

Reliability of a New Radiographic Classification for Developmental Dysplasia of the Hip

Unni Narayanan, MBBS, MSc, FRCS(S),* Kishore Mulpuri, MBBS, MS (Ortho),
MHS(Ed),† Wudbhav N. Sankar, MD,‡ Nicholas M.P. Clarke, ChM, FRCS, FRCS.Ed,§
Harish Hosalkar, MBBS, MD,|| Charles T. Price, MD, FAAP,¶
and International Hip Dysplasia Institute

Background: Existing radiographic classification schemes (eg, Tönnis criteria) for DDH quantify the severity of disease based on the position of the ossific nucleus relative to Hilgenreiner's and Perkin's lines. By definition, this method requires the presence of an ossification centre, which can be delayed in appearance and eccentric in location within the femoral head. A new radiographic classification system has been developed by the International Hip Dysplasia Institute (IHDI), which uses the mid-point of the proximal femoral metaphysis as a reference landmark, and can therefore be applied to children of all ages. The purpose of this study was to compare the reliability of this new method with that of Tönnis, as the first step in establishing its validity and clinical utility.

Methods: Twenty standardized anteroposterior pelvic radiographs of children with untreated DDH were selected purposefully to capture the spectrum of age (range, 3 to 32 mo) at presentation and disease severity. Each of the hips was classified separately by the IHDI and Tönnis methods by 6 experienced pediatric orthopaedists from the United States, Canada, Mexico, United Kingdom, and by 2 orthopaedic senior residents. The inter-rater reliability was tested using the Intra Class Correlation coefficient (ICC) to measure concordance between raters.

Results: All 40 hips were classifiable by the IHDI method by all raters. Ten of the 40 hips could not be classified by the Tönnis method because of the absence of the ossific nucleus on one or both sides. The ICC (95% confidence interval) for the IHDI method for all raters was 0.90 (0.83-0.95) and 0.95 (0.91-0.98) for the right and left hips, respectively. The corresponding ICCs

for the Tönnis method were 0.63 (0.46-0.80) and 0.60 (0.43-0.78), respectively. There was no significant difference between the ICCs of the 6 experts and 2 trainees.

Conclusions: The IHDI method of classification has excellent inter-rater reliability, both among experts and novices, and is more widely applicable than the Tönnis method as it can be applied even when the ossification centre is absent.

Level of Evidence: Level II (diagnostic).

Key Words: developmental dysplasia of the hip, DDH, classification, reliability

(*J Pediatr Orthop* 2015;35:478-484)

Developmental dysplasia of the hip (DDH) is a common disorder in the pediatric population with an overall incidence of approximately 3 to 4 per 1000 live births.¹ Although the clinical examination remains the mainstay for diagnosing DDH in early infancy, not all cases of DDH are detectable by physical examination. Imaging, either by ultrasonography or radiography, have become popular worldwide for screening or confirmation of the diagnosis as well as to classify the severity of the dysplasia. Quantifying the severity of displacement in DDH is useful for prognostication and clinical decision-making. In the newborn and young infant, ultrasounds are superior because the cartilaginous femoral head is not visible on radiographs. Following the appearance of the ossific nucleus in the older infant, anteroposterior (AP) pelvic radiographs replace ultrasonography as the screening imaging modality of choice. Ultrasonography has the limitation of being less accurate with continued growth and ossification of the femoral head. However, radiographic assessment may still be suboptimal or misleading when ossification of the femoral head is delayed or eccentric as in hips with DDH.

Several ultrasonographic and radiographic parameters have been described to define hip dysplasia. Some radiographic parameters such as the center-edge angle of Wiberg and the acetabular angle of Hilgenreiner are used to quantify subtle differences in acetabular dysplasia.^{2,3} In the older child, the center-edge angle of Wiberg is used to evaluate the degree of lateral femoral head coverage by the acetabulum in the frontal plane.² The acetabular index

From the *Hospital for Sick Children, Toronto, ON; †British Columbia Children's Hospital, Vancouver, BC, Canada; ‡Children's Hospital of Philadelphia, Philadelphia, PA; §Scripps Mercy Hospitals, San Diego, CA; ¶Orlando Health, Orlando, FL; and ||University of Southampton, Southampton General Hospital, Hampshire, England.

No authors received support for this study.

The authors declare no conflicts of interest.

Reprints: Charles T. Price, MD, FAAP, Department of Orthopaedic Surgery, Orlando Health, 1222 S. Orange Ave, 5th Floor, Orlando FL 32806. E-mail: charles.price@orlandohealth.com.

Copyright © 2014 Wolters Kluwer Health, Inc. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

angle of Hilgenreiner quantifies the development of the acetabulum by measuring the slope of the boney acetabulum in the frontal plane.³ However, these do not address more severe categories of dislocation for prognostic purposes before treatment. For classification of the full spectrum of severity, the Graf sub-types are well established on ultrasound, whereas the Tönnis method helps classify the severity of DDH using plain radiographs.⁴

The Tönnis method quantifies the severity of DDH using the relative position of the ossific nucleus to Perkin's (P-line) and Hilgenreiner's lines (H-line).⁴ This method relies on the presence of the ossification centre, which can be delayed in appearance and is often eccentrically located within the femoral head.⁵ These issues can therefore make the application of the Tönnis classification in the younger infant quite difficult and potentially unreliable. Despite this limitation, orthopedic surgeons around the world who treat children with DDH continue to utilize plain radiographs and the Tönnis criteria to confirm the diagnosis of DDH, to establish treatment recommendations, and to assess the effect of treatment.^{6,7} Recognizing the limitation of the Tönnis method in young infants, the International Hip Dysplasia Institute (IHDI) group developed a new radiographic classification system to quantify the severity of the displacement of the femoral head, which does not rely on the presence of the ossific nucleus and, therefore, can be applied to children of all ages. The aims of this study were 2-fold: (1) to test the reliability of the IHDI classification of the severity of the hip displacement in DDH both by experienced pediatric orthopedic surgeons as well as by residents in training, and (2) to compare the reliability of the IHDI method with that of the Tönnis classification.

METHODS

The proposed IHDI classification uses the position of the proximal femoral metaphysis (instead of the ossific nucleus) as the important reference landmark to determine the location of the hip. We define the H-point as the midpoint of the superior margin of the ossified metaphysis. As with the Tönnis system, H-line is drawn through the top of the tri-radiate cartilages bilaterally. The standard P-line is then drawn perpendicular to the H-line at the superolateral margin of the acetabulum. Unlike previous classification systems, an additional diagonal line (D-line) is then drawn 45 degrees from the junction of H-line and P-line. The position of the H-point then determines the IHDI grade. In an IHDI grade I hip, the H-point is at or medial to the P-line. In grade II, the H-point is lateral to the P-line and at or medial to the D-Line. For grade III, the H-point is lateral to the D-line and at or inferior to the H-line. Finally, in grade IV hips, the H-point is superior to the H-line (Table 1, Fig. 1).

To test the reliability of the classification system, a purposeful sample of antero-posterior pelvis radiographs of 20 children with untreated DDH across a wide age spectrum from 3 months to 32 months (mean age: 10 mo) were selected by the senior author who was not involved

TABLE 1. Radiographic Measurements for IHDI Classification

Radiographic Measurement	Description
Hilgenreiner's line (H-line)	The single line drawn through the (top of) the triradiate radiate cartilage bilaterally
Perkin's line (P-line)	Perpendicular line from the superolateral margin of the acetabulum
Diagonal line (D-line)	45-degree line from the junction of H-line and P-line
H-point	Midpoint of the superior margin of the ossified metaphysis

in the reliability assessment. The selected sample included both unilateral and bilateral involvement and a spectrum of severity. All radiographs selected were before any treatment.

The radiographs were then analyzed by 8 independent observers who were blinded to the measurements of each other and to the identity of the patients. Observers included: (1) 6 consultant pediatric orthopedic surgeons with a special interest in pediatric hip surgery, (2) 2 orthopedic senior residents who had completed 6 months on a pediatric orthopedic rotation. Any patient identifying information was removed from all radiographs.

All observers were asked to classify both hips on each radiograph with both the Tönnis method (Fig. 2) as well as the IHDI method. Radiographs were supine AP views of the pelvis that were obtained with the hips at rest and with lower limbs held gently (without traction) in the neutral position with the patellae facing forward. Radiographic landmarks in the Tönnis and IHDI methods were defined based on the classic descriptions of these specific parameters from the literature. The technique for each radiographic parameter was agreed upon by the observers a priori, to minimize errors in marking of the landmarks, (ie, drawing the lines) necessary for each classification. For the IHDI classification, it was agreed that an H-point that fell on a line would be scored as the lower grade. In addition, a standardized transparent measurement tool was provided to each observer to aid in the marking of landmarks (Fig. 3).

Statistical Analysis

Statistical analyses were performed utilizing SPSS software (version 12, IBM; Armonk, NY). Intra observer and interobserver reliability for each measure was assessed using intraclass correlation coefficient (ICC).⁸ The ICC reports a value between 0.0 and 1.0, where 1.0 represents perfect agreement or concordance. Although there are no definitive values that clearly differentiate between acceptable and unacceptable agreement, for the purposes of this study we have adopted Munro's correlation strength categories (0.9-1.0 = "very high"; 0.7-0.89 = "high"; 0.5-0.69 = "moderate"; 0.26-0.49 = "low"; 0.0-0.25 = "little if any").⁹ The 2-way mixed model (absolute agreement) was utilized as the observers in this study were not randomly selected and were measuring identical radiographs.

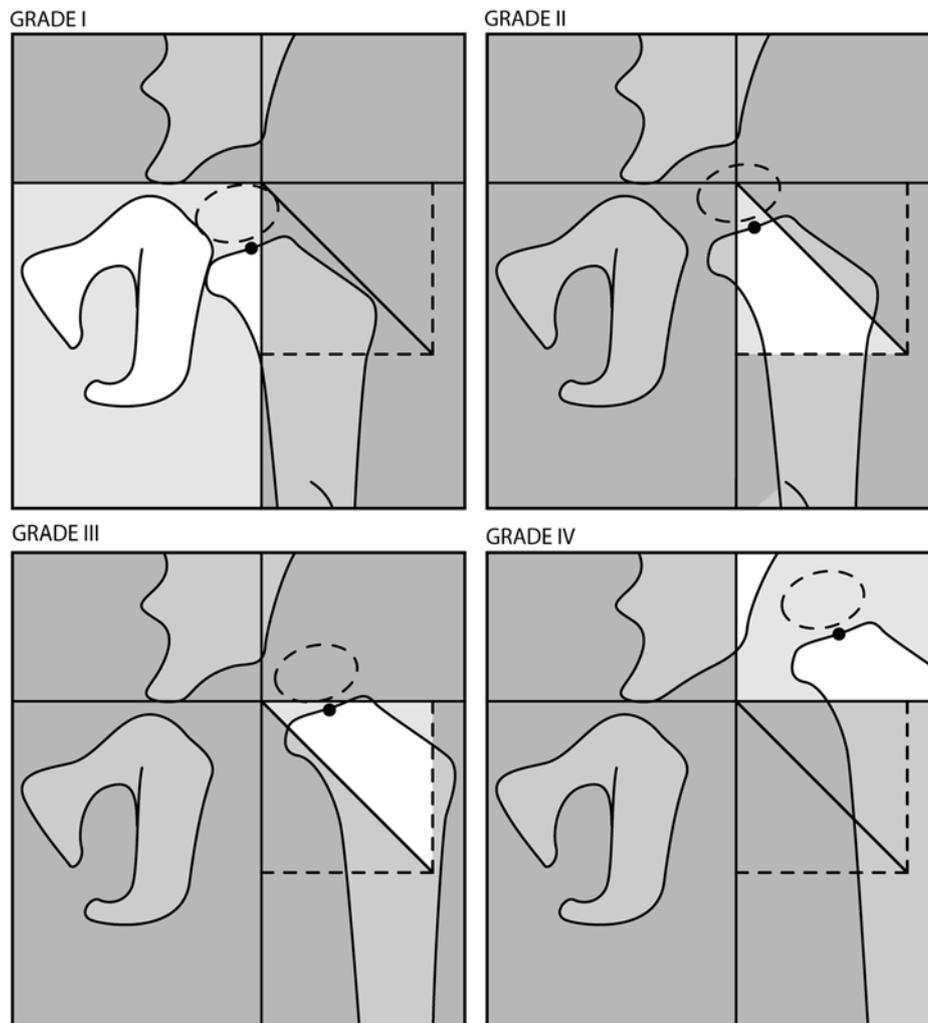


FIGURE 1. IHD classification for DDH (not requiring the presence of an ossific nucleus). H-line is Hilgenreiner’s line drawn through the top of the tri-radiate cartilages bilaterally. P-line is Perkin’s line drawn perpendicular to the H-line at the superolateral margin of the acetabulum. D-line is diagonal line drawn 45 degrees from the junction of H-line and P-line. H-point is the midpoint of the superior margin of the ossified metaphysis. Grade I: the H-point is at or medial to the P-line. Grade II: the H-point is lateral to the P-line and at or medial to the D-line. Grade III: the H-point is lateral to the D-line and at or inferior to the H-line. Grade IV: the H-point is superior to the H-line.

RESULTS

All 20 radiographs (40 hips) were classifiable by the IHD method by all the raters (6 consultant pediatric orthopaedic surgeons and 2 senior orthopedic residents) (Figs. 4A–D). Ten of the 40 (25%) hips could not be classified by the Tönnis method because of the absence of the ossific nucleus, although some raters provided a Tönnis grade based on their assumption of the location of the ossific nucleus.

Interobserver Reliability

We noted a “high to very high” level of agreement between all observers when classifying the severity of DDH with the IHD classification system, with an ICC of 0.90 [95% confidence interval (CI), 0.83–0.95] for the right hip and an ICC of 0.95 (95% CI, 0.91–0.98) for the left hip.

Using the Tönnis classification method, there was a “moderate to high” level of agreement between all observers with an ICC of 0.63 (95% CI, 0.46–0.80) for the right hip and an ICC of 0.63 (95% CI, 0.43–0.78) for the left hip.

For the IHD method, there were no significant differences between the ICCs of the 6 consultant pediatric orthopedic surgeons (0.90 and 0.95) for the right and left hips, respectively, and the 2 trainees (0.95 and 0.98) for the right and left hips respectively.

DISCUSSION

Although ultrasound is typically used for infants under the age of 6 months (who typically have unossified epiphyses), frontal plane pelvis radiographs are still frequently used across the world to diagnose and/or monitor DDH in

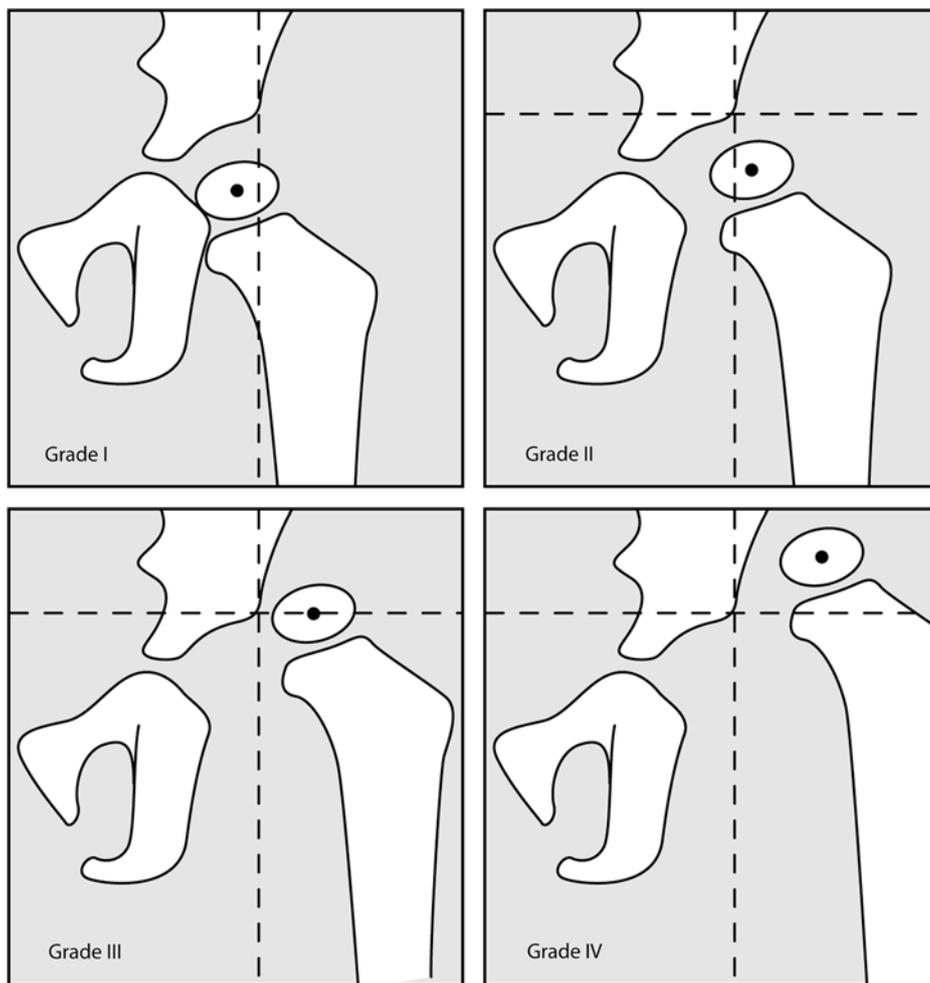


FIGURE 2. Tönnis classification for DDH (relies on the location of the ossific nucleus).⁴ Grade I: the ossification center of the capital epiphysis is medial to the perpendicular line from the superolateral margin of the acetabulum (Perkins's line). Grade II: the ossification center of the capital epiphysis is lateral to the Perkins's line, but below the superolateral margin of the acetabulum (SMA-line). Grade III: the ossification center is at the level of the superolateral margin of the acetabulum. Grade IV: the ossification center is above the superolateral margin of the acetabulum.

older infants and children. Common practice divides severity into (1) dysplasia, (2) subluxation, and (3) dislocation. However, this practice does not allow accurate assessment of intermediate degrees of hip displacement. Although various radiographic parameters have been described to evaluate treatment results or quantify the deformity in acetabular and proximal femoral development, few classification systems exist to evaluate the severity of DDH on plain radiographs.

We are aware of 3 published radiographic methods for evaluation of DDH severity at time of diagnosis. One is the method of Yamamuro and Chene¹⁰ that measures the distance from the center of the metaphysis to the H-line and the distance between the center of the metaphysis and the medial wall of the acetabulum (Fig. 5). Although this measurement has been proven reliable, it is a continuous variable that has not been divided into categories for comparative purposes.

Dyson et al¹¹ described a method of relating the position of the proximal femoral metaphysis to the acetabulum using the metaphyseal-edge (ME) angle.

The ME angle is subtended between P-line and a line drawn from the medial border of the upper femoral metaphysis to the lateral edge of the bony acetabulum. The ME angle is positive if the medial edge of the upper femoral metaphysis lies medial to the lateral edge of the bony acetabulum and negative if it lies lateral to it. In their study, a negative ME angle always indicated a dislocation. Although this system accounts for the presence of an unossified epiphysis, like the method of Yamamuro it does not provide categories of more severe dislocations for comparative purposes.

The final, and most widely used, classification was developed by the Commission for the Study of Hip Dysplasia of the German Society of Orthopedics and Traumatology as reported by Tönnis, and commonly called the Tönnis Classification.^{4,12} The Tönnis method of classifying DDH has been reported to be prognostic. Rosen et al¹³ performed measurements and statistical analysis of initial radiographs, the 1-year follow-up

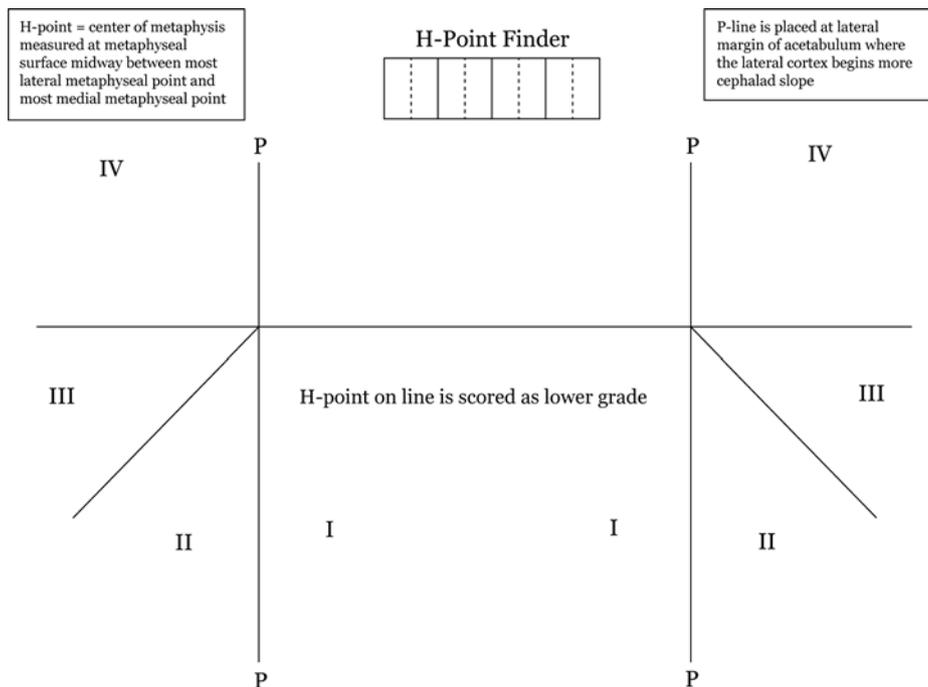


FIGURE 3. Transparent measurement tool provided to each observer.

radiographs, and the most recent radiographs of 81 patients (103 hips) to determine, which measurements could predict the success of treatment. The average follow-up was 49 months (range, 12 to 139 mo), and the average age of the patients at the last follow-up was 65 months (range, 15 to 190 mo). Analysis of the data showed that the measurement with statistically significant predictive value was the Tönnis grade of dislocation determined from the initial radiograph. A single unit increase in the Tönnis grade was associated with a doubling of the odds of failure in patients treated with a Pavlik harness ($P < 0.04$, odds ratio = 2.2) or closed reduction (odds ratio = 2.0).

A major limiting factor of the Tönnis system is that it requires the presence of an ossific nucleus of the femoral head on the plain radiograph. Rosen et al¹³ applied the Tönnis method to newborn infants and others without the ossification center, but the participants in the current study were unable to reliably classify hips by the Tönnis method before ossification of the femoral head. Eccentric position or delayed appearance of the ossific nucleus is a common finding in DDH, particularly in infants aged 6 to 18 months. As this is a common age of presentation in the authors' practices, the need for a more accurate classification system remains.

The Yamamura method of quantifying lateral and proximal displacement of the femoral head was validated by Boniforti and colleagues.^{10,14} A study by Kitoh et al¹⁵ evaluated the prognostic value of the Yamamura method for infants treated with the Pavlik harness. These authors found that the Pavlik harness was more likely to be successful when the distance from H-line to the center of the metaphysis was > 6 mm. No other authors have evaluated the prognostic significance of the Yamamura method.

Although the Tönnis and Yamamura methods have been found useful, neither provides categories of classification that are easily and equally applicable whether the ossification center is present or absent. The purpose of this study was to propose a simple radiographic classification system independent of the ossific center that is reliable both in the hands of experts and novices.

In this study, we demonstrated “high” to “very-high” interobserver reliability for DDH classification by the IHDI system, which was superior to the interobserver reliability of the Tönnis method, as performed by 6 pediatric orthopedic surgeons and 2 orthopedic senior residents. There were no significant differences between the ICCs of the 6 attending pediatric orthopedic surgeons versus the 2 trainees. Furthermore, the Tönnis method could not be used to classify 10 of the hips in this sample, because the ossific nucleus was not visible. In contrast, the IHDI method could be applied in these patients, proving its applicability to infants and children of all ages.

It is difficult to accurately measure 3-dimensional pelvic abnormalities on a frontal plane radiograph, especially when important pelvic landmarks have yet to ossify. The greatest variability can occur in identifying the center of the minimally ossified femoral head as well as the lateral margin of the acetabulum. The location of the H-point, however, was consistent in this cohort, because the femoral metaphysis could be well defined radiographically. This further strengthens the argument that femoral metaphysis is a more reliable and consistent landmark in all children with DDH, irrespective of age and progress of ossification of femoral head.

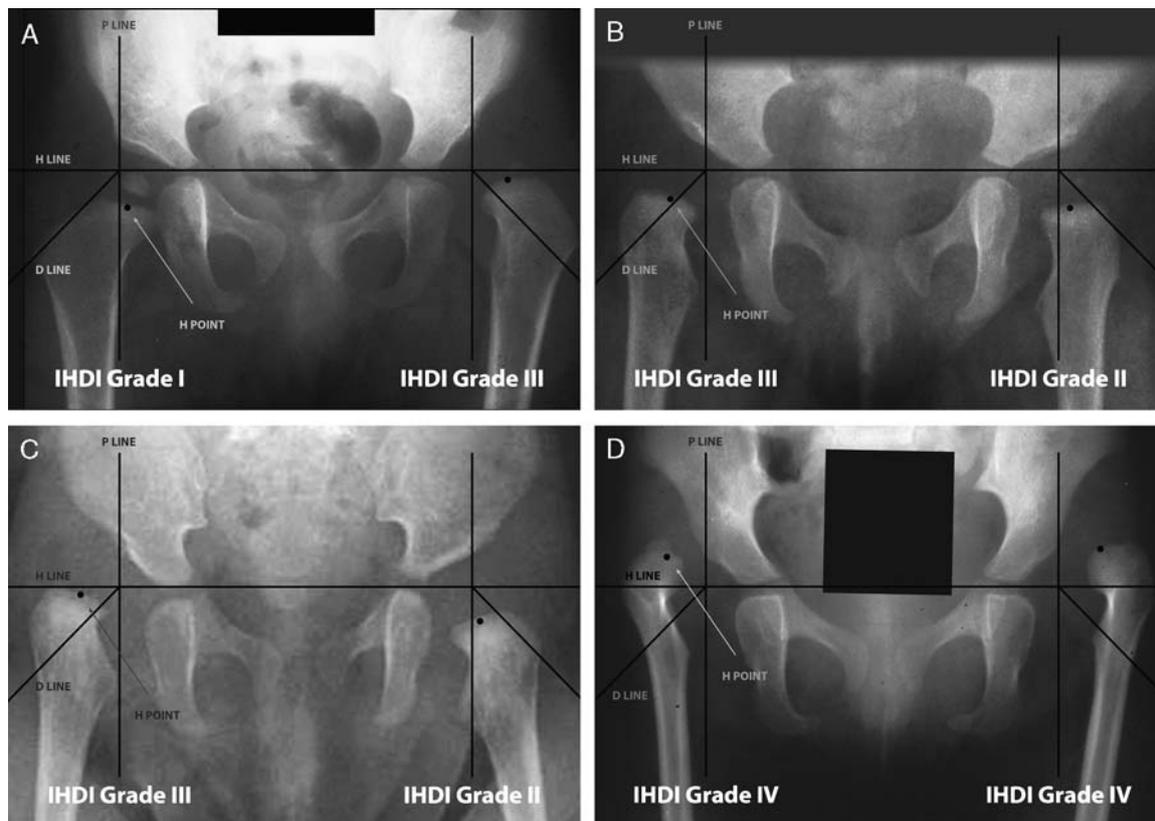


FIGURE 4. (A) Case example of an anteroposterior (AP) pelvic radiograph of an infant with DDH. The right hip was classified as an IHDI grade I hip, whereas the left hip was classified as grade III. (B) Case example of an AP pelvic radiograph of another infant with bilateral DDH. The right hip was classified as an IHDI grade III hip, whereas the left hip was classified as grade II. (C) Third case example of an AP pelvic radiograph of an infant with DDH. The right hip was classified as an IHDI grade III hip, whereas the left hip was classified as grade II. (D) Final case example of an AP pelvic radiograph of an infant with bilateral DDH. Both hips were classified as IHDI grade IV. Note that all hips in the case examples could be classified, although the ossific nucleus was variably present.

One limitation of this study was the potential for selection bias by overrepresenting hips without an ossification center in the selected sample for comparing the 2

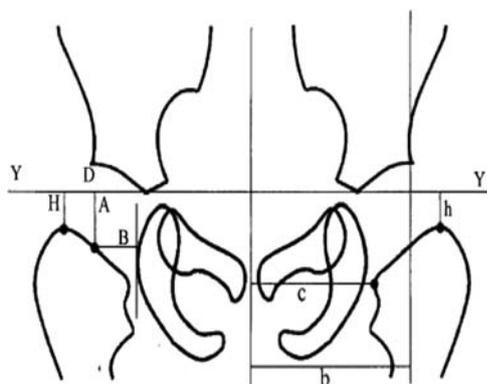


FIGURE 5. Yamamuro measurement method. Y indicates Y-line; A, Yamamuro-A; B, Yamamuro-B; H, Hilgenreiner-H; D, Hilgenreiner-D; h/b and c/b ratio, Smith's ratio for superior and lateral displacement. Reproduced with permission from Boniforti et al.¹⁴

grading systems. To minimize this risk, the sample was selected by an investigator who was not involved in the reliability grading and was based on inclusion of a wide range of severity and ages at presentation (range, 3 to 32 mo; mean, 10 mo) and not on the presence or absence of the ossific nucleus.

The main limitation of the IHDI classification system is its 2-dimensional/single plane assessment obtained from one AP view of the pelvis. In the era of advanced imaging, including the ability to three dimensionally assess the bony and soft-tissue morphology with nonradiation-based ultrasound technology, it is likely that there will be options available in the future to move from traditional biplanar radiographs to nonradiation, nonanesthesia-based dynamic 3-dimensional assessments, which could allow for a more reproducible assessment of hip dysplasia.

If significant variability exists between observers, or between time periods for a single observer, it is difficult to create clinical pathways to treat children with DDH and determine the impact of a certain treatment method over time. Having a simple, reliable classification that has good interobserver reproducibility and is simple to teach and

use, can be invaluable in the management of this condition. Regardless of its limitations, we believe that the new IHDI classification has potentially greater utility in light of its superior reliability and wider applicability across all ages. A future goal of the IHDI is to conduct a prospective multicenter longitudinal study to establish the validity and utility of this classification system both for the purpose of prognostication and for clinical decision making in the management of DDH.

ACKNOWLEDGMENT

The authors thank Allison Crepeau, Patrick Wright, James Kasser, Scott Mubarak, Pablo Castañeda, and John Wedge for their contributions to the reliability grading.

REFERENCES

1. Dezateux C, Rosendahl K. Developmental dysplasia of the hip. *Lancet*. 2007;369:1541–1552.
2. Wiberg G. Studies on dysplastic acetabula and congenital subluxation of the hip joint. *Acta Chir Scand*. 1938;58(suppl):28–39.
3. Hilgenreiner H. Early diagnosis and early treatment of congenital dislocation of the hip. *Med Klin*. 1925;21:1385–1388.
4. Tönnis D. *Congenital Dysplasia and Dislocation of the Hip in Children and Adults*. Berlin, Germany: Springer-Verlag; 1987.
5. Benson M, Fixsen J, Macnicol M, et al. *Children's Orthopaedics and Fractures*. London: Springer London; 2009.
6. Wang TM, Wu KW, Shih SF, et al. Outcomes of open reduction for developmental dysplasia of the hip: does bilateral dysplasia have a poorer outcome? *J Bone Joint Surg Am*. 2013;95:1081–1086.
7. Heesakkers NA, Witbreuk MM, Besselaar PP, et al. Retrospective radiographic evaluation of treatment results of developmental dysplasia of the hip in walking-age children. *J Pediatr Orthop Part B*. 2013;22:427–431.
8. Portney LG, Watkins MP. Statistical measures of reliability. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River, NJ: Prentice-Hall; 2008:557–588.
9. Munro BH. *Statistical Methods for Health Care Research*. 5th ed. Philadelphia: Lippincott Williams & Wilkins; 2005.
10. Yamamuro T, Chene SH. A radiological study on the development of the hip joint in normal infants. *J Jpn Orthop Assoc*. 1975;49:421–429.
11. Dyson PH, Lynskey TG, Catterall A. Congenital hip dysplasia: problems in the diagnosis and management in the first year of life. *J Pediatr Orthop*. 1987;7:568–574.
12. Tönnis D. Radiological classification and diagnosis. *Mapfre Medicina*. 1992;3(suppl):42–45.
13. Rosen A, Gamble JG, Vallier H, et al. Analysis of radiographic measurements as prognostic indicators of treatment success in patients with developmental dysplasia of the hip. *J Pediatr Orthop Part B*. 1999;8:118–121.
14. Boniforti FG, Fujii G, Angliss RD, et al. The reliability of measurements of pelvic radiographs in infants. *J Bone Joint Surg Br*. 1997;79:570–575.
15. Kitoh H, Kawasumi M, Ishiguro N. Predictive factors for unsuccessful treatment of developmental dysplasia of the hip by the Pavlik harness. *J Pediatr Orthop*. 2009;29:552–557.