# Measuring the Acetabular Index: An Accurate and Reliable Alternative Method of Measurement

Benjamin Sherman, DO[1,2], Francois D. Lalonde, MD[1], John A. Schlechter, DO[1,2]

Musculoskeletal Imaging · Original Research

**Keywords**

acetabular index, developmental hip dysplasia, musculoskeletal imaging, pediatric

Submitted: Apr 9, 2020
Revision requested: May 16, 2020
Revision received: Jun 12, 2020
Accepted: Jul 2, 2020
First published online: Apr 28, 2021

This article is available for credit.

The authors declare that they have no disclosures relevant to the subject matter of this article.

**OBJECTIVE.** In children (4 months to 8 years old), radiographic measurements of the acetabular index are the preferred method to assess developmental hip dysplasia. However, the acetabular index has been criticized as having variable reliability owing to difficulty identifying the correct anatomic landmarks. An alternative method of measuring the acetabular index using the ischium is being proposed to avoid the variability of the triradiate cartilage line as a reference point. With the alternative method, the acetabular index is derived by measuring the angle between a line connecting the ischial tuberosities and a line connecting the inferomedial and superolateral edges of the acetabulum. The purpose of this study was to evaluate the accuracy and reliability of this alternative method of measuring the acetabular index compared with the traditional method.

**MATERIALS AND METHODS.** Children 4 months to 8 years old who presented for evaluation of developmental dysplasia of the hip were included. Two physicians, each using both the traditional and the alternative method, measured acetabular indexes on all radiographs. Accuracy was defined as mean absolute error less than 6°. Reliability was calculated by means of intraclass correlation coefficient (ICC).

**RESULTS.** Pelvic radiographs of 40 children (324 hips) were included. The mean age was 23.7 months (range, 4–96 months) and mean acetabular index was 24.2° (range, 8–50°). The alternative method was associated with mean absolute error of 2.50°, which is significantly below the threshold of 6° ($t < 0.001$). Intrarater reliability for the traditional method was high (ICC, 0.81) and for the alternative method was very high (ICC, 0.92). Interrater reliability for the traditional method was high (ICC, 0.89) and for the alternative method was very high (ICC, 0.91).

**CONCLUSION.** Measuring the acetabular index using the alternative method has very high accuracy and intrarater and interrater reliability.

Developmental dysplasia of the hip is a common pediatric orthopedic condition with an estimated incidence of 1.4–80 cases per 1000 births [1–6]. Acetabular dysplasia is characterized by a shallow, immature acetabulum and may present with or without hip instability or dislocation [7]. Instability of the hip causes superior and lateral migration of the femoral head that increases localized contact pressure, which eventually leads to degenerative changes [6]. The severity of subluxation of the femoral head has an inverse relation to the time it takes for symptoms to develop [6]. Mild hip subluxation becomes symptomatic at 40–50 years of age, moderate subluxation at 20–30 years, and severe subluxation at 10–20 years [6]. The diagnosis of acetabular dysplasia is made with imaging studies [8]. Ultrasound is the preferred imaging modality for infants from birth to 4 months or until the femoral head ossifies, after which time radiography becomes the standard method of screening [8–12].

Several radiographic measurements have been described over the years to assess for developmental dysplasia or dislocation of the hip, including the Shenton line, center-edge angle, neck-shaft angle, and the acetabular index [6, 8, 13–16]. The acetabular index is an objective measurement of acetabular dysplasia and can be used to determine the severity of dysplasia and efficacy of treatment.

The traditional method of measuring the acetabular index involves measuring the angle between a line connecting the triradiate cartilages of both hips and a line bisecting the inferomedial and superolateral edges of the acetabulum [17] (Fig. 1). The mean measurement typically decreases with age from 27.5° in the neonatal period to 20.0° by age 2

doi.org/10.2214/AJR.20.23358

*AJR* 2021; 217:172–176

ISSN-L 0361–803X/21/2171–172



[1]Children's Hospital of Orange County, Orange, CA.

[2]Department of Orthopedic Surgery, Riverside University Health System, 26520 Cactus Ave, Moreno Valley, CA 92555. Address correspondence to B. Sherman (DrBenjaminSherman@gmail.com).

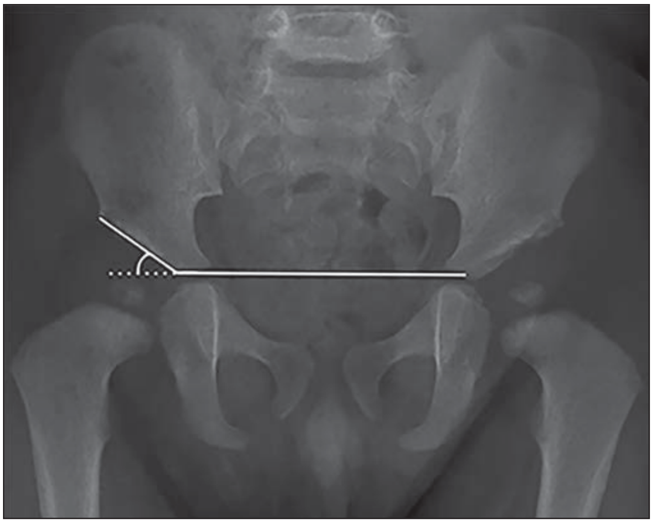

**Fig. 1**—20-month-old healthy boy with bilateral developmental dysplasia of hip. Anteroposterior pelvic radiograph shows traditional method of assessing acetabular index by measuring angle between extension (*dotted line*) of line connecting triradiate cartilages and line bisecting inferomedial and superolateral edges of acetabulum.

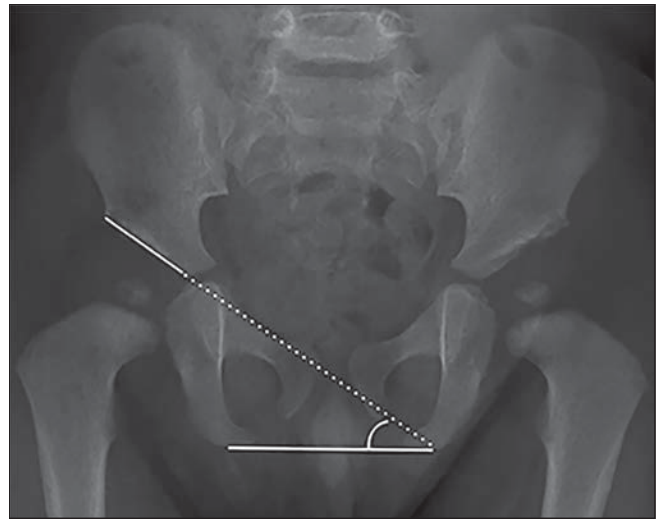

**Fig. 2**—20-month-old healthy boy with bilateral developmental dysplasia of hip. Anteroposterior pelvic radiograph shows alternative method of assessing acetabular index by measuring angle between line connecting lowest points of ischial tuberosity (*solid line*) and line connecting inferomedial acetabular edge and superolateral edge (*dotted and solid line*).

years [6, 17–19]. The acetabular index has been criticized as having variable rates of interobserver and intraobserver reliability. The main criticism is difficulty identifying reliable radiographic anatomic landmarks [8, 18, 20].

An alternative method of measuring the acetabular index is to use the ischium to avoid the variability of the triradiate cartilage line as a reference point. In the alternative method the acetabular index is derived by measuring the angle between a line connecting the ischial tuberosities and a line connecting the inferomedial and superolateral edges of the acetabulum (Fig. 2). This technique has been used by radiologists, orthopedists, and pediatricians in clinical practice to measure the acetabular index; however, the technique and accuracy of the method have, to our knowledge, never been officially described, evaluated, or validated for use in developmental hip dysplasia. The purpose of this study was to evaluate the accuracy and reliability of the alternative method of measuring the acetabular index in comparison with the traditional, reference standard method.

## Materials and Methods

Institutional review board approval was obtained for this study. The orthopedic clinic visits of one of the authors were queried for patients presenting with developmental dysplasia or dislocation of the hip between January 1 and December 31, 2017. Patients younger than 4 months or older than 8 years and those with neuromuscular disorders were excluded. Forty patients met the inclusion criteria. Anteroposterior pelvic radiographs of all patients were examined, including serial radiographs of the same patient if undergoing follow-up at intervals during the year. Only preoperative radiographs were included if the patient underwent reconstructive hip surgery. For patients being treated with an abduction brace or Pavlik harness, only radiographs without the brace or harness were included.

A total of 162 anteroposterior radiographs were selected for review, for a total of 324 hips. Each radiograph was assigned a number by means of the random number generator function of Microsoft Excel (version 2011), and 10 anteroposterior radiographs were duplicated. All radiographs were screened for quality according to the criteria outlined by Clohisy et al. [18]. For a radiograph to be considered acceptable, three criteria had to be met. The coccyx had to be in line with the symphysis pubis; the obturator foramina, iliac wings, and teardrops had to be symmetric; and the distance between the inferior tip of the coccyx and pubic symphysis had to be 1–3 cm. High-quality and poor-quality radiographs are shown in Figures 3 and 4. A total of 123 radiographs were deemed acceptable and 39 unacceptable based on these criteria. Image quality was labeled and kept in a separate spreadsheet, to which the measuring physicians were blinded. No images were removed from final radiographic evaluation. Images were anonymized and uploaded into a folder in the PACS.

The acetabular index was measured twice for both the right and left hips on all 162 anteroposterior pelvic radiographs (324 hips). Both measuring physicians were orthopedic surgeons with fellowship training in pediatric surgery, one with 21 years of postfellowship experience and one with 12 years of postfellowship experience. One physician was randomly assigned to start measuring all assigned radiographs using the traditional method, and the other surgeon started with the alternative method. Minimal directions were given on how to obtain the measurements, and no demonstration was performed. A physician who did not perform the measurements recorded the other physicians' measurements in a spreadsheet, and the record was not shared with the measuring physicians. Once the surgeons measured all radiographs using the assigned method, they remeasured them using the other method. In total, each physician measured 324 radiographs, or 648 hips. All measurements were performed on 1 day in the same order. Radiographs used for intrarater reliability were randomized and included in the 162 anteroposterior pelvic radiographs measured.

Measurements from radiographs that were deemed acceptable were entered in a spreadsheet and used for the primary

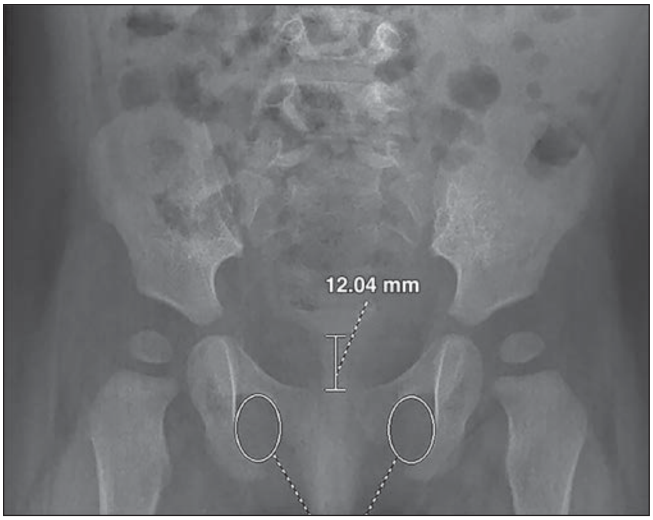


**Fig. 3**—7-month-old healthy girl with very mildly shallow acetabulum bilaterally. Anteroposterior pelvic radiograph depicts criteria used to determine high quality. Coccyx is in line (*solid line*) with symphysis pubis. Iliac wings, teardrops, and obturator foramina (*ovals, bottom dashed lines*) are symmetric. Distance between inferior tip of coccyx and pubic symphysis measures 1–3 cm (*top dashed line*).

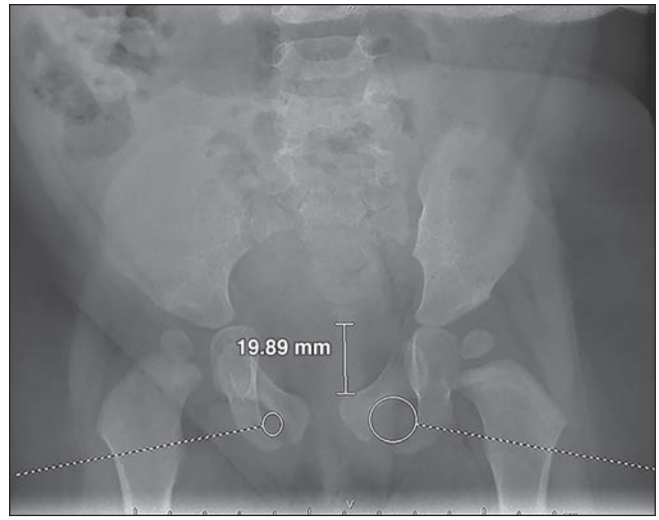


**Fig. 4**—9-month-old healthy girl with developmental dysplasia of left hip. Anteroposterior pelvic radiograph is poor-quality image. Coccyx appears rotated compared with pubic symphysis. Obturator foramen and iliac wings are asymmetric (*circles, dashed lines*). Distance between inferior tip of coccyx and pubic symphysis (*solid line*) does meet criterion.

analysis. Measurements from unacceptable radiographs were entered into a different spreadsheet and analyzed separately. Measurements from the duplicate radiographs were copied into a third spreadsheet for calculation of intraobserver reliability. The data were organized and sent to an independent statistician for analysis. The statistician was first asked to evaluate the accuracy of the alternative method compared with the reference standard. Before initiation of the study, measurements within 6° were considered equivalent on the basis of the work of Broughton et al. [21], who established this range as the threshold of variability. The statistician was also asked to evaluate the interrater and intrarater reliability of both measurement methods and whether age or image quality affected accuracy and reliability.

Reliability statistics were assessed by intraclass correlation coefficient (ICC). Munro correlation categories were used to determine the strength of the relationships within the categories, as follows: 0.0–0.25, little; 0.26–0.49, low; 0.5–0.69, moderate; 0.7–0.89, high; 0.9–1.0, very high [22]. Mean absolute error between the two measurement methods was calculated and correlated with age by means of Pearson correlation coefficient. Mean absolute error was compared between measurements obtained from acceptable and unacceptable radiographs by means of ANOVA after verification of the data for normality and homogeneity of variances. Bland-Altman plots were constructed with difference and mean values for each subject. Repeated-measures ANOVA was used to compare the average of measurements for the traditional and alternative methods for each subject. Accuracy was defined as mean absolute error less than 6°. The data for both raters were combined, and a one-sample $t$ test was performed to evaluate the mean absolute error difference from the threshold of 6°. All analyses were performed with SPSS Statistics for Microsoft Windows software (version 24.0, IBM). An alpha value less than .05 indicated significance.

## Results

### Characteristics and Demographics

The study included 40 patients (mean age, 23.7 months; range, 4–96 months) and 162 radiographs (324 hips). The quality of 123 radiographs was considered acceptable, and the quality of 39, unacceptable. The mean acetabular index measurement was 24.2° (range, 8–50°).

### Primary Outcome: Accuracy

The measurements obtained with the alternative method deviated a mean of 2.50° from those obtained with the traditional method, and this deviation (mean absolute error) was significantly ($t < 0.001$) below the predetermined 6° threshold. When only high-quality radiographs were used, the error decreased slightly to 2.42°. When only poor radiographs were used, the error increased slightly to 2.74°. However, there was no statistically significant difference between the two mean absolute error measurements ($\alpha = .09$). The complete analysis is shown in Table 1.

### Secondary Outcome: Reliability

Interrater reliability for the traditional method was high (ICC, 0.89) and for the alternative method was very high (ICC, 0.91). Intrarater reliability for the traditional method was high (ICC, 0.81) and for the alternative method was very high (ICC, 0.92). Neither age nor radiograph quality was found to affect measurement error.

## Discussion

The acetabular index is an important measurement in the evaluation of hip dysplasia in children 4 months to 8 years old. In many cases, the acetabular index is used to determine the need for treatment and monitor the response to brace and/or surgical treatment. As such, a reliable method of measuring the acetabular index is essential to ensure standardization of care. Variable rates of intraobserver and interobserver reliability for the

**TABLE 1: Accuracy in Comparison of High-Quality and Low-Quality Radiographs**

| Image Quality | No. of Measurements | Mean Absolute Error (°) | 95% Confidence Limits of Intrarater Reliability (°) | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| High | 512 | 2.42 | 2.25 | 2.60 |
| Low | 164 | 2.74 | 2.41 | 3.06 |
| Combined high and low | 676 | 2.50 | 2.34 | 2.66 |

Note—In comparison of high- and low-quality mean absolute error measurements , $\alpha = .09$.

traditional method have been documented in previous studies. Upasani et al. [23] reported on the reliability of the acetabular index measurement in a study that included 21 patients (50 radiographs). Using the Munro categories, they found high intraobserver (0.89) and interobserver (0.70) reliability (ICC, 0.775) for measurements obtained by three observers—one pediatric orthopedic attending surgeon, one pediatric orthopedic surgery fellow, and an orthopedic surgery chief resident.

In a study similar to that conducted by Upasani et al. [23], Dornacher et al. [24] measured the acetabular index on anteroposterior radiographs of 51 10- to 22-month-old patients with hip dysplasia to determine reliability. Intraobserver reliability was very high with a Pearson correlation coefficient of 0.928, and interobserver reliability was high with a Pearson correlation coefficient of 0.875). Broughton et al. [21] found the 95% CI for intrarater reliability to be ± 6.1° and interrater reliability to be ± 5.5° with an overall recommendation that the true measurement be within 6°. Conversely, Kay et al. [20] found more variability in measurement of the acetabular index in their study of 24 pelvic radiographs with measurements obtained by five different examiners (three orthopedic surgery residents, two pediatric orthopedic attending surgeons. The study showed that with good-quality radiographs, the 95% CI for intraobserver reliability was 8.35° and for interobserver reliability was 21.3°. With poor-quality radiographs, reliability decreased significantly. The authors concluded that variability in the examiner's chosen landmarks may be the cause of the measurement's variability.

## *Limitations*

This study had several limitations. Only two physicians performed measurements. Because the measurement ability and experience of measuring physicians can vary, this factor might make the results less translatable than they would be if more reviewers had participated. However, the interrater reliability of this method is high, which supports the premise that physicians have little variation in their measurements. Measurements were performed by pediatric fellowship–trained orthopedic surgeons and not by radiologists. Although these physicians had considerable experience performing pelvic measurements, they might have been less accurate than experienced radiologists, which might have increased the mean standard error of the study. Measurements with both methods were performed on a single day. Although measuring physicians might have remembered their previous measurements, it is unlikely that they did given the high volume of hips measured by each physician. It is possible that the volume of measurements caused fatigue in the measuring physicians, which might have falsely decreased the accuracy of the measurements. However, given that the mean standard error was very low (2.50° variation), this would only strengthen our conclusions.

The strength of our study lies in the methods and strong study design. To our knowledge, this study has the largest sample size to date: 324 hips measured compared with 24 hips measured by Kay et al. [20] and 102 by Dornacher et al. [24]. The high intraobserver and interobserver reliability for the traditional method of measurement of the acetabular index documented in our study closely mirrors the results reported by Upasani et al. [23] and Dornacher et al. Unlike Kay et al., we did not find variability in our measurements based on the quality of the radiographs. The results for the alternative method of measuring the acetabular index showed mild improvement over the traditional method in terms of intrarater reliability with similar interrater reliability. The alternative method had a high degree of accuracy (2.50° difference) compared with the reference standard measurements.

## *Conclusion*

Variations in measurement of the acetabular index can be attributed to chosen landmarks. We found that compared with the traditional method, the alternative method had similar accuracy and greater reliability and is a valid option for measuring the acetabular index on pelvic radiographs of children.

## References

1. Mirdad T. Incidence and pattern of congenital dislocation of the hip in Aseer region of Saudi Arabia. *West Afr J Med* 2002; 21:218–222
2. Munkhuu B, Essig S, Renchinnyam E, et al. Incidence and treatment of developmental hip dysplasia in Mongolia: a prospective cohort study. *PLoS One* 2013; 8:e79427
3. Phelan N, Thoren J, Fox C, O'Daly BJ, O'Beirne J. Developmental dysplasia of the hip: incidence and treatment outcomes in the southeast of Ireland. *Ir J Med Sci* 2015; 184:411–415
4. Kokavec M, Bialik V. Developmental dysplasia of the hip: prevention and real incidence. *Bratisl Lek Listy* 2007; 108:251–254
5. Kolb A, Schweiger N, Mailath-Pokorny M, et al. Low incidence of early developmental dysplasia of the hip in universal ultrasonographic screening of newborns: analysis and evaluation of risk factors. *Int Orthop* 2016; 40:123–127
6. Herring JA. *Tachdjian's pediatric orthopaedics: from the Texas Scottish Rite Hospital for Children.* Elsevier Saunders, 2014
7. Barrera CA, Cohen SA, Sankar WN, Ho-Fung VM, Sze RW, Nguyen JC. Imaging of developmental dysplasia of the hip: ultrasound, radiography and magnetic resonance imaging. *Pediatr Radiol* 2019; 49:1652–1668
8. Keller MS, Nijs EL. The role of radiographs and US in developmental dysplasia of the hip: how good are they? *Pediatr Radiol* 2009; 39(suppl 2):S211–S215
9. Nguyen JC, Dorfman SR, Rigsby CK, et al.; Expert Panel on Pediatric Imag-

ing. ACR Appropriateness Criteria®: developmental dysplasia of the hip—child. *J Am Coll Radiol* 2019; 16(5 suppl):S94–S103
10. Zucker EJ, Lee EY, Restrepo R, Eisenberg RL. Hip disorders in children. *AJR* 2013; 201:[web]W776–W796
11. Mulpuri K, Song KM. AAOS clinical practice guideline: detection and nonoperative management of pediatric developmental dysplasia of the hip in infants up to six months of age. *J Am Acad Orthop Surg* 2015; 23:206–207
12. Tudor A, Sestan B, Rakovac I, et al. The rational strategies for detecting developmental dysplasia of the hip at the age of 4–6 months old infants: a prospective study. *Coll Antropol* 2007; 31:475–481
13. Wiberg G. Studies on dysplastic acetabula and congenital subluxation of the hip joint: with special reference to the complication of osteoarthritis. *Acta Chir Scand* 1939; 83:53–68
14. [No authors listed]. Classic: translation—Hilgenreiner on congenital hip dislocation. *J Pediatr Orthop* 1986; 6:202–214
15. Shenton E. A diagnostic line about the hip joint. *J Phys Ther* 1902; 3:110–112
16. Scoles PV, Boyd A, Jones PK. Roentgenographic parameters of the normal infant hip. *J Pediatr Orthop* 1987; 7:656–663
17. Kleinberg S, Lieberman HS. The acetabular index in infants in relation to congenital dislocation of the hip. *Arch Surg* 1936; 32:1049–1054
18. Clohisy JC, Carlisle JC, Beaulé PE, et al. A systematic approach to the plain radiographic evaluation of the young adult hip. *J Bone Joint Surg Am* 2008; 90(suppl 4):47–66
19. Than P, Sillinger T, Kránicz J, Bellyei A. Radiographic parameters of the hip joint from birth to adolescence. *Pediatr Radiol* 2004; 34:237–244
20. Kay RM, Watts HG, Dorey FJ. Variability in the assessment of acetabular index. *J Pediatr Orthop* 1997; 17:170–173
21. Broughton NS, Brougham DI, Cole WG, Menelaus MB. Reliability of radiological measurements in the assessment of the child's hip. *J Bone Joint Surg Br* 1989; 71:6–8
22. Munro BH. *Statistical methods for health care research.* Lippincott Williams & Wilkins, 2005
23. Upasani VV, Bomar JD, Parikh G, Hosalkar H. Reliability of plain radiographic parameters for developmental dysplasia of the hip in children. *J Child Orthop* 2012; 6:173–176
24. Dornacher D, Cakir B, Reichel H, Nelitz M. Reliability of digital measurement of acetabular index in hip dysplasia to the time children start walking [in German]. *Z Orthop Unfall* 2008; 146:246–250

## FOR YOUR INFORMATION

ARRS is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education activities for physicians.

The ARRS designates this journal-based CME activity for a maximum of 1.00 *AMA PRA Category 1 Credits*™ and 1.00 *American Board of Radiology*©, MOC Part II, Self-Assessment CME (SA-CME). Physicians should claim only the credit commensurate with the extent of their participation in the activity.

To access the article for credit, follow the prompts associated with the online version of this article.