Potential for Change in US Diagnosis of Hip Dysplasia Solely Caused by Changes in Probe Orientation: Patterns of Alpha-angle Variation Revealed by Using Three-dimensional US

**Purpose:** To use three-dimensional (3D) ultrasonography (US) to quantify the alpha-angle variability due to changing probe orientation during two-dimensional (2D) US of the infant hip and its effect on the diagnostic classification of developmental dysplasia of the hip (DDH).

**Materials and Methods:** In this institutional research ethics board–approved prospective study, with parental written informed consent, 13-MHz 3D US was added to initial 2D US for 56 hips in 35 infants (mean age, 41.7 days; range, 4–112 days), 26 of whom were female (mean age, 38.7 days; range, 6–112 days) and nine of whom were male (mean age, 50.2 days; range, 4–111 days). Findings in 20 hips were normal at the initial visit and were initially inconclusive but normalized spontaneously at follow-up in 23 hips; 13 hips were treated for dysplasia. With the computer algorithm, 3D US data were resectioned in planes tilted in 5° increments away from a central plane, as if slowly rotating a 2D US probe, until resulting images no longer met Graf quality criteria. On each acceptable 2D image, two observers measured alpha angles, and descriptive statistics, including mean, standard deviation, and limits of agreement, were computed.

**Results:** Acceptable 2D images were produced over a range of probe orientations averaging 24° (maximum, 45°) from the central plane. Over this range, alpha-angle variation was 19° (upper limit of agreement), leading to alteration of the diagnostic category of hip dysplasia in 54% of hips scanned.

**Conclusion:** Use of 3D US showed that alpha angles measured at routine 2D US of the hip can vary substantially between 2D scans solely because of changes in probe positioning. Not only could normal hips appear dysplastic, but dysplastic hips also could have normal alpha angles. Three-dimensional US can display the full acetabular shape, which might improve DDH assessment accuracy.

---

From the Departments of Radiology and Diagnostic Imaging (J.L.J., V.G.S., L.J.) and Biomedical Engineering (M.M., K.C., R.B.T.), 2A2.41 WC Mackenzie Health Sciences Centre, University of Alberta, 8440-112 St, Edmonton, AB, Canada T6G 2B7. Received February 22, 2014; revision requested April 1; final revision received April 10; accepted April 21; final version accepted April 29. Address correspondence to J.J. (e-mail: jjaremko@ualberta.ca).

© RSNA, 2014
Ultrasoundography (US) is the preferred imaging modality to evaluate developmental dysplasia of the hip (DDH) in infants younger than 6 months (1). DDH is common, occurring in 1.6–28.5 per 1000 infants, depending on definition criteria (2), and one in 1000 infants is born with a dislocated hip (1). Imaging is needed because clinical diagnostic tests such as the Ortolani and Barlow maneuvers lack sensitivity after the neonatal period and for mild disease (3). With US, the most widely used index for evaluating DDH is the acetabular alpha angle, measured by using the Graf method (4) on a standardized two-dimensional (2D) coronal image that must contain a flat horizontal iliac wing, labrum, bony and cartilaginous acetabular roof, and os ischium. These criteria are intended to standardize the plane in which imaging is performed. With the Graf method, a hip with an alpha angle greater than 60° is classified as normal, a hip with an alpha angle from 50° to 59° is classified as having mild dysplasia, a hip with an alpha angle from 43° to 49° is classified as having moderate dysplasia, and a hip with an alpha angle of less than 43° is classified as having severe dysplasia. Other groups prefer different thresholds (eg, classifying a hip with an alpha angle greater than 55° as normal, a hip with an alpha angle of 45°–55° as borderline, and a hip with an alpha angle of less than 45° as abnormal) (5). Treatment decisions are based on the degree of dysplasia and clinical factors, particularly the age of the patient at the time of diagnosis.

Diagnosing dysplasia requires combining clinical and imaging assessment, and, in current practice, the quantitative component of imaging assessment relies heavily on the value of the alpha angle. The alpha angle is subject to wide variability of two main types: interscan and interobserver. Interobserver variability is easily quantified: two people measuring the alpha angle differently on the same 2D US image. Interscan error is more difficult to analyze. Because the US probe is handheld, each scan inevitably is obtained with the probe and the patient in a different position, which, because of the complex three-dimensional (3D) acetabular shape, results in images showing different alpha angles. Quantification of alpha-angle variability due to probe orientation in live patients or in dysplastic hips has not been well established in the literature. This variability is important because it could lead to misdiagnosis of DDH. Missed cases of DDH are particularly of concern because DDH is usually easily treated conservatively if diagnosed early in infancy (6), whereas a later diagnosis is associated with more invasive treatment and lower success rates (1). To avoid this problem, clinical practice often includes follow-up US until initially borderline hips definitely appear normal, with associated costs and parental anxiety, or routine overtreatment until hips normalize. The poor diagnostic accuracy of current US of the hip may account for its lack of beneficial effect on patient care, with rates of late DDH diagnosis or surgery not reduced by using universal or selective US, compared with rates with clinical screening alone, despite higher treatment rates (2).

Three-dimensional US probes now can allow acquisition of images rapidly enough for scans to be feasible even in a noncompliant infant. Because the entire acetabular shape can be acquired in a single 3D data set, the effect of differences in 2D probe orientation can be explored by viewing sections cut along different planes through 3D data. Compared with obtaining repeated 2D measurements, this approach has the advantage of holding all variables constant for the 2D section orientation, allowing the effects of this orientation to be evaluated independently. The purpose of this study was to quantify alpha-angle variability due to changing probe orientation during 2D US of the infant hip and its effect on the diagnostic classification of DDH.

### Advances in Knowledge

- Three-dimensional (3D) ultrasoundography (US) technology can be used to show that interscan variability of alpha angles measured on two-dimensional (2D) US images, solely because of changes in probe positioning, is substantial (upper limit of agreement, 19°) for infants with normal and dysplastic hips.

- Changing probe orientation altered the alpha angle enough to allow a change in the Graf diagnostic category in more than 50% of infant hips imaged and in 72% of hips imaged in infants younger than 1 month.

- Three-dimensional US is feasible in assessing patients for developmental dysplasia of the hip (DDH) because a 3D sweep of the hip can be performed in 3.2 seconds and with less difficulty in positioning than can 2D US.

### Implications for Patient Care

- Assessing DDH in infants by using 2D US is limited by substantial variation caused by probe positioning.

- Three-dimensional US demonstrates the full acetabular shape and, with validation, might offer more reproducible DDH assessment in the future.
Materials and Methods

Patients
This prospective study was approved by the University of Alberta Health Research Ethics Board. Imaging was performed at a tertiary pediatric hospital (Stollery Children’s Hospital Pediatric Orthopedic Clinic, Edmonton, Alberta, Canada) from October 2012 to November 2013. At the first routine clinical 2D US of the hip in each patient, written informed consent was obtained from a parent to add 3D US of the hip at the same visit. The imaging indication was clinical suspicion of DDH because of laxity at examination, asymmetric hip creases, or other risk factors such as a family history of a positive finding for DDH. Because dysplasia can be unilateral or bilateral, we included each hip separately. Considering only the 3D scans showing the entire range of planes necessary for this study, we had 36 hips in 35 patients, and 26 (74%) of them were female. Patients underwent US at the mean age of 41.7 days (range, 4–112 days; for female infants, mean age was 38.7 days [range, 6–112 days]; for male infants, mean age was 50.2 days [range, 4–111 days]). They received routine care at a pediatric orthopedic clinic from one or more of five clinic surgeons who were blinded to 3D US images and findings. We observed clinical care for at least 6 months to classify each imaged hip as normal at the first orthopedic assessment (category 0; n = 20), questionable abnormal initially but with findings that resolved spontaneously at follow-up imaging and clinical examination (category 1; n = 23), or dysplastic and requiring treatment by using a Pavlik harness and/ or surgery (category 2; n = 13).

Imaging
All imaging was performed by using platforms (Philips iU22; Philips Healthcare, Andover, Mass). We performed conventional 2D US in both hips by using a 12-MHz linear transducer (12 L5; Philips Healthcare), including static coronal imaging in the Graf standard plane and axial dynamic imaging to assess for hip stability by using our usual protocol. Our usual protocol is to gather images showing the best Graf representation of each hip with and without measurements for alpha angle, beta angle, and coverage, as well as axial dynamic images. Two-dimensional US images were interpreted by a pediatric radiologist, with results and images made available to referring clinicians. In addition, two study team members, including a radiologist, technologist, or medical or graduate student trained by the study radiologists, used a high-resolution 13-MHz 3D linear (13VL5; Philips Healthcare) transducer in the coronal orientation to obtain a 3D US data set at each hip. With this transducer, we performed a 3.2-second automated sweep through a range of ±15° starting at 0° to generate a 3D data set of 256 US sections that were 0.13 mm thick, each containing 411 × 192 pixels measuring 0.11 × 0.20 mm. These 3D scans were not released for use in clinical treatment.

Image Processing
Images were analyzed off-line by using custom software (Matlab R2010a-2012; MathWorks, Natick, Mass) that allowed viewing of 3D US images followed by markup of points, lines, and curves replicating the functionality of standard cardiac magnetic resonance imaging workstation analysis of features such as ventricular contours. The major obstacle to performing this analysis was the proprietary 3D data file format. A key function of the custom software was to extract manufacturer-specific pixel spacing information from the 3D US Digital Imaging and Communications in Medicine data. For each hip, an observer (M.M., a graduate student in biomedical engineering and radiology) trained under the lead radiologist (J.L.J.) traced the relevant anatomic contours (acetabular rim, femoral head cartilage surface, os ischium) on the original 3D US data set by using a customized semi-automated interactive interface that allowed contours traced on selected sections to be interpolated to intervening sections, then reviewed and corrected by using a nudging tool. A 3D model then was generated for each structure by creating a patch surface connecting the traced and interpolated contours (Fig 1). The training process involved initial combined review (by J.L.J. and M.M.) of anatomy on multiplane and 3D-reformatted computed tomographic scans of a normal infant pelvis and a physical model of this pelvis generated by a 3D printer. This process was followed by a preliminary trial in which surface models from 3D US were traced by both observers in 15 hips in eight patients ranging from normal to severely dysplastic, separate from the final study, and consensus review of models produced by two authors (J.L.J. and M.M.) to ensure anatomically appropriate appearance. Tracings in the final study were reviewed and approved by the lead radiologist (J.L.J.).
As a reference orientation, we required an initial central plane located in the middle of the acetabulum in each patient. We generated this central plane in a simple way from two landmarks. The observer located the corner point of the acetabular angle in the most anterior and posterior sections in which this angle could be defined, points A and P on Figure 2a. This central plane then was generated automatically perpendicular to the

**Figure 2**: Model demonstrating the central plane and rotated planes. (a) Central plane is obtained through the middle of the acetabulum as the plane perpendicular to the line joining the corner points of the most anterior (A) and posterior (P) positions in which an acetabular angle can be drawn at the midpoint (M) of this line. (b) Planes were cut at angles rotated from the central plane at 5° intervals in axis 1, representing probe tilt around a craniocaudal axis (ie, rotating the base of the transducer from the table toward the ceiling). (c) Planes were similarly cut at 5° intervals around a transverse axis (axis 2), representing probe tilt as if twisting the transducer around a line joining the two hips.
midpoint of line A–P (Fig 2a). Three-dimensional data were resampled along this plane to produce the 2D US image that would result from a 2D probe placed in exactly that location and orientation. From this reference central plane, 2D US images were extracted along orientations rotated in 5° increments from −50° to 50° away, as if slowly rotating a 2D US probe. Images were produced by rotation around axis 1 oriented cranio-caudally (Fig 2b) and axis 2 oriented transversely between hips (Fig 2c).

Each of the resulting 21 images obtained at 5° increments of probe orientation for each axis was reviewed for quality by two observers (J.L.J., a radiologist with dual pediatric and musculoskeletal fellowship training and 10 years of imaging experience [observer 1], and M.M., a biomedical engineering graduate student preparing a thesis on US for hip dysplasia, trained and supervised by J.L.J. [observer 2]). Image quality was scored on the basis of the number of major Graf standard plane criteria met (maximum, four): flat horizontal iliac wing, labrum visible, os ischium present, and midportion of femoral head visible. As a nearly spherical structure, the midpoint of the femoral head is best determined by observing that the image saved is that in which the head has the largest possible radius. We rated the femoral head acceptable on an image if it was plausible from review of that image that the middle one-third of the femoral head was included, regardless of whether other 3D US images of the region (which would not be available to a radiologist reporting 2D US findings) might show a larger femoral head radius. Only 18% (428 of 2352) of the images reviewed had a quality score of 4 from both observers. On each of these, the observers separately measured the alpha angle between acetabular roof and iliac lines, and these observers were blinded to the alpha angle measured on other images of that hip, as well as to each other’s work and to clinical data. For intraobserver variability, the alpha angle was remeasured by the same observer on each of 118 images obtained in this fashion in a subset of 15 hips, 1 month after the initial reading session, and this observer was blinded to previously measured alpha angles.

Statistical Analysis

Descriptive statistics were recorded as the mean ± standard deviation. We computed differences in alpha angles between observers and between observations at various plane orientations. We characterized variability in alpha angles as the upper limit of agreement in these tests (7). For interobserver variability (assessment of the same image by different observers), this was the same as the repeatability coefficient (1.96 multiplied by the standard deviation), as also measured by Gwynne Jones et al (5), whereas for interscan variability, this was the mean difference plus 1.96 multiplied by the standard deviation (ie, the upper limit of the 95% confidence interval [CI] for the alpha angle). Interobserver variability was assessed by using mean difference and intraclass correlation coefficients calculated by using a two-way mixed effects model. We used univariate analysis of variance to compare means in dysplastic and normal hips. Statistics were calculated with software (SPSS 20; SPSS, Chicago, Ill).

Results

Section Orientation

In all 56 hips, at least one image had an acceptable quality score of 4. The range of simulated probe orientations providing images with a quality score of 4, indicating that four main features required in a Graf standard-plane image are present (flat horizontal iliac wing, labrum visible, os ischium present, and midportion of femoral head visible), was 24° ± 10 (standard deviation) in axis 1 (maximum, 45°) and 23° ± 8.4 in axis 2 (maximum, 40°). The image with the central plane had a quality score of 4 in 38 (68%) of 56 hips and was within a 10° rotation from an image with a quality score of 4 in 50 (89%) of 56 hips.

Variability of Alpha Angle with Section Orientation

The alpha angles recorded at each hip by the radiologist across all images with a quality score of 4 varied by a mean of 6.9° ± 4.8 in different probe orientations along axis 1 (range, 0°–21°), 10.8° ± 7.1 in axis 2 (range, 0°–31°); combining both axes, the angles varied by a mean of 13.9° ± 7.1 (upper limit of 95% CI, 28°). Ranges were similar for observer 2. The proportion of this variability occurring for alpha angles greater than 60° may not be clinically relevant because those angles are considered normal. When we considered any measured alpha angle greater than 60° to be equal to 60°, the remaining variability was 3.5° ± 6.9 (range, 0°–24°; upper limit of 95% CI, 19°) across the two axes. For nearly all hips and all axes, changes in virtual probe orientation produced 2D images showing visually substantial variation in the acetabular shape corresponding to the numerical variation in alpha angle. For example, the hip in Figure 3 had alpha angles measured at 52°–70° on images with a quality score of 4, depending on the section orientation, an 18° range that was slightly higher than the mean range of 13.9°.

To assess the clinical significance of this variation, we considered whether observed differences in alpha angle caused by simulated probe rotation led to a change in the Graf diagnostic category. Considering only 3D US images with a quality score of 4, 26 (46%) of 56 hips had images with findings within only one Graf category; 20 (36%) of 56 hips had images with findings in two Graf categories (Fig 4); nine (16%) of 56 hips had images with findings in three Graf categories (Fig 5); and one (2%) had images with findings in all four Graf categories. Overall, 54% of hips had findings in more than one Graf category. Changes in diagnosis were significantly more common at younger ages: Findings on images in 18 (72%) of 25 hips scanned in infants younger than 31 days were classified in two or more Graf categories versus findings on images in just nine (29%) of 31 hips scanned in older infants (P = .002).
Variability of Alpha Angle between and within Observers

Four hundred twenty-eight images had a quality score of 4 and were of acceptable quality in 56 hips. In these images, interobserver variation in alpha angle was a mean of 0.1° ± 5.1° (repeatability coefficient, 10°; coefficient of variation, 0.048). The intraclass correlation coefficient was 0.89. Alpha angles recorded by the two observers were within 5° and 10° of each other in 73% (313 of 428) and 97% (415 of 428) of cases, respectively. Intraobserver variability was lower, with the mean difference between alpha angles of 1.8° ± 3.7° (repeatability coefficient, 7.4°) and measurements within 5° and 10° of each other in 82% (351 of 428) and 96% (411 of 428) of cases, respectively.
Normal versus Dysplastic Hips

We observed clinically normal hips that appeared dysplastic in some planes (Fig 3) and clinically dysplastic hips that had normal 60° alpha angles in certain planes, either with the acetabular contours appearing grossly abnormal (Fig 5) or essentially normal (Fig 6). The alpha angles we observed at 3D US, averaged from acceptable images across all simulated probe orientations in both axes, clearly indicated (P < .001) clinical diagnostic categories: 3D US images in dysplastic hips requiring treatment (category 2; n = 13) had a mean alpha angle of 49.5° versus that in images in initially borderline hips requiring follow-up but no treatment (category 1; n = 23; mean angle, 63.2°) and that in images in hips not requiring treatment or follow-up (category 0; n = 20; mean angle, 71.4°) (Fig 7).

Discussion

We used 3D US technology in infants with normal and dysplastic hips to examine the potential alpha angle interscan variation solely caused by US probe orientation. Our most important
finding was that changing probe orientation altered the alpha angle enough to lead to a change in the Graf diagnostic category in more than one-half of the imaged hips overall and in nearly three-quarters of hips imaged in patients younger than 1 month. We also found that a high-quality image satisfying Graf criteria could be obtained at orientations varying by an average of 24°, or in some patients by as much as 45°, in each axis. This finding means that the sonographer often has a relatively wide range of probe positions in which acceptable images can be generated and must rely on skill and experience to select the best image to save within this range. The interscan variability of alpha angles measured on 2D US images in those different probe orientations was substantial (upper limit of 95% CI: 19°, 28°) and higher than interobserver variability (repeatability coefficient, 10°).

DDH diagnosis from the alpha angle measured on a standardized 2D US image is based on the assumption that this is a one-of-a-kind clinically representative image (ie, that if any trained individual performs US in that hip that day, any image obtained containing appropriate landmarks will have similar acetabular contours and alpha angle). The risk of misdiagnosis and overtreatment because of incorrect probe angulation is a recognized hazard of 2D US, with Graf et al (8) even advocating use of a probe guide-frame system to avoid tilt. Still, to our knowledge, investigators in only one previous study directly assessed the effect of changes in US transducer angle on alpha angle measurements (9). They used ex vivo 2D US in five infant cadaver hips in a water bath and showed that the alpha angle in a normal hip could vary from 48° to 65° (ie, from moderately dysplastic to normal) by changing probe position. Our 3D US study results help confirm these findings in vivo and demonstrate that dysplastic hips can be shown to meet normal 60° alpha angles in probe orientations producing images that still meet Graf standard plane criteria. The range of alpha angle variability in our study compares well with data in that study, with an upper limit of agreement of 19° versus a range of 17° in their study (9).

The 19° upper limit of agreement of interscan variability reported here is greater than the approximately 8° reported in a larger clinical study (10). This finding is expected because we included all US images in each hip that met Graf standards, whereas in clinical practice, sonographers are trained to move the probe through a range of orientations to select the best image of the hip, with the steepest possible alpha angle, deepest acetabular fossa, and largest femoral head radius of all possible images meeting Graf standards. A sonographer typically would not select an image at the margins of the acceptable range of probe orientations. However, in a squirming, crying infant, even the most experienced sonographer has a limited opportunity truly to assess all possible imaging planes with the probe. The hazard for the interpreting radiologist or surgeon is that the image may be of diagnostic quality but not actually represent the best image of the acetabulum that could have been obtained. Our study results demonstrate the wide possible range of interscan variability in alpha angle measurement. Careful sonographer training and experience, and thorough radiologist assessment of other factors such as acetabular rounding, coverage, and dynamic hip stability, are critical to minimize the practical effect of this variability, particularly in the youngest infants. This is especially true given the relatively high superimposed interobserver variability in alpha angle measurement once an image is obtained. The reliability of US diagnosis of hip dysplasia is likely considerably greater in clinical practice, when the alpha angle, acetabular coverage, acetabular morphology, and hip stability are assessed together, than when the alpha angle alone is used for diagnosis. Testing this reliability was outside the scope of our study, which focused on the limitations of hinging diagnosis primarily on the alpha angle.

We did not assess acetabular coverage formally in this study. Our preliminary testing of this index showed an even wider range of variation with plane orientation than the alpha angle, as expected, because this index involves the 3D geometric relationship between two bones (femoral head and acetabulum) rather than only the shape of the acetabulum.
intrinsic flaws in using the 2D Graf standard plane as a diagnostic method for DDH. Simply by rotating the US probe, the sonographer could change the diagnostic category for more than one-half of the hips in this study by at least one, and often two, levels (eg, from normal to moderately dysplastic). Because the image on the screen still appears diagnostic, and there is no record of the position of the 2D US probe, this error invisibly degrades the accuracy and reliability of 2D US assessment of DDH. Conceptually, 3D US of DDH could reduce or eliminate this error by removing dependence on the specific plane of image acquisition.

Three-dimensional US ideally would allow firm diagnosis of every hip as normal or dysplastic at the initial scan. Although our study was not designed or powered to define these diagnostic thresholds in 3D US, hips classified in different diagnostic categories showed substantially different mean alpha angles across all 3D US imaging planes. We expect that diagnostic accuracy of US for DDH could be improved in the future by considering this additional acetabular shape information available from 3D US, both visually and by measuring alpha angles and novel indexes being developed.

Our study had limitations. As a pilot study, the sample size was relatively small, and some of our earlier scans included a range of the acetabulum too narrow to be included; however, our data set demonstrates the full range from normal to severely dysplastic hips. The clinical diagnoses and treatment decisions were made by one of five orthopedic surgeons on the basis of their own experience and preferences, and they were not validated by any external reference standard. However, this situation represents typical clinical practice, and our conclusions regarding the variability of US measurements do not depend on the specific diagnosis in each patient. The follow-up interval was relatively short at more than 6 months, but this is the most clinically important period because the decision to begin treating DDH is generally within weeks of initial assessment. The probe movements we assessed in 3D data were limited to pure rotation in one of two planes, and we did not consider combinations of two rotations or rotation and translation. These assessments would have been challenging to perform mathematically, and although they might have revealed further increased variation in alpha angles, they would not alter our conclusion that this variation is an important hazard of 2D US assessment of DDH.

In conclusion, we found that 3D US of the infant hip is feasible and images can be analyzed quantitatively. After an initial flurry of studies in the 1990s (11–14), 3D US was not pursued in hip dysplasia assessment. Those early reports describe slow manual sweeps to produce scans and cumbersome post-processing. The probe we used can help obtain a 3D scan in 3.2 seconds with less difficulty in positioning than in 2D US. In our experience, almost all infants are able to stay still for this brief time, especially if distracted by toys, sucrose, a warmed blanket, and warmed US gel. Computer analysis time per hip is less than 10 minutes. As technology advances, 3D scanning and analysis times probably will continue to shorten, potentially further improving the practical application of this technology in the future.

Disclosures of Conflicts of Interest: J.L.J. disclosed no relevant relationships. M.M. disclosed no relevant relationships. V.G.S. disclosed no relevant relationships. R.B.T. disclosed no relevant relationships. K.C. disclosed no relevant relationships. B.B. disclosed no relevant relationships. R.B.T. disclosed no relevant relationships.

References