Reproducibility and Reliability of Plain Radiographic Parameters of DDH in Children

1Bomar J D; 2Upasani V; 1Gaurav P; +Hosalkar H
+Rady Children’s Hospital, Orthopedic Department, San Diego, CA,
2University of California, San Diego, Department of Orthopaedic Surgery, San Diego, CA
hhosalkar@rchsd.org

INTRODUCTION: Several radiographic parameters have been used to define and monitor the progress of hip dysplasia in infancy and childhood. The purpose of this study was to assess the inter- and intra-observer reliability for four common radiographic measures of hip dysplasia (Sharp’s angle, center-edge angle, acetabular index and femoral head-neck angle).

METHODS: Fifty antero-posterior pelvis radiographs from 21 patients with DDH were examined by three blinded observers. The average age at time of x-ray was 10 months, ranging from 4 to 33 months of age. Radiographs were obtained digitally and printed on paper. Identifying information was removed from all radiographs. Each observer drew the lines necessary to measure neck shaft angle, center edge angle of Wiberg, sharp angle, and acetabular index. Lines were drawn bilaterally. In total, eight individual measures were evaluated for each of the 50 x-rays. However, the right center edge angle of Wiberg was not measured on one radiograph because that hip was dislocated. A fourth participant used a protractor to measure and record each angle, from each observer. One week later, the 50 AP pelvis radiographs were put in random order and printed out again, and the exercise outlined above was repeated.

Statistical Analysis. The intra- and inter-observer reliability for each measure was evaluated using the intraclass correlation statistic (ICC) in SPSS (version 12, IBM, Armonk, New York, USA). For the purposes of this study an ICC of 0.7 was chosen as the cut-off point between acceptable and unacceptable agreement.

All measures were also categorized as either requiring treatment, or not requiring treatment. Neck shaft angle measures were categorized as not requiring treatment if they were ≤ 130°. Center edge angle measures were categorized as not requiring treatment if they were > 25°. Sharp angle measures were categorized as not requiring treatment if they were ≤ 42°. Acetabular index measures were categorized as not requiring treatment if they were ≤ 30°.

RESULTS: Inter-observer reliability. At time period one we found an acceptable level of agreement between observers when measuring the neck shaft angle (ICC = 0.868, 95% confidence interval (CI) 0.812 to 0.908). The three observers also tended to agree on whether or not treatment was required based on the neck shaft angle (k = 0.598, 95% CI 0.078 to 1.000) and the acetabular index (k = 0.538, CI 0.365 to 0.711) (see Table).

At time period two we found an acceptable level of agreement between observers when measuring neck shaft angle (ICC = 0.900, 95% CI 0.864 to 0.928), center edge angle (ICC = 0.742, 95% CI 0.613 to 0.828), and acetabular index (ICC = 0.775, 95% CI 0.703 to 0.836). The three observers also tended to agree on whether or not treatment was required based on the same three measures: neck shaft angle (k = 0.740, 95% CI 0.196 to 1.000), center edge angle (k = 0.472, 95% CI 0.088 to 0.935), acetabular index (k = 0.493, 95% CI 0.330 to 0.656) (see Table).

Intra-observer reliability. When we compared observer one’s measures at different time periods we found an acceptable level of agreement when measuring neck shaft angle (ICC = 0.808, 95% CI 0.853 to 0.930) and acetabular index (ICC = 0.721, 95% CI 0.609 to 0.805). When these measures were evaluated for whether or not treatment would have been necessary they tended to be in agreement for the neck shaft angle (k = 0.852, 95% CI 0.566 to 1.000), center edge angle (k = 0.497, 95% CI 0.178 to 0.816), and acetabular index (k = 0.532, 95% CI 0.146 to 0.918).

When we compared observer two’s measures at different time periods we found an acceptable level of agreement when measuring neck shaft angle (ICC = 0.931, 95% CI 0.895 to 0.955), center edge angle (ICC = 0.730, 95% CI 0.618 to 0.813), and acetabular index (ICC = 0.702, 95% CI 0.587 to 0.789). When these measures were evaluated for whether or not treatment would have been necessary they tended to be in agreement for the neck shaft angle (k = 0.740, 95% CI 0.395 to 1.000), sharp angle (k = 0.662, 95% CI 0.043 to 1.000), and acetabular index (k = 0.458, 95% CI 0.270 to 0.646).

When we compared observer three’s measures at different time periods we found an acceptable level of agreement when measuring neck shaft angle (ICC = 0.893, 95% CI 0.826 to 0.932) and acetabular index (ICC = 0.804, 95% CI 0.721 to 0.863). When these measures were evaluated for whether or not treatment would have been necessary they tended to be in agreement for all four measures: neck shaft angle (k = 0.580, 95% CI 0.010 to 0.950), center edge angle (k = 0.402, 95% CI 0.141 to 0.663), sharp angle (k = 1.000, 95% CI 1.000 to 1.000), and acetabular index (k = 0.615, 95% CI 0.425 to 0.805).

DISCUSSION: Based on the results of this study we can conclude that Acetabular index and neck shaft angle are reliable and reproducible indices in patients with suspected hip dysplasia for diagnosis and follow-up.

SIGNIFICANCE: Bi-planar radiographs continue to provide reliable information for diagnosis of DDH in children less than three years. True three dimensional assessment of the acetabulum remains important in patients with acetabular dysplasia. Advanced non-radiation imaging techniques including 3D ultrasound may prove to be the future for diagnosis, prognosis, and management of this condition in your children.

ACKNOWLEDGEMENT: We would like to thank Christine Farnsworth, MS for helping to coordinate this project.