Intra-Observer and Inter-Observer Reliability of Shaft Condylar Angle and Lateral Capitellohumeral Angle: Evaluation Based on Reliability in Different Ages and Levels of Experience

Piyanuch Musikachart, MD1, Thanase Ariyawatkul, MD2, Jidapa Wongcharoenwatana, MD2, Tananthorn Piamthipmanas, MD2, Suchitphon Chanchoo2, Perajit Eamsobhana, MD2

Departments of 1Biochemistry, Faculty of Medicine, Siriraj Hospital and 2Orthopaedic Surgery, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand

Objectives: The aims of this paper were: (i) to examine the intra-observer and inter-observer reliability of the shaft-condylar angle (SCA) and the lateral capitellohumeral angle (LCHA); (ii) to study the influence of experience level on the inter-observer and intra-observer reliability; and (iii) to determine the influence of the the age of the patients on reliability.

Method: A retrospective cohort study was conducted. The study reviewed 81 elbow radiographs. The patients were aged between 2 and 13 years. All the images taken between 2000 and 2017 were independently measured by a senior pediatric orthopaedic surgeon, a pediatric orthopaedic surgeon, a pediatric orthopaedic fellow, an orthopaedic chief resident, a general practitioner, and a pediatric orthopaedic research assistant. Measurement was performed two times within a 2-week interval. Inexperienced observers (general practitioner and research assistant) were supervised by senior pediatric orthopaedic surgeons for at least 30 radiographs before performing the measurement. Inclusion criteria were as follows: (i) age 2–13 years; and (ii) no previous elbow fracture. Exclusion criteria: elbow radiographs do not show true lateral view. The intraclass correlation coefficient (ICC) was used to calculate the reliability.

Results: The mean values of SCA and LCHA were 43° and 48°, respectively. For SCA, intra-observer reliability was excellent (ICC = 0.85) for one observer, good (range = 0.73–0.76) for three observers, and moderate (0.59) for one observer. Inter-observer reliability was moderate (0.48, 0.58), whereas the reliability categorized by age group showed excellent agreement (0.88–0.94). For LCHA, intra-observer reliability was excellent (0.84–0.89) for three observers and good (0.66–0.80) for two observers. Inter-observer reliability was moderate (0.44–0.45). Conversely, the reliability classified by age group showed excellent agreement (0.83–0.91).

Conclusion: Intra-observer reliability for LCHA and SCA were excellent to good for most observers. Inter-observer reliability was moderate for LCHA and SCA. Reliability classified by age group showed excellent to good agreement. Reliability was influenced by the level of experience, especially for non-medical staff.

Key words: Intra-observer and interobserver reliability; Lateral capitellohumeral angle ; Shaft condylar angle

Introduction

Fractures around the elbow are among the most common points of fracture in children1–5. These fractures are important because of pathophysiological defects in this area leading to nonunions6,7, neurovascular complications8, deformities7,9–11, and limitation of children’s activities12. For...
this reason, the recognition of injuries is essential. Anteroposterior (AP) and lateral images are recommended for use in the assessment of abnormality with signs and parameters. The shaft-condylar angle (SCA) and the lateral capitellohumeral angle (LCHA) are radiographic parameters used for evaluating the sagittal alignment of the elbow. The capitellum is one of the six ossification centers around the elbow joint, which grows throughout childhood. Using the capitellum for measurement of the angles could be unreliable due to the capitellum changing shape over time. Many signs (such as fat pad and sail signs) and parameters of elbow radiographs in pediatric patients were gradually developed to improve the accuracy of diagnosis. Irshad et al. reported that the sensitivity and specificity of the fat pad sign to diagnose radial head/neck fractures were 85.40% and 50%, respectively. The NPV and PPV were 87 and 47, respectively. The anterior humeral line (AHL) is another radiographic parameter used for appraising the severity and post-reduction achievement of supracondylar fractures in children. The anterior humeral line indicated for the line passed through the middle one-third of the capitellum in normal children. However, Herman et al. reported that the proportion between the AHL of the humerus and the anterior one-third of the capitellum in normal children below 4 years of age were increased. Likewise, Ryan et al. showed that all the AHL in the age range 0–5 years intersected with the middle one-third of the capitellum, but in 30% of children under the age of 2 years, it intersected with the anterior one-third of the capitellum. The previous reports provide evidence of the significance of the capitellum as one of the landmarks affecting reliability in elbow measurement.

The SCA is commonly used in lateral radiographic measurement of the elbow. In 2016, SCA was one of the radiographic parameters used to evaluate the outcome difference between the closed reduction and surgical fixation groups of patients with Garland type-II supracondylar fractures. A report of a late presented pediatric supracondylar fracture by Mulpreek et al. found that SCA can be the intraoperative reference for correction of sagittal alignment of the pediatric elbow. In terms of reliability, Suangyanon et al. was the only study that reviewed the normal side of elbow radiographs of 58 pediatric patients below 14 years of age. They detected excellent and good intra-observer and inter-observer reliability, respectively, using the Pearson correlation coefficient for calculation. However, research tended to focus on the overall reliability of six angles rather than the SCA. The reliability of the angle based on the age of the patients and the experience of the observer was still questionable.

The LCHA was measured by two lines which involved a part of the capitellum. Preliminary work on the reliability of LCHA was carried out using normal elbow radiographs in 75 children, aged 0–12 years, measured by five observers on three occasions. The reliability of intra-observers was moderate to good and inter-observer reliability was moderate to fair. Recently (2018), two studies reported that LCHA was reliable. A retrospective cohort study of 62 patients, aged 0–12 years, read by six observers was presented. Intra-observer reliability was moderate to excellent and inter-observer reliability was moderate in both periods of measurement. According to the other report, by Suangyanon et al., the correlation coefficient of intra-observers and inter-observers for LCHA was excellent and good, respectively. Apparently, the range of correlation coefficients for LCHA in previous studies is still varied and subject to disagreement.

The aim of this research was: (i) to study the intra-observer and inter-observer reliability of SCA and LCHA; (ii) to study the influence of experience level of observers on the intra-observer reliability; and (iii) to identify the age of the patients, which could affect reliability.

**Materials and Method**

The study was approved by the Institutional Review Board. A retrospective study was conducted on the data from 2000 to 2017. Data including age, gender and side of the patients were collected. The angles were measured using a radiograph archiving and communication system (PACS). PACS is the standard medical imaging program for collecting radiographic information. Inclusion criteria were as follows: (i) age 2–13 years; (ii) having normal side of elbow radiographs; and (iii) documented true lateral view. Radiographs revealing previous trauma, previous infection and inflammation, underlying bone disease, congenital disorder, and problems which interfere with the alignment of the elbow were excluded. The images were taken using a standard direct digital radiographic technique. The sample size was set as 120 radiographs based on the mean value of a previous study.

Eighty-one radiographs were independently measured by six observers (a senior pediatric orthopaedic surgeon, a pediatric orthopaedic surgeon, a pediatric orthopaedic fellow, an orthopaedic chief resident, a general practitioner and a pediatric orthopaedic research assistant) using PACS software. A pediatric orthopaedic surgeon described and demonstrated the method of measurement to other observers for consistency of measurement. Inexperienced observers (general practitioner and research assistant) had to practice under the supervision of pediatric orthopaedic surgeons for at least 30 radiographs before generating them by themselves.

All observers repeated the measurement on all radiographs after a period of 2 weeks from initial measurement to decrease the recall bias.

**Shaft Condylar Angle**

The SCA is one of the angles of elbow radiographs. The angle is defined as the angle between the axis of the humeral and the capitellum (Fig. 1A).

- A line was drawn on a lateral view along the anatomical axis of the humerus.
- A second line was drawn to bisect the capitellum into equal parts.
The area of intersection of these two lines should be at the metaphysis of the humerus.

- The anterior portion of the angle is SCA.

**Lateral Capitellohumeral Angle**

The LCHA is the importance angle to assessment the elbow abnormality on lateral view. The angle is formed from the axis of the humeral and the capitellum (Fig. 1B).

- A line was established along the anterior border of the humerus.
- A second axis was the line connecting the anterior and the posterior point on the proximal border of the capitellum.
- The two intersecting lines should connect at the proximal border of the capitellum.
- LCHA were assessed by measuring the anterior part of the angle.

Narrowing of the angles can be considered as the flexion hyperextension deformities. Abnormality of the angles can be considered as the deformities such as hyperextension or limit of elbow flexion. Supracondylar fracture and elbow dislocation can be possibly developed in both increased and decreased depending on mechanical force and type of fracture.

Data including the angle, sex, age, and side were collected. Intra-observer and inter-observer reliability of LCHA and SCA were analyzed using intraclass correlation coefficients (ICC) based on the 95% confidence interval for absolute agreement. A correlation coefficient of 0–0.20 indicated poor reliability. Those from 0.21 to 0.40 indicated fair reliability. Those from 0.41 to 0.60 indicated moderate reliability. Those from 0.61 to 0.80 indicated good reliability, and those greater than 0.81 indicated excellent reliability. ANOVA, the independent t-test and the Mann–Whitney test were used to compare the differences among age groups, and between genders and sides. Mean and standard deviation were used to describe the variability of the data. P-values of <0.05 were statistically significant. The software program used to analyze the data was SPSS for Windows Version 18.0 (SPSS, Chicago, IL, USA).

**Results**

**Patient Characteristics**

Eighty-one elbow radiographs of children 2–13 years of age were included and classified following the ranges 2–4, 5–7, 8–10, and 11–13 years. The mean age was 7.33 ± 2.83 years. Evaluations were carried out by six observers on two occasions. There were 54 (66.67%) males and 27 (33.33%) females. There were 35 (43.21%) radiographs of the left side and 46 (56.79%) of the right side (Table 1).

**Demographics of Radiographic Parameters**

Mean ± SD of SCA and LCHA classified by age, gender, and side are shown in Table 2. The mean SCA was 43.31° ± 4.66° and the mean LCHA was 47.79° ± 5.13°. The results did not reveal a statistically significant difference in angles among age groups. The mean scores by gender and side were different, although not statistically significantly.

**Intra-Reliability and Inter-Reliability of Shaft Condylar Angle and Lateral Capitellohumeral Angle**

Intra-observer ICC of the angles are presented in Table 3. Intra-observer reliability for SCA was excellent for one observer (pediatric orthopaedic fellow), good for three observers, but was only moderate for one observer (research assistant).
Fractures around the elbow are among the most common fractures in children. Anatomical structures of pediatric elbows are growing over time. Thus, parameters for evaluation of the injury were established to improve the accuracy of diagnosis. Recent research has demonstrated the reliability of practical parameters. Sagittal radiographs including shaft condylar angle (SCA) and lateral capitellohumeral angle (LCHA) have commonly been used to estimate the quality of reduction. Our study aimed to determine the intra-observer and inter-observer reliability of SCA and LCHA categorized by age group and level of experience in six observers. Of the 81 elbow radiographs of 2–13 years of age were included. The results indicated that intra-observer reliability was moderate to excellent for SCA and good to excellent for LCHA among all observers. In contrast, inter-observer reliability was moderate.

**Shaft Condylar Angle**

The mean score of the SCA was $43.31 \pm 4.66^\circ$, which was a much higher value than in Suangyanon et al. (2018), in which the value was $40.10^\circ \pm 6.24^\circ$ and $39.10^\circ \pm 6.25^\circ$, as confirmed by two orthopaedic pediatric fellows. ICC analysis was used to calculate the reliability of the angle. Although they found that the intra-observer reliability was excellent with ICC of 0.961, we found good reliability (ranging = 0.59–0.85). Our ICC was moderate on both occasions of measurement (0.48, 0.58) but the Suangyanon et al. reported good (0.747). The discrepancy in the results can be explained by the different methods and the characteristics of the population study. In Suangyanon et al., two observers examined 50 radiographs of patients aged less than 14 years. Our test was performed on 81 radiographs of 2–13-year-old patients by six observers. The reason that we did not include patients 0–2 years old was because the configuration of the capitellum was very small in size and of spherical shape. Evaluation was difficult and this could affect the reliability of measurement. Shank et al. described the

### Table 2: Mean value with P-value of the angles categorized by age group, gender, and side

| Demographics       | Shaft condylar angle (°) mean ± SD (minimum–maximum) | P value | Lateral capitellohumeral angle (°) Mean ± SD (minimum–maximum) | P value |
|--------------------|------------------------------------------------------|---------|---------------------------------------------------------------|---------|
| Total              | 43.31 ± 4.66 (29.92–51.08)                          |         | 47.79 ± 5.13 (37.25–65.75)                                    |         |
| Age group (years)  |                                                      |         |                                                              |         |
| 2–4                | 42.86 ± 4.66 (35.42–49.58)                           | 0.591   | 47.43 ± 4.81 (40.50–55.83)                                    | 0.369   |
| 5–7                | 43.42 ± 4.48 (34.58–51.08)                           |         | 46.69 ± 4.81 (37.25–53.83)                                    | 0.369   |
| 8–10               | 42.52 ± 5.40 (28.92–50.67)                           |         | 48.42 ± 6.14 (40.00–65.75)                                    |         |
| 11–13              | 44.57 ± 4.02 (37.25–50.92)                           |         | 49.34 ± 4.37 (38.67–57.92)                                    |         |
| Gender             |                                                      |         |                                                              |         |
| Male               | 44.45 ± 4.58 (28.92–51.08)                           | 0.001   | 46.68 ± 5.22 (37.25–65.75)                                    | 0.002   |
| Female             | 41.01 ± 3.99 (34.58–49.08)                           |         | 50.00 ± 4.18 (47.25–57.92)                                    |         |
| Side               |                                                      |         |                                                              |         |
| Left               | 42.02 ± 4.83 (28.92–51.08)                           | 0.030   | 49.05 ± 4.97 (37.25–65.75)                                    | 0.070   |
| Right              | 44.29 ± 4.34 (34.17–50.92)                           |         | 46.86 ± 5.11 (37.25–57.92)                                    |         |

### Table 3: The intraclass correlation coefficients (ICC) with 95% confidence interval (CI) of and intra-observer reliability categorized by six observers

| Observers                              | Shaft condylar angle (ICC) | 95% CI      | Lateral capitellohumeral angle (ICC) | 95% CI      |
|----------------------------------------|---------------------------|-------------|--------------------------------------|-------------|
| A senior pediatric orthopaedic surgeon| 0.74                      | 0.62–0.82   | 0.66                                 | 0.51–0.76   |
| A pediatric orthopaedic surgeon        | 0.76                      | 0.64–0.84   | 0.88                                 | 0.82–0.92   |
| A pediatric orthopaedic fellow         | 0.85                      | 0.77–0.90   | 0.84                                 | 0.76–0.90   |
| An orthopaedic chief resident          | 0.74                      | 0.62–0.82   | 0.89                                 | 0.84–0.93   |
| A general practitioner                 | 0.73                      | 0.62–0.82   | 0.80                                 | 0.71–0.87   |
| A pediatric orthopaedic research assistant | 0.59                   | 0.43–0.72   | 0.67                                 | 0.53–0.78   |
limitation of LCHA in those aged 0–2 years. The SCA also have a capitellum as one of the measurement landmarks. The unreliability of LCHA measurement at younger age might reflect the limitation of SCA measurement as well.

**Lateral Capitellohumeral Angle**

The mean value of LCHA was 47.79° ± 5.13° (the mean values for other studies were: 50.80° ± 6.00°, 44.70° ± 8.70°, 51.81° ± 7.56°, and 48.87° ± 7.88°). Intra-observer reliability was good to excellent, ranging from ICC = 0.67–0.89. The previous finding in the literature (Hasegawa et al., Suangyanon et al.). They reported intra-observer reliability ranging from ICC = 0.61–0.95 (good to excellent agreement) to ICC = 0.975 (excellent agreement), respectively. Another study by Shank et al. worked on 71 radiographs, aged 0–12 years, assessed by five observers. They showed moderate to good reliability (0.52–0.76).

For inter-observer reliability, our data was moderate for both measurements (0.44, 0.45). The value comparable with the older study. They evaluated 62 radiographs of children 2–11 years of age with six observers and reported moderate inter-observer reliability on two occasions of measurement (0.56, 0.51). However, they demonstrated fair reliability in the three periods of measurement among all observers (0.42, 0.34, 0.36). Subsequently, the highest reliability was verified with ICC of 0.67 (good agreement). These data need to be compared with caution due to the

| Period of measurement | Shaft condylar angle (°) | 95% CI | Lateral capitellohumeral angle (°) | 95% CI |
|-----------------------|--------------------------|--------|----------------------------------|--------|
| First measurement     | 0.48                     | 0.38–0.58 | 0.44                             | 0.34–0.54 |
| Second measurement    | 0.58                     | 0.48–0.67 | 0.45                             | 0.35–0.55 |

| Age group (years)     | Shaft condylar angle (°) | 95% CI | Lateral capitellohumeral angle (°) | 95% CI |
|-----------------------|--------------------------|--------|----------------------------------|--------|
| 2–4                   | 0.55                     | 0.33–0.78 | 0.45                             | 0.28–0.72 |
| 5–7                   | 0.64                     | 0.49–0.78 | 0.57                             | 0.42–0.73 |
| 8–10                  | 0.72                     | 0.58–0.85 | 0.63                             | 0.46–0.79 |
| 11–13                 | 0.55                     | 0.34–0.77 | 0.49                             | 0.28–0.73 |
measurement based on different demographic data of the patients and individual experience of the observers with pediatric radiographs.

Level of Experience

There is no available published evidence of the relationship between experience level and reliability of SCA and LCHA measurement. Few studies have explored the influence of different backgrounds of observers on orthopaedic-radiographic assessment. Previous studies have demonstrated that sensitivity in radiographic diagnosis of femoroacetabular impingement and acetabular dysplasia is improved with surgical experience, including an attending orthopaedic hip surgeon, an attending musculoskeletal radiologist, an orthopaedic sports fellow, and a third-year orthopaedic surgery resident. Nevertheless, false positive diagnosis increased with surgical experience level25. The outcomes were not statistically significantly different among all different experience levels of observers for detection of abnormal wrists and for chest radiographs26. Another study corroborated the results: inexperienced observers (two medical students and one research assistant) were able to interpret severe knee osteoarthritis radiographs using Kellgren–Lawrence (K–L) grade and the Osteoarthritis Research Society International (OARSI) Summary Score27. In addition, the reliability of joint space width was increased by repeated measurement, as reported by Ornetti et al.28.

This study was carried out by six observers with different levels of experience. The research assistant established the lowest ICC value of intra-observer reliability for SCA (Table 3); the reliability was only moderate (0.59) but the others verified good to excellent reliability (0.73–0.85). Table 4 shows the data for intra-observer reliability sorted by age group. The research assistant also generated the smallest score in almost all age groups (5–13 years). Likewise, the intra-observer reliability of measurement for LCHA was dependent on skill level. This was not only limited by the lowest experience but also the greatest experience for ages 2–7 years (Tables 4 and 5). This corresponded with Shank et al., with five observers, including two pediatric orthopaedic surgeons, two orthopaedic residents, and one pediatrician. The intra-observer reliability of the pediatrician was lower than that of the others (0.52 vs 0.65–0.76)22. Conversely, Hasegawa et al. found no correlation in reliability among six readers with different experience in orthopaedic surgery.23

Age-Related Reliability of measurement

The inter-observer reliability for SCA and LCHA among all age groups was moderate to good (range, 0.55–0.72 and 0.45–0.63, respectively). Interestingly, the reliability decreased for the extremely young (2–4 years) and the older age (11–13 years) group. The reliability was increased in the middle-aged group (5–10 years) for both angles, which could be explained by the changing capitellum. As mentioned previously, the capitellum is one of the significant landmarks of measurement but it changes over time. In younger children, the structure is small, leading to difficulty identifying the axis of the anterior border. In older children, the structure is beginning to fuse with the humerus, leading to difficulty measuring as well.

Limitations

Our research has two potential limitations. First, less experienced observers could compromise the overall results of measurements. Second, the sample size was calculated for 120 radiographs, but the present study used 81 radiographs.
20. Mulpruek P, Angsanuntsukh C, Woratanarat P, Sa-Ngasooongsong P, Tawonsawatruk T, Chanplakom P. Shaft-Condylar Angle for surgical correction in neglected and displaced lateral humeral condyle fracture in children. Acta Orthop Belg, 2015, 81: 384–391.

21. Suangyanon P, Chalayon O, Worawuthangkul K, Kaewponsawan K, Ariyawatkul T, Eamsobhana P. Pediatric elbow measurement parameters: evaluation of the six angles in inter- and intra-observer reliability. J Clin Orthop Trauma, 2018. https://doi.org/10.1016/j.jcot.2018.07.019.

22. Shank CF, Wiater BP, Pace JL, et al. The lateral capitellohumeral angle in normal children: mean, variation, and reliability in comparison to Baumann’s angle. J Pediatr Orthop, 2011, 31: 266–271.

23. Hasegawa M, Suzuki T, Kuroiwa T, et al. Reliability of radiographic measurement of lateral capitellohumeral angle in healthy children. Medicine (Baltimore), 2018, 97: e0314.

24. Skibo L, Reed MH. A criterion for a true lateral radiograph of the elbow in children. Can Assoc Radiol J, 1994, 45: 287–291.

25. Schottel PC, Park C, Chang A, Knutson Z, Ranawat AS. The role of experience level in radiographic evaluation of femoroacetabular impingement and acetabular dysplasia. J Hip Preserv Surg, 2014, 1: 21–26.

26. Margolis SA, Nilsson KA, Reed RL. Performance in reading radiographs: does level of education predict skill? J Contin Educ Health Prof, 2003, 23: 48–53.

27. Kiara K, Collins JE, Gurary E, et al. Reliability and accuracy of cross-sectional radiographic assessment of severe knee osteoarthritis: role of training and experience. J Rheumatol, 2016, 43: 1421-1426.

28. Ornetti P, Maillefert JF, Paternotte S, Dougados M, Gossec L. Influence of the experience of the reader on reliability of joint space width measurement. A cross-sectional multiple reading study in hip osteoarthritis. Joint Bone Spine, 2011, 78: 499–505.