgenol* 1996;167:1395-1398.
12. **Sanders JO, Khoury JG, Kishan S, et al.** Predicting scoliosis progression from skeletal maturity: a simplified classification during adolescence. *J Bone Joint Surg [Am]* 2008;90-A:540-553.
13. **Risser JC.** The Iliac apophysis; an invaluable sign in the management of scoliosis. *Clin Orthop* 1958;11:111-119.
14. **Shuren N, Kasser JR, Emans JB, Rand F.** Reevaluation of the use of the Risser sign in idiopathic scoliosis. *Spine* 1992;17:359-361.
15. **Little DG, Sussman MD.** The Risser sign: a critical analysis. *J Pediatr Orthop* 1994;14:569-575.
16. **Wang WW, Xia CW, Zhu F, et al.** Correlation of Risser sign, radiographs of hand and wrist with the histological grade of iliac crest apophysis in girls with adolescent idiopathic scoliosis. *Spine* 2009;34:1849-1854.
17. **Li DT, Linderman GC, Cui JJ, et al.** The proximal humeral ossification system improves assessment of maturity in patients with scoliosis. *J Bone Joint Surg [Am]* 2019;101-A:1868-1874.

18. **Sauvegrain J, Nahum H, Bronstein H.** Study of bone maturation of the elbow. *Ann Radiol (Paris)* 1962;5:542-550.
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* 2004;29:47-50.
20. **Tanner JM, Whitehouse RH, Takaishi M.** Standards from birth to maturity for height, weight, height velocity, and weight velocity: british children, 1965. I. *Arch Dis Child* 1966;41:454-471.
21. **Tanner JM, Whitehouse RH, Takaishi M.** Standards from birth to maturity for height, weight, height velocity, and weight velocity: british children, 1965. II. *Arch Dis Child* 1966;41:613-635.
22. **Tanner JM, Landt KW, Cameron N, Carter BS, Patel J.** Prediction of adult height from height and bone age in childhood. A new system of equations (TW Mark II) based on a sample including very tall and very short children. *Arch Dis Child* 1983;58:767-776.
23. **Greulich W, Pyle S.** *Radiographic atlas of skeletal development of the hand and wrist.* Redwood City, California: Stanford University Press, 1959.
24. **Tanner J, Healy M, Goldstein H, Cameron N.** *Assessment of skeletal maturity and prediction of adult height(TW3).* Philadelphia, Pennsylvania: Saunders Ltd, 2001.
25. **van Rijn RR, Lequin MH, Robben SG, Hop WC van KC, van Kuijk C.** Is the Greulich and Pyle atlas still valid for Dutch Caucasian children today? *Pediatr Radiol* 2001;31:748-752.
26. **Soudack M, Ben-Shlush A, Jacobson J, et al.** Bone age in the 21st century: is Greulich and Pyle's atlas accurate for Israeli children? *Pediatr Radiol* 2012;42:343-348.
27. **Moradi M, Sirous MMP.** The reliability of skeletal age determination in an Iranian sample using Greulich and Pyle method. *Forensic Sci Int* 2012;223:e1-4.
28. **Patil ST, Parchand MP, Meshram MMKN.** Applicability of Greulich and Pyle skeletal age standards to Indian children. *Forensic Sci Int* 2012;216:e1-4.
29. **Benjavongkulchai S, Pittayapat P.** Age estimation methods using hand and wrist radiographs in a group of contemporary Thais. *Forensic Sci Int* 2018;287:e1-e8.
30. **Mansourvar M, Ismail MA, Raj RG, et al.** The applicability of Greulich and Pyle atlas to assess skeletal age for four ethnic groups. *J Forensic Leg Med* 2014;22:26-29.
31. **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;29:2910-2923.
32. **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;81:172-173.
33. **King DG, Steventon DM, O'Sullivan MP, et al.** 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;67:848-851.
34. **Gerges M, Eng H, Chhina H, Cooper A.** Modernization of bone age assessment: comparing the accuracy and reliability of an artificial intelligence algorithm and shorthand bone age to Greulich and Pyle. *Skeletal Radiol* 2020;49:1449-1457.
35. **Heyworth BE, Osei DA, Fabricant PD, et al.** The shorthand bone age assessment: a simpler alternative to current methods. *J Pediatr Orthop* 2013;33:569-574.
36. **Luk KDK, 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;14:315-325.
37. **Yamamoto Y, Shigematsu H, Cheung PWH, et al.** How do we follow-up patients with adolescent idiopathic scoliosis? Recommendations based on a multicenter study on the distal radius and ulna classification. *Eur Spine J* 2020;29:2064-2074.
38. **Satoh M.** Bone age: assessment methods and clinical applications. *Clin Pediatr Endocrinol* 2015;24:143-152.
39. **Okuda A, Shigematsu H, Fujii H, et al.** Reliability comparison between 'distal radius and ulna' and 'simplified Tanner-Whitehouse III' assessments for patients with adolescent idiopathic scoliosis. *Asian Spine J* 2020;14:280-286.
40. **Cheung JPY, Cheung PWH, Luk KDK.** When should we wean bracing for adolescent idiopathic scoliosis? *Clin Orthop Relat Res.* 2019;477:2145-2157.
41. **Li Y, Mao S, Shi B, et al.** Utilization of distal radius and ulna classification scheme in predicting growth peak and curve progression in idiopathic scoliosis girls undergoing bracing treatment. *Eur Spine J* 2020;29:770-778.
42. **Ylikoski M.** Growth and progression of adolescent idiopathic scoliosis in girls. *J Pediatr Orthop B* 2005;14:320-324.
43. **Siegel V.** Reproducibility in research. *Dis Model Mech* 2011;4:279-280.
44. **Cheung JPY, Samartzis D, Cheung PWH, Cheung KM, Luk KD.** Reliability analysis of the distal radius and ulna classification for assessing skeletal maturity for patients with adolescent idiopathic scoliosis. *Global Spine J* 2016;6:164-168.