Peer reviewed article

Inter-observer agreement of ultrasonographic measurement of alpha and beta angles and the final type classification based on the Graf method

E. A. Simon[a], F. Saur[b], M. Buerge[c], R. Glaab[d], M. Roos[e], G. Kohler[a]

[a] Department of Orthopaedics/Traumatology Cantonal Hospital, Frauenfeld, Switzerland
[b] Paediatrician, Frauenfeld, Switzerland
[c] Department of Radiology, Cantonal Hospital, Frauenfeld, Switzerland
[d] Department of Surgery, Cantonal Hospital, Frauenfeld, Switzerland
[e] University of Zürich, ISPM, Department of Biostatistics, Zürich, Switzerland

Summary

Background: The aim of our study was to investigate the agreement of the assessment of hip ultrasonograms by different observers.

Methods: In the period from June 3rd to December 9th 2002, four different (by experience and field) groups of observers rated all first time sonograms obtained in our hospital. The results in terms of angle and type classification were compared.

Results: 158 ultrasonographic images were evaluated. The inter-observer agreement for the classification "normal" (type I) versus "abnormal" (types IIa+ to IV) varied from 83% to 90% with kappa coefficients which indicated moderate (kappa 0.55) to substantial (kappa 0.71) inter-observer agreement. For one pair of observers, a better agreement could be demonstrated for the assessment of immature hips than for mature ones. The deviation for the α-angle was 0 to 16° with a standard deviation of 3.15° (95% CI 2.95, 3.37), and for the β-angle 0 to 26° with a standard deviation of 6.1° (95% CI 5.7, 6.5). The intra-class correlation coefficient was estimated to be 0.72 and 0.34 for the alpha and beta angles respectively. If the hip was immature there was no increase in the discrepancy in assessment between observers. The least agreement existed between the less experienced and the most experienced. It has not been possible to make a statement on the discrepancy with regard to initial signs of instability or decentralization of the hip joints because of the small number of hips of this type.

Conclusions: Although the spread in measured α- and β-angles is large, the inter-observer agreement for the classification showed good results. No disagreement occurred in the diagnosis of normal vs. dysplastic hips, so no severe cases have been missed. The experience and training of the investigators seemed to play an important role with regard to variability and agreement. The agreement in the assessment of immature hips was better than that of mature hips. Therefore, ultrasound examination of infant hips would appear to be a trustworthy screening method.

Key words: ultrasound; hip dysplasia; screening; inter-observer agreement

Introduction

Developmental dysplasia of the hip is characterized by an abnormal relation between the femoral head and the acetabulum. These abnormalities range from dysplasia alone, where the anatomy of the developing articulation surfaces is abnormal, to a potential partial and complete dislocation of the hip.

Developmental dysplasia of the hip occurs at a rate of 2–4% making it the most frequent abnormality of the locomotor apparatus [1, 2]. For this reason, there has always been the need to search for a method to recognize this abnormality early so that adequate treatment can be commenced. The aim of thorough screening is the early detection of immature, unstable, decentralized and dislocated hip joints, i.e. all hips that are at risk of dysplasia or dislocation and, therefore, of early attrition in later life [3]. Apart from the initial finding, it is crucial to the outcome that treatment commences as early as possible. Treatment be-

None of the authors has an affiliation or financial agreement with any company.
We did not receive any financial support for the making of this study.

comes more intensive, longer and more invasive the older the patient is at the time of diagnosis [4–7].

Since the neonatal hip joint cannot be definitively assessed by conventional radiology (the femoral nucleus of ossification is not yet fully developed), clinical and additional ultrasound examinations are performed, if possible, in the first few days after birth [8]. Harcke, Terjesen and Graf [17], have formulated different ultrasound examination techniques for the hips of newborns and infants, with Graf's method being the most established in Europe.

In the literature, the range of deviations reported for measurement of angles based on coronal ultrasonograms according to Graf is broad [5, 9–11]. Furthermore, there is still insufficient evidence of agreement in terms of classification between observers [7, 12–16]. Therefore, there is much debate in literature about the value of Graf's method. Part of the discussion is that ultrasound screening is only reliable if the examiners are experienced because of the challenging technique.

For this reason we initiated our study with the aim of representing, in the context of quality assurance, the agreement of assessments between different observers.

Patients and methods

Duration / Groups of observers

In the period ranging from 3. 6. 2002 to 9. 12. 2002, sonograms obtained by one sonographer were evaluated by four groups of observers; they were not present when the images were obtained. The observer groups were blinded in terms of anamnesis and clinical examination of the infants. Since different teams in our hospital evaluate hip sonographs, variabilities would have clinical impact. The groups were comprised of a team from the radiology department under the supervision of a radiologist with FMH permit for sonography of infant hips, a paediatrician, likewise with a permit, an orthopaedic specialist with a permit, and a registrar for orthopaedics without a permit. To obtain a FMH permit for sonography of infant hips it is compulsory to attend three courses in hip ultrasonography, including a final examination as well as presenting a certain number of investigations. After two and five years, a recertification is necessary.

The registrar in orthopaedics and the registrars in radiology had been trained by their team according to Graf's principles.

Sonography

The infant was correctly positioned in a standard positioning tray. The sonograms were obtained by the same qualified and certified sonographer using an ATL HDI 5000 Philips sonography device with linear probe and an ultrasound frequency of 7.5–10 MHz. The sonographic images were based on the standardized methodological criteria according to Graf [17].

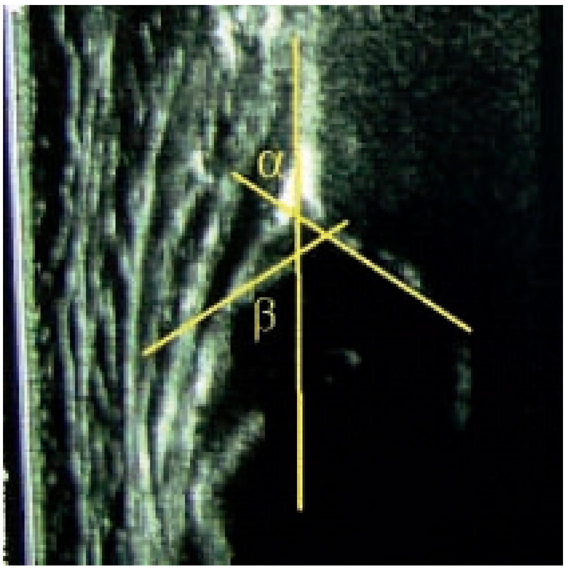


Figure 1

Example of assessment with alpha and beta angles marked on a sonogram.

Classification

Classification was performed according to the Graf method. The "bone angle" alpha (α) and the "cartilage angle" beta (β) were traced onto separate sheets of tracing paper and measured (figure 1) so that no residues could be seen on the sonogram. The examiners were familiar with the classification although there were differences in the levels of training. Entries were made on a standardized, pre-printed form and included information as name, age, side, α- and β-angles, hip type, bony roof, superior bony rim and cartilaginous acetabular roof.

As in the Dutch studies by Roovers et al. [5] the original 10 groups were suitably combined to form 4 main groups for screening purposes: type I = normal, types IIa/IIa^+/IIa^- = immature, types IIb/IIc/D = dysplasia, types IIIa/IIIb/IV = dislocation. The angles of types IIa/IIa^+/IIa^- matched so they were combined into one group for this study. If it had been for the purpose of prognosis and therapy this would not have been possible because between types IIa^+ and IIa^- is the cut off line concerning a necessary therapy.

Later, division into two groups was performed – normal (type I) and abnormal (types IIa^+ up to IV) in order to calculate inter-observer agreement.

Demographic data

During the period ranging from 3. 6 to 9. 12. 2002 all infants who were assigned to the Department of Radiology for an ultrasound assessment were included. Ninety percent had prior clinical and ultrasound assessment in the first days of life.

Sonographic records were obtained as follows: 162 hip sonographs/81 infants, 55 (68%) girls and 26 (32%) boys. The ages of our patients at the time of investigation ranged from 11–188 days (median of 41 days). Both sides were investigated. Sonographic follow up images were excluded. Four sonograms were excluded because they did not fulfil the standardized methodological criteria according to Graf.

Statistics

Inter-observer agreement was estimated by calculating the proportion of agreement (%) and Cohen's kappa with the corresponding 95% confidence interval (95% CI) and the index of average positive (p_{pos}) and negative (p_{neg}) agreement [33, 34]. Cohen's kappas were interpreted according to Landis and Koch [20, 21]. A kappa index of +1.0 indicates complete correlation and 0.0 indicates pure chance [22]. It is known that Cohen's kappa is difficult to interpret as a valid measure and is not undisputed [20, 32–34]. This is why it should be complemented by the index of average positive (p_{pos}) and negative (p_{neg}) agreement as suggested in the literature [33, 34].

For every child both joints were examined. Each joint was evaluated by four observer groups, which made their measurements on the alpha and beta angles independently of each other. It is a nested design as each child has two joints and each joint had one ultrasonogram done which contributes to four measurements of an angle. Such a design can be analysed by applying the hierarchical ANOVA with random effects [18]. The merits of this method are twofold. First, one is able to estimate the residual spread of the measurements together with the corresponding 95% CI. Second, one can compute the intra-class correlation coefficient ICC which is known to measure the level of inter-observer agreement. ICC describes the level of dependence of the measurements of the four observer groups at the same hip within a child.

A question came up: did the measurements of the observers agree better on hips which were healthy than on disordered ones? It was supposed that disordered hips are more difficult to evaluate and consequently the measurements of the four observer groups tend to disagree on them. Therefore, for every child a distinction was made between disordered hip (side 2) and a healthy one (side 1).

In order to investigate the agreement more profoundly we used the method suggested by Bland and Altman [25, 29]. Consider figure 3 where plots for the alpha angle measurements by the registrar (AWS2R3) and a well-trained paediatrician (AWS2R4) on side 2 (disordered hip) are given. First, Bland and Altman suggested plotting the measurements of both observers against each other, together with the identity line representing the perfect agreement. Second, they suggested plotting the difference between the measurements of both observers (AWS2R3–AWS2R4) against their mean (AWS2R3+AWS2R4)/2, together with the estimated bias and the limits of agreement. The bias between two observers can be estimated by the mean difference of their measurements together with the corresponding 95% CI (table 5). If the mean difference is significantly different from zero (95% CI does not include 0), there is a bias between observers. In our case the registrar tended to measure smaller alpha angles than did the paediatrician, as the mean difference is negative. Moreover, Bland and Altman suggested the computation of limits of agreement (low and up) together with their 95% CI.

Another way of evaluating the agreement between two raters is by computating the concordance correlation coefficient (CCC) as suggested by Lin [27, 28]. He gave some arguments why the usual Pearson correlation coefficient alone is inappropriate to describe the agreement between two observers. Instead, one can correct it in an appropriate way to obtain the concordance correlation coefficient. The correction accounts for both the bias (location shift) between observers and the difference in dispersion (scale shift) of their measurements. We obtained the estimates for CCC together with the corresponding 95% CI.

Results

General

The sonograms of two infants had to be excluded from the evaluation because the quality of the images was too poor.

Two sonograms showing severe pathological conditions were looked at separately for the analysis of angle measurements. A 24° difference in an α-angle, as measured by the orthopaedist and the registrar, remained an isolated case, this hip was assessed by all observers as pathological. There were more girls (68%) in the study than boys (32%). The left hip was identified as immature or more severely affected in 58% of the cases.

Inter-observer agreement of the Graf classification

The distribution of diagnoses is shown in figure 2. The division of the hips into I = normal, IIa/IIa⁺/IIa = immature, IIb/IIc/D = dysplasia, IIIa/IIIb/IV= dislocation is shown in table 1. The variability of the dichotomy normal vs. abnormal was 75.3% (119/158 sonograms) in all observer groups per child. The inter-observer agreement per observer pair was determined with reference to the dichotomy: normal (I) and abnormal (IIa⁺ to IV), an example of which is shown in table 2. The proportion of inter-observer agreement for the pairs of observers for the classification normal (type I) versus abnormal (IIa⁺ to IV) varied from 83% to 90%, see table 3. There was a difference between the right (fair to moderate) and the left sides (moderate to substantial) in terms of kappa values. Overall, the inter-observer agreement between the pairs of observers was good with regard to the classification of normal (type I) versus abnormal (IIa⁺ to IV), see table n. The best agreement at 90%, kappa coefficient 0.7 (95% CI 0.68; 0.70), with fairly good p_{pos} = 0.94 and p_{neg} = 0.72 existed between the orthopaedist and the paediatrician. The least agreement at 83%, kappa 0.5 (95% CI 0.53; 0.58) with fairly good p_{pos} = 0.89 and poor p_{neg} = 0.6 existed between the registrar and the paediatrician, see table 3.

The agreement between the radiology team and the orthopaedist in the diagnosis of immature hips was better CCC = 0.5 (95% CI 0.32; 0.65) than for normal hips CCC = 0.29 (95% CI 0.08; 0.47). A similar tendency was observed for other pairs of observers.

No disagreement occurred in the diagnosis of normal vs. dysplastic hips, so that no severe cases have been missed.

Intra-class correlation coefficient for the assessment was in the middle range with ICC = 0.6.

Inter-observer agreement of the alpha and beta angles

The angle values were compared by four groups of observers. The variability for the α-angle was between 0° and 16°, and for the β-angle it was between 0° and 26°. The standard deviation for the observers was lower for the α-angle than for the β-angle, see table 4.

Figure 2

Observer ratings of the Graf classification of 158 sonograms; distribution of types of the Graf classification by the four groups of observers in absolute numbers.

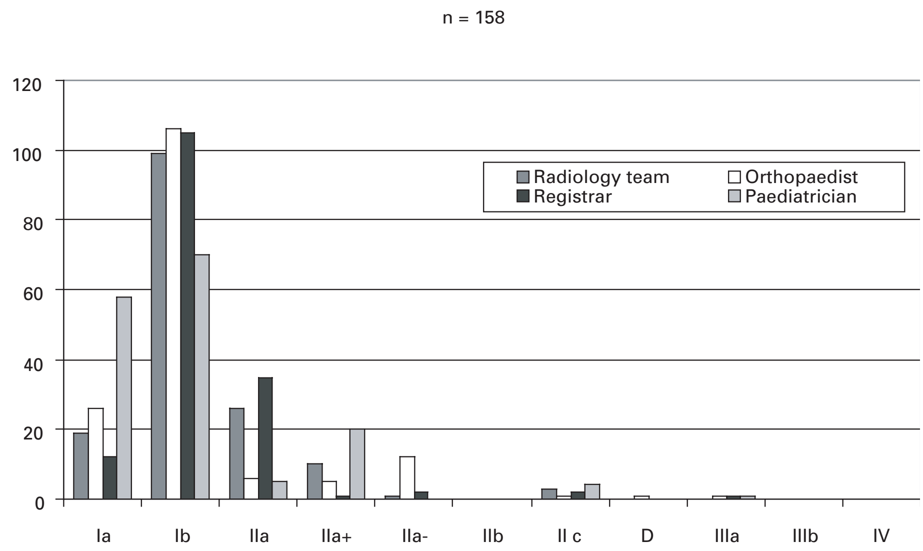


Table 1

Distribution of Graf classification of 158 ultrasonograms by four observers with regard to normal, immature, dysplasia and dislocation hip type.

hip type	radiology team (%)	orthopaedist (%)	registrar (%)	paediatrician (%)
I (normal)	74.7	83.5	74.0	81.0
IIa/ IIa^+/IIa^- (immature)	23.4	14.6	24.1	15.9
IIb/IIc/D (dysplasia)	1.9	1.3	1.3	2.5
IIIa/IIIb/IV (dislocation)	0.0	0.6	0.6	0.6

Table 2

Inter-observer agreement between the paediatrician and the registrar in orthopaedics with regard to normal and abnormal hips (absolute numbers).

	paediatrician		
Registrar	normal	abnormal	total
Normal	111	6	117
Abnormal	21	20	41
Total	132	26	158

Table 3

Inter-observer agreement: proportion of agreement *(italic)* positive and negative index of proportional agreement (normal); kappa value and 95% kappa confidence interval (bold) of the four groups of observers on the dichotomous classification normal vs. abnormal.

	Radiology team	orthopaedist	registrar	paediatrician
Radiology team		0.63 (0.61–0.65)	0.71 (0.68–0.73)	0.58 (0.56–0.61)
Orthopaedist	85% $p_{pos} = 0.91$ $p_{neg} = 0.67$		0.64 (0.62–0.66)	0.69 (0.68–0.71)
Registrar	87% $p_{pos} = 0.87$ $p_{neg} = 0.75$	86% $p_{pos} = 0.91$ $p_{neg} = 0.68$		0.55 (0.53–0.57)
Paediatrician	84% $p_{pos} = 0.89$ $p_{neg} = 0.63$	90% $p_{pos} = 0.94$ $p_{neg} = 0.7$	83% $p_{pos} = 0.89$ $p_{neg} = 0.60$	

Table 4

Mean values and standard deviations (SD) in degrees for the alpha and beta angles of 158 ultrasonograms.

	alpha		beta	
	Mean (°)	SD (°)	Mean (°)	SD (°)
Radiology team	62.2	5.4	62.3	6.3
Orthopaedist	63.4	5.2	61.5	7.0
Registrar	61.8	5.0	67.0	6.0
Paediatrician	64.1	5.5	57.4	6.0

Table 5

Bland/Altman angle analysis for the paediatrician and the registrar with mean difference, 95% CI and limits of agreement low and up.

angle	side	mean difference (95% CI)	low (95% CI)	up (95% CI)
alpha	disordered healthy	–2.36 (–3.3, –1.4) –2.29 (–3.3, –1.3)	–10.6 (12.2, –8.9) –11.1 (–12.8, –9.3)	5.8 (4.2, 7.5) 6.5 (4.7, 8.3)
beta	disordered healthy	8.6 (7.3, 9.9) 10.6 (9.2, 11.9)	–2.3 (–4.5, –0.14) –0.69 (–2.9, 1.6)	19.6 (17.4, 21.8) 21.8 (19.5, 42.1)

The estimated residual standard deviations from the hierarchical ANOVA are equal to 3.15° (95% CI 2.95, 3.37) for the α-angle and 6.1° (95% CI 5.7, 6.5) for the β-angle.

The intra-class correlation coefficient revealed a better agreement (ICC = 0.72) for the α-angle than (ICC = 0.34) for the β-angle.

Biases were found between different observers. The maximal bias was found between the paediatrician and the registrar. The registrar measured the α-angle to be 2° smaller and the β-angle to be 9° greater than those measured by the paediatrician at disordered hips (side 2) (table 5). This is visualised by the Bland Altman plots in figure 3. The minimal mean differences were found for the orthopaedist and the paediatrician. The orthopaedist measured the α-angle to be only 0.5° smaller and the β-angle to be 4° greater than those measured by the paediatrician.

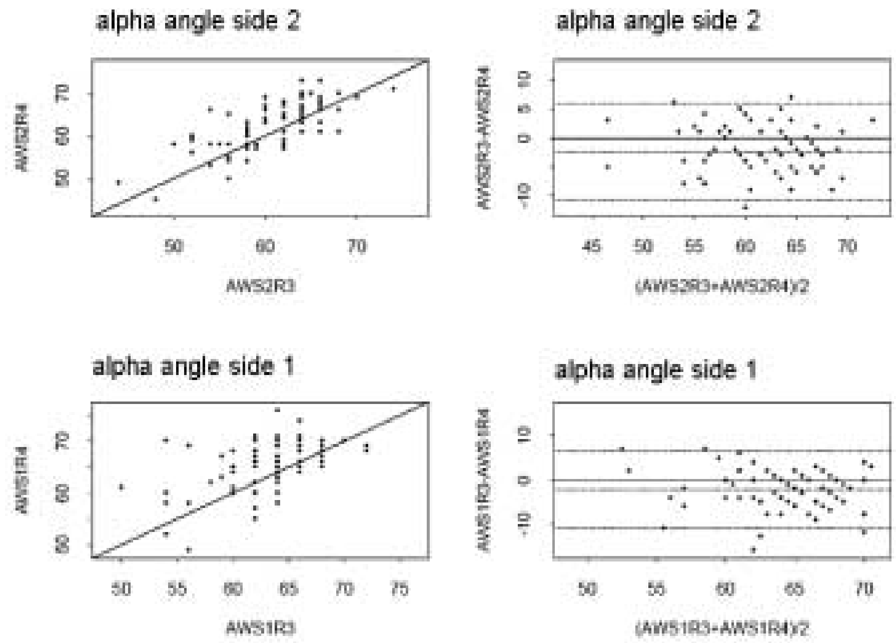


Figure 3

Bland/Altman plots for alpha angles of the paediatrician and the registrar in orthopaedics (side2 disordered), (side1 healthy).

Discussion

The cost of general sonographic screening amounts to CHF 11.5 m. per year for the investigation of all children born in Switzerland. Compared with the alternative, namely, clinical screening, general sonographic screening is about CHF 7 m. more expensive each year [12]. Since the result of screening is that treatment and other subsequent costs are reduced [7], cost-benefit analysis is required. To this end the agreement of an investigation must be known. In addition, there is much debate in the literature on ultrasound screening for developmental dysplasia of the hip because its value remains unclear. Part of the discussion is that Graf's technique is only reliable if the examiners are experienced.

The distribution of our patient sample corresponds to a distribution described as normal in the literature [1].

The high proportion of type II hips (Graf classification) in our study compared to studies in which the sonograms from the first weeks of life were evaluated arises from the fact that our sonograms were performed on out-patients who had been referred to us [10, 22, 30]. The gender distribution in our patient sample shows a significant predominance of girls which is in agreement with the data given in the literature [1, 2]. Similarly, published reports describe a significantly higher incidence of immaturity on the left side as opposed to the right [1, 2], and this tendency was also recorded in our study.

Agreement of the diagnoses

As it has been shown, the observer making the diagnosis plays an important role. The variability (difference in measurements) was significantly higher between well-trained and less well-trained personnel.

In addition, there is a tendency for experienced examiners to assess a hip as normal more frequently than inexperienced examiners. The inexperienced examiners classified the most hips as immature, which led to increased costs due to follow-up examinations and treatments. The greatest agreement at 90% existed between the paediatrician and the orthopaedist who both had besides the permit long-term experience in sonogram evaluations, which seemed to explain the good and comparable results. The poorest agreement was found for the paediatrician and the registrar. These results support our statement that the degree of experience and training has an important influence on the agreement of assessment.

The registrar for orthopaedics had no permit and was, as the team from the Department of Radiology, less experienced. The comparable results of those two groups lead to the hypothesis that observers with less experience tend to assess sonograms towards greater pathology.

It is of particular clinical relevance that no disagreement occurred amongst the observers for the dichotomy normal vs. dysplastic. Thus, no pathological condition was missed despite the sometimes wide scatter in the measurement of angles.

A higher agreement in the assessment of immature hips could be demonstrated for the Radiology team and the orthopaedist by estimating CCC together with the corresponding 95% CI. This tendency was also apparent for other observer groups. This led to the assumption that attention increased during the assessment procedure as soon as the first glance suggests an immature hip. A comparison across observer groups of first glance assessment with regard to the central position of the hip was not performed in our study. It would however support the requirement that sonograms should first be assessed and described according to standard nomenclature before they should be "measured up". A good agreement for inter-observer was found based on the dichotomous characteristics "normal" vs. "abnormal" based on the kappa coefficients. Roovers et al. [5] reported comparable results. Since the kappa coefficient depends on the true prevalence, studies can only be correctly compared if there is agreement of the group divisions [22]. For this reason, we selected group divisions in agreement with those of Roovers et al. [5]. Rosendahl et al. [15], Baron et al. and Ömeroglu et al. [13, 14] described better results for the kappa values but because the division of groups is not in agreement with ours, no direct comparison can be made. This fact has not always been adequately taken into account in other studies [5]. A poor agreement is clear in the study by Disat et al. [25]. The computation of the hierarchical ANOVA, the intra-class correlation coefficient (ICC), the application of the method by Bland-Altman to obtain the estimates of bias and the limits of agreement together with the estimation of the concordance correlation coefficient (CCC) by Lin seem to be appropriate statistical tools to study the agreement between observer groups. Moreover, the inter-observer agreement found in our study is comparable with the results of other studies [5, 13–15].

Agreement of angles

In the literature an accuracy of ± 11°, on average 3.2°, is given for the α-angle and of ± 28°, on average 11.9°, for the β-angle for hip sonographies according to Graf [10].

Niethard and Roessler assume an error margin in the determination of the α- and β-angles of ± 10° [11]. The current study by Roovers et al. reports average standard deviations of 3.2° and 6° for the α- and β-angles, respectively [5]. Our study produced similar results and can be regarded as comparably good. Graf assumes deviations for both angle values of ± 4° [9]. Radiology could also be used to assess hip centralization and roofing. The maximum error, as established by experimental investigation, for the radiographic measurement of the AC angle is 0°–8°, on average 2.3°. In terms of accuracy, sonography is thus comparable with radiographic imaging. Apart from x-ray exposure, radiological examination can only take place when the infant is older and the osseous structures can be adequately visualized [10]. The radiograph, therefore, gains in importance after the first year of life when sonography of the hip is rendered unintelligible by the echodense femoral nucleus.

Although the variability of the α- and β-angles when measured on the same sonogram was high between observers, it did not lead to any false negative assessments since there was a tendency to classify hips as more severely affected than they actually were in cases of uncertainty. This may however lead to over-treatment and an increased number of follow-ups. The dependence of observer measurements per child was represented by the intra-class correlation coefficient. ICC for the α-angle (ICC = 0.72) was clearly better than for the β-angle (ICC = 0.34); this is known from the literature [5].

Based on the agreement limits it could be significantly demonstrated that the maximal differences existed for the observer pair "paediatrician and registrar" and the minimal differences between "orthopaedist and paediatrician".

The finding that the registrar measured the α-angle about 2° smaller and the β-angle 9° greater than measured by the paediatrician supports the theory that the qualification of the examiner plays an important role as does the fact that the most minimal deviation existed between the paediatrician and the orthopaedist. The orthopaedist measured the α-angle at only about 0.5° smaller and the β-angle 4° greater than the paediatrician did.

To substantiate these findings further studies are required. Apart from demonstrating reproducibility, these studies would show to what extent the scatter depends on the level of training.

Conclusions

The inter-observer agreement of our study corresponds to results found in the literature.

Evaluation of hip sonograms is comparable to radiographic evaluation with regard to the accuracy of angle measurement.

Examiners with less experience assess sonograms with a tendency towards greater pathology so that pathological findings will not be missed, however, this may lead to an increased number of check ups and treatments.

Pathological findings are generally identified as such by all examiners.

The disagreement between trained and untrained personnel is statistically significant, thus regular quality controls and training of sonographers as well as standardized training and re-certification remain essential.

Correspondence:
Dr. med. Gregor Kohler, Leitender Arzt
Orthopädie und Traumatologie
Kantonsspital Frauenfeld
Postfach
CH-8501 Frauenfeld
E-Mail: gregor.kohler@stag.ch

References

1 Liebe A. Diagnostik der Säuglingshüfte. Stellenwert der Anamnese und des klinischen Befundes gegenüber der Hüftsonographie nach Graf. Orthopädie aktuell 2202;112.
2 Schambes M. Ergebnisse des Hüftultraschallscreenings von 1993–1998, Klinik Hirslanden, Zürich, Interdisziplinäre Fortbildung November 2002; 1–2.
3 Henner HD, Otto R CH, Ramzin MS, Zweifel HJ. (eds.) Ultraschalldiagnostik, Springer Verlag Berlin, Heidelberg; 198:20–368.
4 Graf R. Sonographie der Säuglingshüfte und therapeutische Konsequenzen. Ein Kompendium. 5. Ed., Thieme; 137–149.
5 Roovers EA, Boere-Boonekamp M, Geertsma T, Zielhuis G, Kerkhof A. Ultrasonic screening for developmental dysplasia of the hip in infants. Reproducibility of assessments made by radiographers. J Bone Joint Surg. 85–B; 5:726–30.
6 Schilt M. Die konservative Therapie der angeborenen Hüftdysplasie und Luxation; Monograph, Symposium der Universität Irchel, Schweizer Gesellschaft für Ultraschall in Medizin und Biologie; 91–102.
7 www.sssss.ch.
8 Schilt M. Optimaler Zeitpunkt des Hüftsonographie-Sceenings, Ultraschall in Med. 2001; Georg Thieme Verlag; 22: 39–47.
9 Graf R. Die anatomischen Strukturen der Säuglingshüfte und ihre sonographische Darstellung. Morphol Med1982;2:29–38.
10 Melzer Ch. Röntgenbild-Sonographie-Anatomie, Orthopädische Klinik Giessen, Angeborene Hüftdysplasie und -luxation vom Neugeborenen bis zum Erwachsenen. Symposium Zürich Universität-Irchel, Nov. 1993, 69 ff.
11 Niethard FU, Roesler H. Die Genauigkeit von Längen- und Winkelmessungen im Röntgenbild und Sonogramm des kindlichen Hüftgelenkes. Z Orthop 1987;125:170–6.
12 Albano T. Umsetzung und Kosten des sonographischen Screenings in der Schweiz. Clinical Research Hirslanden, Ärztefortbildung Klinik Hirslanden Zürich, Monograph, 28. Nov. 2002; 1–9.
13 Bar on E, Meyer S, Harari G, Porat S. Ultrasonography of the hip in developmental hip dysplasia. J Bone Joint Surg (BR) 1990;80–B:321–4.
14 Ömeroglu H, Bicimoglu 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.
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 SBU/SE, Evaluation med. Interventionen, Hüftultraschall bei Neugeborenen, Prävention von Hüftleiden, Anterior-dynamic ultrasound in the early detection in newborns (ALERT), HTA-Newsletter, Nov. 2002; 12: 2 ff.
17 Graf R. Sonographie der Säuglingshüfte und therapeutische Konsequenzen. Ein Kompendium. 5. Ed., Thieme; 25 ff.
18 Ahrens H. Varianzanalyse, Wissenschaftliche Taschenbücher WTB, Akademie-Verlag, Berlin; 1998
19 Willet W. Nutrial Epidemiologie, New York, Oxford University Press, 1990; 96–114.
20 Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
21 Landis JR, Koch GC. A review of statistical methods in the analysis of data arising from observer reliability studies (part 1), Statist Neerland 1975;29:151–61.
22 Holle R, Victor N. Methoden zur Konstruktion und Evaluierung klinischer Scores. Ruprecht-Karls-Universität Heidelberg, Institut für Medizinische Biometrie und Informatik. Abteilung Medizinische Biometrie; Nr. 27, 47 ff.
23 Precht M, Kraft R. Bio-Statistik 2: 252 ff.
24 Altman DG, Machin D, Byrant T, Gardner M. Statistics with confidence, second Edition; 20–25.
25 Bland LM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
26 Altman DG. Practical Statistics for medical Research, Chapman & Hall, London; 403–406.
27 Lin LI. Assay validation using the concordance correlation coefficient. Biometrics 1992;48:599–604.
28 Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989;45:255–68.
29 Bland LM, Altman DG. Measurement error. BMJ 1996;313: 744.
30 Joller R. Sonographie der Säuglingshüfte – Erste Ergebnisse eines Screening – Programms im Kanton Uri, Symposium Zürich Universität – Irchel; Nov. 1993; 163 ff.
31 Cohen J. A Coefficient of Agreement for Nominal Scales. Educ Psychol Meas 1960;20:37–46.
32 Byrt T, Bishop J, Carlin JB. Bias, Prevalence and Kappa. J Clin Epidemiol 1993;46:423–9.
33 Cicchetti DV, Feinstein AR. High Agreement but Low Kappa: II. Resolving the Paradoxes. J Clin Epidemiol 1990;43:551–8.
34 Feinstein AR, Cicchetti DV. High Agreement but Low Kappa: I. The Problems of Two Paradoxes. J Clin Epidemiol 1990;43: 543–9.
35 Dias JJ, Thomas ICH, Lamont AC, Mody BS, Thompson JR. The reliability of ultrasonic assessment of neonatal hips. J Bone Joint Surg (Br) 1993;75–B;479–82.

Swiss Medical Weekly

Official journal of
the Swiss Society of Infectious disease
the Swiss Society of Internal Medicine
the Swiss Respiratory Society

The many reasons why you should choose SMW to publish your research

What Swiss Medical Weekly has to offer:

- SMW's impact factor has been steadily rising, to the current 1.537
- Open access to the publication via the Internet, therefore wide audience and impact
- Rapid listing in Medline
- LinkOut-button from PubMed with link to the full text website http://www.smw.ch (direct link from each SMW record in PubMed)

- No-nonsense submission – you submit a single copy of your manuscript by e-mail attachment
- Peer review based on a broad spectrum of international academic referees
- Assistance of our professional statistician for every article with statistical analyses

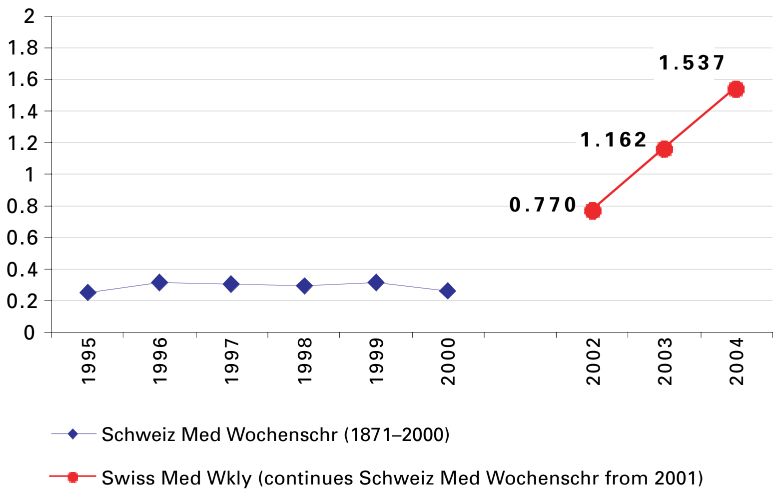
- Fast peer review, by e-mail exchange with the referees
- Prompt decisions based on weekly conferences of the Editorial Board
- Prompt notification on the status of your manuscript by e-mail
- Professional English copy editing
- No page charges and attractive colour offprints at no extra cost

We evaluate manuscripts of broad clinical interest from all specialities, including experimental medicine and clinical investigation.

We look forward to receiving your paper!

Guidelines for authors:
http://www.smw.ch/set_authors.html

Impact factor Swiss Medical Weekly



EMH FMH Schwabe
Editores Medicorum Helveticorum

All manuscripts should be sent in electronic form, to:

EMH Swiss Medical Publishers Ltd.
SMW Editorial Secretariat
Farnsburgerstrasse 8
CH-4132 Muttenz

Manuscripts: submission@smw.ch
Letters to the editor: letters@smw.ch
Editorial Board: red@smw.ch
Internet: http://www.smw.ch