

# The reliability of ultrasonography in developmental dysplasia of the hip

## How reliable is it in different hands?

Mehmet Müfit Orak, Tolga Onay<sup>1</sup>, Talat Çağırılmaz<sup>2</sup>, Cenk Eliboğlu<sup>3</sup>, Funda Dinç Eliboğlu<sup>3</sup>, Tuncay Centel<sup>4</sup>

### ABSTRACT

**Background:** Developmental dysplasia of the hip (DDH) is the most common skeletal dysplasia. Two principal methods used in early diagnosis of DDH are clinical examination and ultrasonographic investigation. Dogruel *et al.* found a low specificity of clinical examination in patients with DDH. Additionally, Kamath *et al.* stated that ultrasonography performed by a radiologist in routine clinical practice is more reliable than physical examination performed by the average clinician. In clinical practice, the application and assessment of hip ultrasonography are completed by a single person. This assessment determines the followup of the patient. Thus, hip ultrasonography performed on the same person by different individuals under the same conditions will yield a more accurate assessment of the reliability of ultrasonographic assessment of DDH. Although inter-observer reliability was high in many previous studies of ultrasound image evaluation, reliability rates vary among studies of the application of ultrasonography.

**Materials and Methods:** Inter-examiner reliability of hip ultrasonography was analyzed among four investigators who separately evaluated 100 hips (50 infants). The obtained bone structure angles ( $\alpha$ ), cartilage structure angles ( $\beta$ ), and distribution of hip types were compared among the investigators. All infants were brought to the hospital for a healthy child followup examination, according to the country's health policy. Babies between 0 and 6 months were included in the study. Babies with any neuromuscular disorders, neural tube defects or any type of genetic anomalies were excluded from the study. The study was explained to the families of all infants and written informed consent was obtained.

**Results:** There was a significant difference in the hip type determined by the investigators with respect to  $\alpha$  and  $\beta$  angles ( $P < 0.01$ ,  $P < 0.01$ ,  $P = 0.002$ ). The average alpha measurements of the first orthopedist, second orthopedist, first radiologist, and second radiologist were  $67.38 \pm 6.24$ ,  $65.60 \pm 5.84$ ,  $65.44 \pm 4.59$ , and  $62.59 \pm 4.50$ , respectively. The average beta measurements of the first orthopedist, second orthopedist, first radiologist, and second radiologist were  $53.85 \pm 8.86$ ,  $50.74 \pm 7.80$ ,  $44.77 \pm 6.30$ , and  $44.39 \pm 5.81$ , respectively. Agreement among the results obtained by the clinicians was investigated in dual comparisons. The relative agreement according to the alpha angle ranged from 3.6% to 44.5%, and the relative concordance according to the beta angle ranged from 0.9% to 45.3%. Agreement regarding hip typing was determined to range from 19.1% to 42.6%.

**Conclusion:** Sonographic evaluation of the hip appears to vary depending on the investigator.

**Key words:** Developmental dysplasia of hip, hip ultrasonography, reliability

**MeSH terms:** Hip dysplasia, congenital, hip dislocation, congenital, ultrasonography, ultrasonic diagnosis

Department of Orthopedic and Traumatology, Fatih Sultan Mehmet Education and Research Hospital, <sup>1</sup>Department of Orthopedic and Traumatology, Marmara University Education and Research Hospital, Istanbul, <sup>2</sup>Department of Orthopedic and Traumatology, Biga State Hospital, Çanakkale, <sup>3</sup>Department of Radiology, Muğla University Education and Research Hospital, Muğla, <sup>4</sup>Department of Orthopedic and Traumatology, Istanbul Memorial Hospital, Istanbul, Turkey

**Address for correspondence:** Dr. Tolga Onay, Department of Orthopedic and Traumatology, Marmara University Education and Research Hospital, Istanbul, Turkey. E-mail: [onaytolga@yahoo.com](mailto:onaytolga@yahoo.com)

Access this article online	
Quick Response Code:	Website: <a href="http://www.ijonline.com">www.ijonline.com</a>
	DOI: 10.4103/0019-5413.168753

### INTRODUCTION

Developmental dysplasia of the hip (DDH) is the most common skeletal dysplasia.<sup>1</sup> Two principal methods used in early diagnosis of DDH are clinical examination and ultrasonographic investigation. Dogruel *et al.*<sup>2</sup> found a low specificity of clinical examination in patients with DDH. Additionally, Kamath *et al.*<sup>3</sup> stated that ultrasonography performed by a radiologist in routine clinical practice is more reliable than physical examination performed by the average clinician.

The ultrasonographic investigation technique developed by Graf in 1980 is gaining wide acceptance as a radiological evaluation method in patients with DDH.<sup>4</sup> In Graf's method, DDH is classified into (four) main groups and nine subgroups based on the patient's age and ultrasonographic measurements.<sup>5</sup>

Many treatment algorithms have been developed based on Graf's DDH classification system. Ultrasonography has a highly important role in screening for DDH.<sup>6-9</sup> Therefore, the reliability of ultrasonographic assessment of the hips has been investigated in many studies. Most of these studies involved ultrasonographic hip images taken by a single person but measured and interpreted by various individuals.<sup>10-13</sup>

In clinical practice, the application and assessment of hip ultrasonography are completed by a single person. This assessment determines the followup of the patient. Thus, hip ultrasonography performed on the same person by different individuals under the same conditions will yield a more accurate assessment of the reliability of ultrasonographic assessment of DDH. However, very few studies have been designed in this way.<sup>14-17</sup>

The aim of the present study was to investigate the inter-examiner reliability of the results of hip ultrasonography, performed and assessed by different clinicians for the diagnosis of DDH in the same hip.

## MATERIALS AND METHODS

This prospective study was approved by the local ethics committee. Fifty infants from 0 to 6 months of age who presented to the hospital from September 2012 to December 2012 were included in the study. All infants were brought to the hospital for a healthy child followup examination, according to the country's health policy. Babies between 0 and 6 months were included in the study. Babies with any neuromuscular disorders, neural tube defects or any type of genetic anomalies were excluded from the study. The study was explained to the families of all infants and written informed consent was obtained. Both hips of each infant were sonographically assessed by four different clinicians (two radiologists and two orthopedists) using the method described by Graf.<sup>4</sup>

A consensus meeting regarding the Graf method was held before the study began. Ultrasonographic investigations were performed on a special hip ultrasonography table while supporting the infants in the lateral decubitus position. All ultrasonography examinations were performed by using a 5-MHz linear transducer and the

same ultrasound device (LOGIQ P5; GE Healthcare, US). According to Graf, standard images were obtained in coronal plane. Each examiner obtained a standard plain image according to Graf's description. The deepest point of the acetabulum, the lower iliac margin at the triradiate cartilage, the labrum and the chondroosseous border of the proximal femur were identified as reference points.<sup>18</sup> The images were printed on papers after being obtained. In total, 100 hips of 50 infants were evaluated. Four ultrasonographic investigations of each hip were performed for a total of 400 ultrasonographic investigations. The two orthopedists (T.O. and T.Ç.) participating in the study completed their orthopedic residencies in the same clinic and were extensively involved with the management of pediatric orthopedic problems. Similarly, the two radiologists (C.E. and F.E.) participating in the study completed their residency training in the same clinic. All investigators had experience with >500 hip ultrasonography examinations. Each measurement was performed successively with a maximum 2-min interval between measurements. Each clinician performed his or her own ultrasonographic examination alone and freely. Measurements were performed manually. The 'baseline' is the line of the ilium as it intersects the bony and cartilaginous portions of acetabulum. The 'inclination line' is the line along the margin of the cartilaginous acetabulum. The third is the 'acetabular roofline' along the bony roof. The intersection of the roofline and the baseline forms the 'alpha angle', whereas the intersection of the inclination line and the baseline forms the 'beta angle'. The alpha and beta angles of the hip joints were recorded and the hips were typed according to Graf's classification system.

## Statistical analysis

The number cruncher statistical system 2007 and the power analysis and sample size 2008 statistical software (329 North 1000 East Kaysville, Utah 84037 USA) programs were used for the statistical analysis. Quantitative study data were compared and descriptive statistical methods were used (mean, standard deviation, median, frequency, and the ratio). Variance analysis in repeated measurements was used for followup measurements of the data, and Bonferroni-adjusted tests were used for pairwise

**Table 1: Averages according to the alpha and beta measurements of the investigators**

Investigator	Mean±SD	
	Alpha measurements	Beta measurements
Orthopedist-1	67.38±6.24	53.85±8.86
Orthopedist-2	65.60±5.84	50.74±7.80
Radiologist-1	65.44±4.59	44.77±6.30
Radiologist-2	62.59±4.50	44.39±5.81

SD=Standard deviation

comparisons. The intraclass correlation test was used to determine the level of correlation between the repeated measurements. Cohen's kappa test was used to assess the levels of agreement among the hip types. Statistical significance was defined as  $P < 0.05$  and  $P < 0.01$ .

## RESULTS

One hundred hips of 50 infants were evaluated (23 [46.0%] female, 27 [54.0%] male). Their ages ranged from 1 to 179 days (mean:  $64.36 \pm 54.47$  days).

The average alpha measurements of the first orthopedist, second orthopedist, first radiologist, and second radiologist were  $67.38 \pm 6.24$ ,  $65.60 \pm 5.84$ ,  $65.44 \pm 4.59$ , and  $62.59 \pm 4.50$ , respectively. A highly statistically significant difference was found in the distribution of the alpha measurements among the four clinicians ( $F = 46.65$ ) (repeated-measures test,  $P < 0.01$ ) [Table 1].

The average beta measurements of the first orthopedist, second orthopedist, first radiologist, and second radiologist were  $53.85 \pm 8.86$ ,  $50.74 \pm 7.80$ ,  $44.77 \pm 6.30$ , and  $44.39 \pm 5.81$ , respectively. A highly statistically significant difference was found in the distribution of the beta measurements among the four clinicians ( $F = 60.89$ ) (repeated-measures test,  $P < 0.01$ ) [Table 1].

The hips were divided into four subgroups according to Graf's classification system as follows: Mature (tip 1a-1b), immature (tip 2a), minor dysplasia (tip 2b-2c-D), and major dysplasia (3a-3b-4). Investigation of the hip types according to this classification system revealed the following. Of the 100 hips, the first orthopedist typed 94 (94.0%) as mature, 4 (4.0%) as immature and 2 (2.0%) as having minor dysplasia. The second orthopedist typed 89 (89.0%) as mature, 10 (10.0%) as immature and 1 (1.0%) as having minor dysplasia. The first radiologist typed 96 (96.0%) as mature and 4 (4.0%) as immature. Finally, the second radiologist typed 84 (84.0%) as mature, 15 (15.0%) as immature and 1 (1.0%) as having minor dysplasia [Table 2].

In total, 75 hips were evaluated as mature (Ia/Ib) by all of the examiners. No hips were evaluated as immature (IIa) or as having minor dysplasia (IIb/IIc/D) by all of the

examiners. In addition, none of the examiners found any major dysplasia (IIIa/IIIb/IV). There was a statistically significant difference in the hip typing among the four physicians (Cochran's  $Q = 14.66$ ,  $P = 0.002$ ).

Agreement among the results obtained by the clinicians was investigated in dual comparisons. The relative agreement according to the alpha angle ranged from 3.6% to 44.5%, and the relative concordance according to the beta angle ranged from 0.9% to 45.3%. Agreement regarding hip typing was determined to range from 19.1% to 42.6% [Table 3].

## DISCUSSION

Developmental dysplasia of the hip is one of the most common skeletal dysplasias of childhood and frequently causes permanent disability if the diagnosis is delayed or treatment is not performed.<sup>19</sup> Ultrasonography for DDH, which was popularized by Graf in the 1980s, has gained wide acceptance and now plays an important role in the early diagnosis of DDH.<sup>4</sup> Harcke and Kumar<sup>20</sup> emphasized that the person who performs the ultrasonographic examination can achieve the necessary basic skills and techniques after performing at least 100 ultrasonography procedures. Graf<sup>18</sup> suggested that the method is tied to the established standards, with definite rules, and is independent of repeatable experience and skill. In the present study, each specialist had had experience with >500 hip ultrasonography procedures. This experience is sufficient for the application of Graf's method.

The present technique is widely used and has been evaluated for reliability by many investigators. However, previous studies commonly involved ultrasonography performed by one person but interpreted by different clinicians. It was shown that different results can be obtained even on the same ultrasonography image.<sup>21,22</sup> In such studies, the reliability of Graf hip ultrasonography was reportedly poor to good.<sup>10-13</sup> Although various authors have reported that the performance of ultrasonography may vary depending on the individual,<sup>23-25</sup> very few studies have examined the reliability in this regard. Bar-On *et al.*<sup>14</sup> stated that in the same ultrasonographic examination, intra-observer agreement was substantial while inter-observer agreement

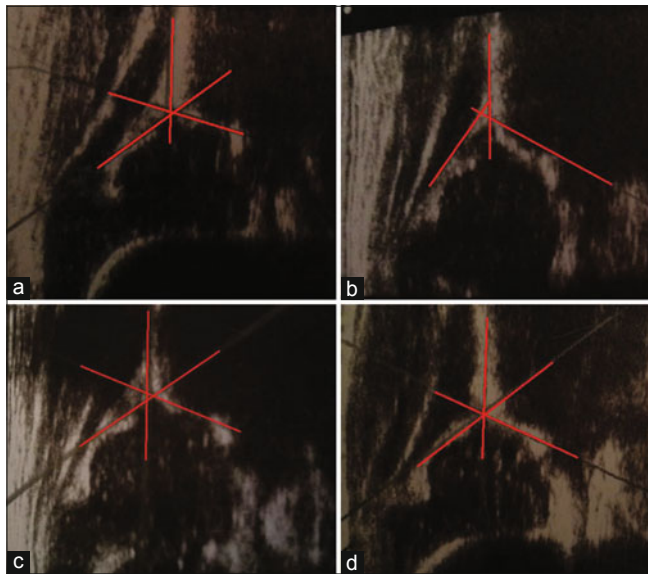
**Table 2: Distribution of the types**

Hip type	n (%)			
	Orthopedist-1	Orthopedist-2	Radiologist-1	Radiologist-2
I (normal)	94 (94)	89 (89)	96 (96)	84 (84)
IIa (immature)	4 (4)	10 (10)	4 (4)	15 (15)
IIb/IIc/D (minor dysplasia)	2 (2)	1 (1)	0	1 (1)
IIIa/IIIb/IV (major dysplasia)	0	0	0	0

**Table 3: Assessments of alpha and beta measurements and dual agreement of the physicians according to the types**

Investigator	Alpha measurement		Beta measurement		Hip types	
	Interclass correlation	P	Interclass correlation	P	Cohen's kappa	P
Orthopedist-1-orthopedist-2	0.445	0.001**	0.156	0.060	0.426	0.001**
Orthopedist-1-radiologist-1	0.148	0.070	0.453	0.001**	0.370	0.001**
Orthopedist-1-radiologist-2	0.312	0.001**	0.320	0.001**	0.363	0.001**
Orthopedist-2-radiologist-1	0.036	0.361	0.009	0.536	0.251	0.011*
Orthopedist-2-radiologist-2	0.324	0.001**	0.025	0.598	0.191	0.051
Radiologist-1-radiologist-2	0.119	0.118	0.080	0.214	0.252	0.001**

\*P&lt;0,05, \*\*P&lt;0,01

**Figure 1:** (a-d) Four different ultrasonography images of the same hip, which are interpreted by 4 different examiner. Standard plain was obtained for all ultrasonography images. Alpha angle is measured 71, 57, 68 and 72, while beta angle is measured 52, 49, 55, and 54, respectively

was moderate. However, the reliability and agreement were markedly low in ultrasonographic examinations performed by two different people on 150 hips. The mean intra-observer reliability coefficient was 0.29 (-0.03 to +0.49), and the mean inter-observer coefficient was 0.28 (0.12–0.55).<sup>14</sup> Rosendahl *et al.*<sup>15</sup> reported low agreement in ultrasonography performed by two different people. Roovers *et al.*<sup>16</sup> reported an agreement rate of 94.8% in hip ultrasonography examinations performed and assessed by two different examiners on 48 hips of 24 patients. Peterlein *et al.*<sup>17</sup> performed a similar study, in which ultrasonography was performed on each newborn by three investigators with different levels of experience. Interestingly, they found no statistically significant difference between investigators measurements even when the experience levels were highly different from each other.

In our study, statistically significant differences were found among the results of the four investigators. The agreement among their results was investigated in dual comparisons. An agreement rate of 3.6–44.5% was determined according to the alpha angle, 0.9–45.3%

according to the beta angle, and 19.1–42.6% according to the hip type. Although Graf suggested that the method is tied to definite standards and rules and is independent of personal experience, the differences in agreement among the investigators in our study seem to support the idea that hip ultrasonography is dependent on the individual. In previous studies, it was shown that the disagreement was higher for beta angle and measurement in pathological hips.<sup>22</sup> Our present study also had higher disagreement for beta angle.

The limitations of study are relatively low number of pathological hips and sample size. However, the strength is the performance of the ultrasonography evaluation of the same infant by four different clinicians. The few similarly designed studies that have been performed in the past involved much fewer<sup>14-17</sup> clinicians.

Different results can be obtained even on same ultrasonography image. Copuroglu *et al.*<sup>21</sup> also suggested that a major reason that made the difference between the observers was to find the correct landmarks for measuring the angles on an ultrasonographic image. He emphasized that the observer had problems to identify the anatomical structures or they did not handle the correct definitions.

When the ultrasonography images in our study were evaluated, it was seen that standard plain was obtained for all images [Figure 1a-d]. However, these images are not an exact replica of each other. But, the most important reason of different results was the choice of the diverse reference point when the alpha and beta angles were measured. A large number of self performed examinations and training in potential mistakes may improve ultrasonographic measurements.<sup>22</sup>

## CONCLUSION

Ultrasonography of normal hips has low inter-observer reliability. It should be kept in mind that ultrasonographic evaluation in the followup and treatment of DDH may vary, depending on the practitioner.

## REFERENCES

1. Loder RT, Skopelja EN. The epidemiology and demographics of hip dysplasia. *ISRN Orthop* 2011;2011:238607.
2. Dogruel H, Atalar H, Yavuz OY, Sayli U. Clinical examination versus ultrasonography in detecting developmental dysplasia of the hip. *Int Orthop* 2008;32:415-9.
3. Kamath S, Mehdi A, Wilson N, Duncan R. The lack of evidence of the effect of selective ultrasound screening on the incidence of late developmental dysplasia of the hip in the Greater Glasgow Region. *J Pediatr Orthop B* 2007;16:189-91.
4. Graf R. The diagnosis of congenital hip-joint dislocation by the ultrasonic Comboud treatment. *Arch Orthop Trauma Surg* 1980;97:117-33.
5. Graf R. Classification of hip joint dysplasia by means of sonography. *Arch Orthop Trauma Surg* 1984;102:248-55.
6. Bialik GM, Eidelman M, Katzman A, Peled E. Treatment duration of developmental dysplasia of the hip: Age and sonography. *J Pediatr Orthop B* 2009;18:308-13.
7. Demirhan M, Dikici F, Eralp L, Onen M, Göksan B. A treatment algorithm for developmental dysplasia of the hip for infants 0 to 18 months of age and its prospective results. *Acta Orthop Traumatol Turc* 2002;36:42-51.
8. Eidelman M, Katzman A, Freiman S, Peled E, Bialik V. Treatment of true developmental dysplasia of the hip using Pavlik's method. *J Pediatr Orthop B* 2003;12:253-8.
9. Guille JT, Pizzutillo PD, MacEwen GD. Development dysplasia of the hip from birth to six months. *J Am Acad Orthop Surg* 2000;8:232-42.
10. Roposch A, Graf R, Wright JG. Determining the reliability of the Graf classification for hip dysplasia. *Clin Orthop Relat Res* 2006;447:119-24.
11. Simon EA, Saur F, Buerge M, Glaab R, Roos M, Kohler G. Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the Graf method. *Swiss Med Wkly* 2004;134:671-7.
12. Omeroglu H, Biçimoglu A, Koparal S, Seber S. Assessment of variations in the measurement of hip ultrasonography by the Graf method in developmental dysplasia of the hip. *J Pediatr Orthop B* 2001;10:89-95.
13. Dias JJ, Thomas IH, Lamont AC, Mody BS, Thompson JR. The reliability of ultrasonographic assessment of neonatal hips. *J Bone Joint Surg Br* 1993;75:479-82.
14. Bar-On E, Meyer S, Harari G, Porat S. Ultrasonography of the hip in developmental hip dysplasia. *J Bone Joint Surg Br* 1998;80:321-4.
15. Rosendahl K, Aslaksen A, Lie RT, Markestad T. Reliability of ultrasound in the early diagnosis of developmental dysplasia of the hip. *Pediatr Radiol* 1995;25:219-24.
16. Roovers EA, Boere-Boonekamp MM, Geertsma TS, Zielhuis GA, Kerkhoff AH. Ultrasonographic screening for developmental dysplasia of the hip in infants. Reproducibility of assessments made by radiographers. *J Bone Joint Surg Br* 2003;85:726-30.
17. Peterlein CD, Schüttler KF, Lakemeier S, Timmesfeld N, Görg C, Fuchs-Winkelmann S, *et al*. Reproducibility of different screening classifications in ultrasonography of the newborn hip. *BMC Pediatr* 2010;10:98.
18. Graf R. The use of ultrasonography in developmental dysplasia of the hip. *Acta Orthop Traumatol Turc* 2007;41 Suppl 1:6-13.
19. 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:S211-5.
20. Harcke HT, Kumar SJ. The role of ultrasound in the diagnosis and management of congenital dislocation and dysplasia of the hip. *J Bone Joint Surg Am* 1991;73:622-8.
21. Copuroglu C, Ozcan M, Aykac B, Tuncer B, Saridogan K. Reliability of ultrasonographic measurements in suspected patients of developmental dysplasia of the hip and correlation with the acetabular index. *Indian J Orthop* 2011;45:553-7.
22. Hell AK, Becker JC, Rühmann O, Lewinski Gv, Lazovic D. Inter-and intraobserver reliability in Graf's sonographic hip examination. *Z Orthop Unfall* 2008;146:624-9.
23. Wientroub S, Grill F. Ultrasonography in developmental dysplasia of the hip. *J Bone Joint Surg Am* 2000;82-A: 1004-18.
24. Portinaro NM, Pelillo F, Cerutti P. The role of ultrasonography in the diagnosis of developmental dysplasia of the hip. *J Pediatr Orthop* 2007;27:247-50.
25. Clinical practice guideline: Early detection of developmental dysplasia of the hip. Committee on Quality Improvement, Subcommittee on Developmental Dysplasia of the Hip. American Academy of Pediatrics. *Pediatrics* 2000;105:896-905.

**How to cite this article:** Orak MM, Onay T, Cagirmaz T, Elibol C, Elibol FD, Centel T. The reliability of ultrasonography in developmental dysplasia of the hip How reliable is it in different hands?. *Indian J Orthop* 2015;49:610-4.

**Source of Support:** Nil, **Conflict of Interest:** None.

## Staying in touch with the journal

### 1) Table of Contents (TOC) email alert

Receive an email alert containing the TOC when a new complete issue of the journal is made available online. To register for TOC alerts go to [www.ijoonline.com/signup.asp](http://www.ijoonline.com/signup.asp).

### 2) RSS feeds

Really Simple Syndication (RSS) helps you to get alerts on new publication right on your desktop without going to the journal's website. You need a software (e.g. RSSReader, Feed Demon, FeedReader, My Yahoo!, NewsGator and NewzCrawler) to get advantage of this tool. RSS feeds can also be read through FireFox or Microsoft Outlook 2007. Once any of these small (and mostly free) software is installed, add [www.ijoonline.com/?rssfeed.asp](http://www.ijoonline.com/?rssfeed.asp) as one of the feeds.