



Središnja medicinska knjižnica

Šmigovec I., Đapić T., Trkulja V. (2014) *Ultrasound screening for decentered hips in children with severe cerebral palsy: a preliminary evaluation.* Pediatric Radiology, 44 (9). pp. 1101-9. ISSN 0301-0449

<http://www.springer.com/journal/247>

<http://link.springer.com/journal/247>

<http://dx.doi.org/10.1007/s00247-014-2956-6>

<http://medlib.mef.hr/2342>

University of Zagreb Medical School Repository

<http://medlib.mef.hr/>

Ultrasound screening for decentered hips in children with severe cerebral palsy: a preliminary evaluation

Igor Šmigovec¹, Tomislav Đapić¹, Vladimir Trkulja²

¹Department of Orthopedic Surgery, University Hospital Center Zagreb, Zagreb University School of Medicine, Zagreb, Croatia

²Zagreb University School of Medicine, Zagreb, Croatia

Corresponding author: Igor Šmigovec, MD
Department of Orthopedic Surgery
University Hospital Center Zagreb
Zagreb University School of Medicine
Šalata 6, 10000 Zagreb, Croatia
e-mail: ortopedija@yahoo.com
phone: +385959043165

Authors have no conflict of interest to declare. This study received no funding/support.

Abstract

Background. Ultrasound (US) is routinely used for hip screening in children with developmental hip disorder, whereas standard hip surveillance in children with cerebral palsy (CP) is based on repeated X-ray assessments. **Objective.** To evaluate US as a diagnostic tool in screening for decentered hips in children with CP. **Patients and Methods.** Prospective, diagnostic single-center assessor-blind study. Consecutive CP patients (age 2-8 years) with severe motor disability (unable to walk unaided) underwent US (lateral longitudinal scan to determine lateral head distance, LHD) and X-ray hip assessment (migration percentage, MP). Diagnostic properties of LHD in detecting hips with $MP \geq 33\%$ (requires preventive treatment) were evaluated overall ($n=100$) and for hips assessed at the age 24-60 ($n=38$) or $60 < \text{to} \leq 96$ months ($n=68$). Fifty hips underwent US assessment by two investigators to evaluate inter-rater reliability and agreement. **Results.** Prevalence of $MP \geq 33\%$ was 22.0% overall and 26.2% and 19.4% in the two age-based subsets, respectively. LHD well discriminated hips with $MP \geq 33\%$ (areas under the ROC curves 94%, 99% and 92%, respectively). At the optimum cut-off values of LHD (5.0, 5.0 and 4.8 mm, respectively), sensitivity was 95.5%, 100% and 100% overall and in the two age-based subsets, respectively, whereas specificity was 85.9%, 96.4%, 72.0%, respectively. Consequently, positive predictive value was relatively low, but negative predictive value was 98.5% (95% CI 92.1-100) overall and 100% (97.5% one-sided CI 87.2-100) and 100% (97.5 one-sided CI 90.2-100) in the two age-based subsets, respectively. Inter-rater reliability was high (intraclass correlation coefficient= 0.98, 95%CI 0.97-0.99) and 95% limits of agreement were reasonably narrow (-1.203 mm to 0.995 mm). **Conclusion.** In children with CP, US can be reliably used in screening for decentered hips and can greatly reduce the need for repeated X-ray assessments.

Key words: cerebral palsy, hip migration, ultrasound, X-ray, diagnosis

Introduction

At birth, children with cerebral palsy (CP) have normal hips but eccentric forces from spastic adductors and flexors tend to change the structure of the developing joint. Muscle imbalance may result in lateralization (decentered hips) and, eventually, in more advanced abnormalities – hip dysplasia and hip dislocation [1]. Hip migration in CP patients is rather common, particularly among children who cannot walk unaided. The changes become noticeable around the age of two and the risk of hip dislocation is highest before seven years of age [1]. Hip migration and/or displacement can cause severe pain, loss of seating balance and comfort, progression of scoliosis (if present), whereas progressive hip contracture makes hygiene difficult [2, 3]. However, advanced hip lateralization is preventable and successful screening programs have been developed based on repeated clinical and radiological (X-ray) evaluations, typically in 6-8-month intervals and, where needed, introduction of early preventive treatments [4-6]. The classical X-ray diagnosis of decentered hips is based on determination of the migration percentage (MP) which defines the part (%) of the femoral head not covered by the acetabulum [7] (Fig. 1), whereas diagnosis of acetabular dysplasia is usually based on determination of the acetabular index (AI) [8] (Fig. 1). The aim is to detect hips with MP indicative of a need for preventive treatment: typically, when $MP \geq 33\%$, intensive conservative or preventive surgical treatment is indicated, although some authors suggest $MP \geq 25$ as a threshold [9, 10]. The limitation of this screening scheme lies with the fact that infants are repeatedly exposed to X-rays. Children suffering from developmental dysplasia of the hip (DDH) also require repeated radiographic assessments in order to prevent hip dislocation. The use of ultrasound in DDH has been widely accepted during the past 20 years. Determination of the lateral head distance (LHD) which defines, in millimeters, the part of the femoral head not covered by the bony acetabulum (Fig. 1) shows good concordance with MP values (children ≥ 2 years of age) and ultrasound has replaced X-ray diagnostics in this setting to a great extent [11]. However, the experience with ultrasound in DDH cannot be simply transferred to children with CP – the biomechanics of hip changes in the two conditions is considerably different. In DDH, acetabular dysplasia is a primary disorder and femoral head migration

occurs as a consequence, typically in the anterior and lateral direction [10]. In CP, muscle imbalance forces femoral head to migrate typically in the posterior and lateral direction and acetabular dysplasia is a secondary development. We therefore aimed to assess diagnostic value of the ultrasound measure LHD in screening for decentered hips in children with severe CP at the age between 2 (the earliest time to expect lateralization) and 8 years (risk of relevant lateralization disappears).

Patients and methods

General design and ethics

This was a prospective, single-center, investigator-blind diagnostic study conducted between March 2010 and April 2012. The study was approved by the local University School of Medicine Ethics Committee.

Standard hip surveillance procedure, eligibility for the present analysis and patient flow

The standard hip surveillance procedure in children with CP (age 2-8 years) is based on clinical assessment (hip abduction with the hips and knees extended) and on hip X-rays [4]. Hip abduction is assessed every 6 months: if $>45^\circ$, regular clinical checks are continued without X-rays; if $\leq 45^\circ$, a child is considered at a high risk of a relevant hip lateralization (MP $\geq 33\%$) and a hip X-ray is indicated (always both hips). If MP $< 25\%$, regular clinical checks are continued and X-ray is repeated annually, whereas if MP is 25% to 32%, X-ray is repeated every 6 months. In non-ambulatory children with CP, i.e., those with a severe bilateral spastic CP level IV or V according to the gross motor functions classification system (GMFCS) [12], hip abduction is most commonly $\leq 45^\circ$ and X-rays are repeated in 6-8 month intervals or annually, depending on the actual MP. Monitoring is continued until a child is 8 years old. Should MP $\geq 33\%$ be detected, monitoring after corrective interventions follows the same general scheme [4].

Eligible for the present analysis were consecutive children with CP who met the following criteria: a) informed consent (parents/guardians); b) age 24 to 96 months (inclusive); c) GMFCS level IV or V.

Exclusion criteria were: a) previous surgical interventions to the hip; b) other conditions affecting the

hip, e.g., transient synovitis, Legg-Calve-Perthes disease or slipped capital femoral epiphysis.

Generally, patients (hips) were assessed regardless of whether examined for the first time or at follow-up visits. In line with the indication, both hips underwent X-ray and ultrasound (US) evaluation. However, once MP $\geq 33\%$ had been detected (unilaterally or bilaterally), further radiological assessments for this particular child were not considered for the present analysis. Hence, some patients (hips) provided X-ray/US data on more than one occasion (Fig. 2), but at each assessment the investigators were fairly unprejudiced about the local (anatomical) hip condition.

Primary objective and methods to ascertain investigator blind

The primary objective was to assess whether the ultrasound measure *lateral head distance* (LHD) [11, 13] (Fig. 1) could be meaningfully used in hip screening in children with severe CP, i.e., whether it could detect the level of hip lateralization defined as MP $\geq 33\%$. Clinical evaluation was performed by the same experienced investigator (IŠ, 7 years of experience as a pediatric orthopedic surgeon) and, when indicated, anteroposterior hip X-ray followed. Radiographs were evaluated by another experienced investigator (TĐ, 17 years of experience as a pediatric orthopedic surgeon, >5000 hip X-ray assessments in children) unaware of the clinical/US status. Ultrasound assessment was subsequently performed by the investigator unaware of the X-ray finding (IŠ, >3000 US hip assessments in children). Fifty hips underwent US examination by both investigators (IŠ and TĐ; TĐ with >6000 US hip assessments), independently, on the same day. No formal randomization was employed for this purpose: patients (hips) were evaluated consecutively, with no specific selection criteria or assessment order, until 50 hips were analyzed. In this case, US was performed before the X-ray evaluation.

X-ray assessment (the “standard of truth”)

All patients (hips) underwent the same standard anteroposterior X-ray assessment (RADSpeed apparatus, Shimadzu Corp., Kyoto, Japan; tube voltage range 42-50 KV; tube current range 28-40 mA; entrance surface dose range 0.2-0.7 mGy; dose-area product range ~ 20 -350 mGy*cm²) [4]: children were positioned with the pelvis horizontal, flat and level and the legs were positioned almost parallel

to each other (neutral abduction/adduction) with the patellae facing upwards; in the case of a fixed flexion deformity giving a lordosis, the legs were raised until the lumbar spine flattened. Migration percentage (MP) and acetabular index (AI) (Fig. 1) were determined. In the case of the “Gothic arch” formation of the lateral acetabular margin, the midpoint of the arch was to be used as a reference point for calculation of MP (Fig. 1) [14]. Figure 3 shows a marked bilateral hip decentration in a 39 months old girl with MP=55%.

Ultrasound assessment (the “test procedure”)

All patients underwent the same standardized real-time ultrasound assessment (Logiq 200 apparatus, GE Medical Systems, Milwaukee, WI, USA) being positioned in the same manner as for the X-ray, and using a method validated in DDH [11, 13]. Briefly [11, 13], scanning was performed from the lateral aspect of the hip region with a 5 MHz linear transducer. With the transducer kept parallel with the long axis of the body, a lateral longitudinal scan was employed showing the lateral anatomic landmarks of the hip joint. When depicting the central part of the femoral head, the distance from the lateral tangent of the ossification center to the lateral bony acetabular rim – lateral head distance (LHD, millimeters) - was measured to represent the uncovered part of the femoral head and indicate lateralization of the femoral head relative to the acetabulum. When the entire bony femoral head was medial to the acetabular bony rim, LHD was given a negative sign.. The anatomical structures on ultrasound image are depicted 90° rotated to those on a radiograph (Fig. 1). Figure 3 shows LHD in a decentered hip (MP=55%) in a 39 months old girl.

Sample size and power considerations

We expected around 25% of the examined hips to have MP \geq 33% (“event prevalence”). Considering the experience with ultrasound in DDH [11, 13], we expected LHD to have a high discriminative power defined as an area under the receiver-operating characteristic (ROC) curve of at least 90%. With the assumed event prevalence of 25% and in order to achieve 90% power to detect a difference between an area under the ROC curve of 90% (highly discriminative test) and an area under the ROC

curve of 75% (modestly discriminative test) as significant at a two-sided alpha level 0.05, we calculated that at least 87 hips needed to be examined.

Data analysis

The present analysis refers to 100 X-ray/US hip images obtained from 36 children all of whom underwent bilateral radiological hip assessment at least once, whereas in 9 children the hips were assessed on repeated occasions (Fig. 2). Although the collected data were clustered by nature (two hips “within” a child; repeated assessments “within” a hip [child]) we assumed that in this specific case they could be considered “numerically independent”, i.e., that each pair of radiological data (X-ray, US *per* hip) represented a different anatomical condition and corresponding (different) radiological values (MP, LHD). The assumption was based on the following reasoning: a) two hips (assessed at one time-point) in one child are not anatomically identical, i.e., each has specific key radiological measures; b) at this age and condition, anatomical hip particulars should not be expected to be identical when looked at 6-8 months apart (i.e., hip lateralization may progress or reduce between visits). To evaluate this assumption, we first determined agreement between radiological parameters (MP, LHD) established for two hips (left and right) at one visit (all included children, radiological data obtained at the first assessment point), and between radiological parameters established on two separate assessments of the same hips (children i.e., hips assessed on the first and the second assessment point, Fig. 2). As depicted in Fig. 4a, radiological measures (MP by X-ray, LHD by US) taken on two hips in the same child on one occasion showed a considerable (numerical) disagreement – the 95% limits of agreement extended between -33.4% and 31.9% for MP and between -5.1 to 5.2 millimeters for LHD. Similarly, although somewhat less, the disagreement was considerable regarding repeated assessments on the same hips (Fig. 4b). In the next step, prevalence of MP $\geq 33\%$ was analyzed by fitting several generalized linear mixed models with logit link and binomial distribution (logistic models) with LHD as a fixed effect and different modeling of random effects. The first model included “patients” as a G-side random effect to account for spatial clustering (“hips” within “patients”) and “assessment point” as an R-side (residual) random effect to account for

temporal clustering (repeated assessments on the same hips). However, the distribution of random effects was far from normal, residual variability was not properly modeled (Pearson $\chi^2/df= 0.07$) and the amount of variability accounted for by the R-side random effect was negligible (covariance parameter estimate 0.029, standard error 0.163). The subsequent models included only G-side random effects with maximum likelihood estimation (Gauss-Hermite quadrature): a) a model with one G-side random effect to account for clustering of “hips” within “patients” (36 patients with a maximum of 8 observations *per* patient); b) a model with one G-side random effect that considered “hips” (nested in “patients”) as clusters of radiological observations (1-4 *per* hip); c) a model with two G-side random effects – (i) observations clustered in “hips(patients)” and (ii) “patients” (as a third level). The last model showed the best fit: random effects were closest to the normal distribution, -2Log Likelihood was the lowest (51.77), residual variance was modeled better than in other models, however still poorly (Pearson $\chi^2/df= 0.21$) and the test of covariance parameter estimates was not significant (df= 2; -Log Likelihood= 56.02, $\chi^2=4.26$, p=0.119). The estimated effect of LHD was OR=3.44, 95% CI 1.60-7.41. On the other hand, a fixed-effect model (LHD as a single independent) yielded only a slightly higher -2Log Likelihood (56.03), but with practically normal distribution of residuals and with no signs of overdispersion (Pearson $\chi^2/df= 1.02$) which, generally, could indicate a need to account for correlation within data. The effect of LHD was OR=3.00, 95% CI 1.79-5.02. Overall, we concluded that the most appropriate way to analyze the data is to treat them as independent observations. Discriminative properties of LHD in respect to MP $\geq 33\%$ were evaluated by logistic regression and the receiver operating characteristics (ROC) curve analysis. Diagnostic performance of LHD was determined at a cut-off value with an optimal combination of sensitivity and specificity (sensitivity, specificity, false-positive rate, false-negative rate, proportion of correctly classified hips, positive predictive value [PPV], negative predictive value [NPV], probability of an event with a negative test, likelihood ratio for a positive test [LR+] and a negative test [LR-] and diagnostic odds ratio). Since the probability of finding incident hips with MP $\geq 33\%$ declines with age and since children with once detected MP $\geq 33\%$ were not further included in the present analysis, we assumed

that the event prevalence would be lower for hips assessed at older age. Since event prevalence can affect tests of diagnostic performance (i.e., PPV, NPV), the analysis was performed for the entire data set and also separately for hips assessed at the age 24-60 months and hips assessed at the age 60< and ≤96 months. Hips that underwent ultrasound assessment by two independent investigators were used to assess inter-rater reliability (intraclass correlation coefficient, ICC) and limits of agreement for LHD. We used SAS for Windows version 9.3 (SAS Inc., Cary, NC, USA).

Results

The present analysis includes a total of 100 radiological (X-ray and ultrasound, US) hip assessments, 38 of which refer to hips examined at the age of 24-60 months and 62 refer to hips assessed at the age of 60< and ≤96 months. Migration percentage (MP) ≥33% was detected in 22 (22%) cases overall and in 10 (26.3%) and 12 (19.4%) cases in the two age-based subsets, respectively. Other radiological particulars are summarized in Table 1.

Does LHD discriminate between hips with MP≥33% and MP<33%?

Based on the areas under the ROC curves (all estimates >90%, all 95% CIs entirely >85%), lateral head distance (LHD) discriminated well between the hips with MP ≥33% and those with MP <33% - considering all hips, as well as considering hips assessed at the age 24-60 months and those assessed at the later age (Fig. 5). Cut-off LHD values with optimum levels of sensitivity and specificity were 5 mm, 5 mm and 4.8 mm considering all hips and those in the two age-based subsets, respectively (Table 2). In all three (sub)sets, LHD showed very high sensitivity, i.e., 95%-100% (Table 2).

Considering hips assessed at the age 24-60 months, specificity was also high (96.4%), consequently false positive and false negative rates were low. In this subset, LHD showed properties of an excellent diagnostic test both considering its “exclusion” and “confirmation” properties – NPV was 100% leaving no probability of MP ≥33% with a negative test (LHD <5 mm), and PPV exceeded 90% indicating <10% probability of “no event” with a positive test (LHD ≥5 mm) (Table 2). However, considering all hips (cut-off 5 mm) and those assessed at the older age (cut-off 4.8 mm), specificity

was relatively modest (86% and 72% , respectively) leaving a high false positive rate resulting in a poor PPV (Table 2). Still, negative predictive value in both sets was almost absolute (Table 2). In all three sets, general diagnostic test characteristics, i.e., likelihood ratios for a positive or a negative test and the overall diagnostic odds ratios were excellent (Table 2).

Concordance of lateral head distance measurement by ultrasound between two raters

Measurements of LHD by two raters in 50 hips were very similar regarding the means and inter-subject variability (very close between-subject standard deviations) (Table 3). Intraclass correlation coefficient (0.980) indicated a high inter-rater reliability and within-subject variability (variance coefficient) was minimal (Table 2). The limits of agreement indicated that in 95% of the cases the difference between the two raters was approximately between -1.0 and 1.0 millimeter (Fig. 6).

Discussion

In children with cerebral palsy (CP), migration of the femoral head is rather common but severe changes of the hip joint can be precluded by timely introduction of preventive treatments. Hence, patients undergo regular clinical and X-ray assessments in order to detect whether the threshold values of migration percentage (MP) are reached [6]. In children with developmental hip dysplasia (DDH), ultrasound has greatly replaced X-ray assessments in the screening programs [11].

Considering the substantial biomechanical differences between hip lateralization in CP and DDH, we aimed to evaluate ultrasound as a diagnostic aid in screening for decentered hips specifically in children with CP. In the present preliminary analysis we embraced CP patients who are typically closely monitored for hip lateralization, i.e., those with the most severe motor disability (GMFCS levels IV and V; children with lower levels are not routinely submitted to regular X-ray assessments) at the age between 24 and 96 months since incident hip migration is rare in children above or below this age. The limitations of the study are relatively modest single-center sample and inter-rater assessment of the method based only on two assessors. However, we have ascertained a fair level of internal validity by employing an appropriate standard of truth and independence of ultrasound and

X-ray evaluations. The present results strongly suggest that using the approach implemented in hip screening in DDH (lateral longitudinal scan) [11], ultrasound measure LHD can be successfully used in screening for decentered hips in children with CP, as well. In respect to discrimination between hips with MP $\geq 33\%$ and those with MP $< 33\%$, a commonly used threshold for introduction of preventive treatments [9], LHD shows properties of an excellent screening procedure – it is quick, simple and straightforwardly excludes the need for an X-ray assessment as a definite diagnostic tool. It should be noted that NPV was consistently high at different event prevalence - 26% (age 24-60 months), 19% (age 60< and ≤ 96 months) and 22% (overall). However, extension of the current study is underway that should enlarge the sample size across tighter age-based subsets possibly resulting in different prevalence of MP $\geq 33\%$ and should allow for further evaluation of predictive values of LHD. Also, it should enable validation of LHD (and possibly other ultrasound parameters) against other diagnostic MP thresholds (e.g., 25%, 60%) and acetabular index (AI). Only four hips in the current sample had AI $> 30^\circ$ (acetabular dysplasia), which was too few for a meaningful evaluation of discriminative properties of LHD (although all four hips had LHD well above 5 mm). Finally, inter-rater evaluation of LHD measurement indicated that it could be successfully implemented by different, comparably experienced ultrasound assessors.

Conclusion

The ultrasound measure lateral head distance could be used as a screening test for decentered hips in children with CP that could considerably reduce the need for repeated X-ray assessments.

References

1. Rang M, Silver R, de la Garza J (1986) Cerebral palsy. In: Lovell WW, Winter RB, eds. *Pediatric Orthopedics*, 2nd ed. Philadelphia, PA: JB Lippincott Co.; p. 345-396
2. Hodgkinson I, Jindrich ML, Duhaut P et al (2001) Hip pain in 234 non-ambulatory adolescents and young adults with cerebral palsy: a cross-sectional multicentre study. *Dev Med Child Neurol* 43:806-808
3. Knapp DR Jr (2002) Untreated hip dislocation in cerebral palsy. *J Pediatr Orthop* 22:668-671
4. Dobson F, Boyd RN, Parrott J et al (2002) Hip surveillance in children with cerebral palsy. *J Bone Joint Surg Br* 84-B:720-726
5. Hägglund G, Andersson S, Düppe H et al (2005) Prevention of dislocation of the hip in children with cerebral palsy. The first ten years of a population-based prevention programme. *J Bone Joint Surg Br* 87-B:95-101
6. Elkamil AI, Andersen GL, Hägglund G et al (2011) Prevalence of hip dislocation among children with cerebral palsy in regions with and without a surveillance programme: a cross sectional study in Sweden and Norway. *BMC Musculoskelet Disord* 12:284. doi: 10.1186/1471-2474-12-284
7. Reimers J (1980) The stability of the hip in children: a radiological study of the results of muscle surgery in cerebral palsy. *Acta Orthop Scand Suppl* 184:1-100
8. Hilgenreiner H (1925) Zur frühdiagnose und frühbehandlung der angeborenen hüftgelenksverrenkung. *Med Klin* 21:1385-1388, 1425-1429
9. Hägglund G, Lauge-Pedersen H, Persson M (2007) Radiographic threshold values for hip screening in cerebral palsy. *J Child Orthop* 1: 43–47
10. Miller F (2005) *Cerebral palsy*. New York: Springer Inc.
11. Tegnander A, Terjesen T (1999) Reliability of ultrasonography in the follow-up of hip dysplasia in children above 2 years of age. *Acta Radiologica* 40:619-624

12. Palisano R, Rosenbaum P, Walter S et al (1997) Development and reliability of a system to classify grossmotor function in children with cerebral palsy. *Dev Med Child Neurol* 39:214–223
13. Terjesen T, Rundén TO, Johnsen HM (1991) Ultrasound in the diagnosis of congenital dysplasia and dislocation of the hip joints in children older than two years. *Clin Orthop Relat Res* 262:159-169
14. Parrott J, Boyd RN, Dobson F et al (2002) Hip displacement in spastic cerebral palsy: repeatability of radiologic measurement. *J Pediatr Orthop* 22:660-667.

Fig. 1 Schematic representation of the X-ray and ultrasound assessment of hips in children with cerebral palsy. An anteroposterior X-ray of a hip (left) is used to detect migration percentage (MP= $a/b \times 100$) that defines the proportion of the femoral head not covered by the acetabulum. Distance a is a distance between the acetabular rim and the lateral tangent of the femoral head, and distance b is a distance between the medial and the lateral tangent of the femoral head (both in millimeters). Acetabular index (AI) is an angle and is expressed in degrees. In the case of the “Gothic arch” formation of the lateral acetabular margin (middle), the midpoint of the arch is used as a reference point for calculation of MP. The diagram on the right depicts elements of the hip and surrounding structures as visualized by a longitudinal lateral ultrasound scan. Lateral head distance (LHD) is a distance (in millimeters) between the lateral tangent of the bony femoral head and iliac baseline and defines the part of the femoral head not covered by the acetabulum. LHD determined by the ultrasound corresponds to distance a determined on the anteroposterior radiograph.

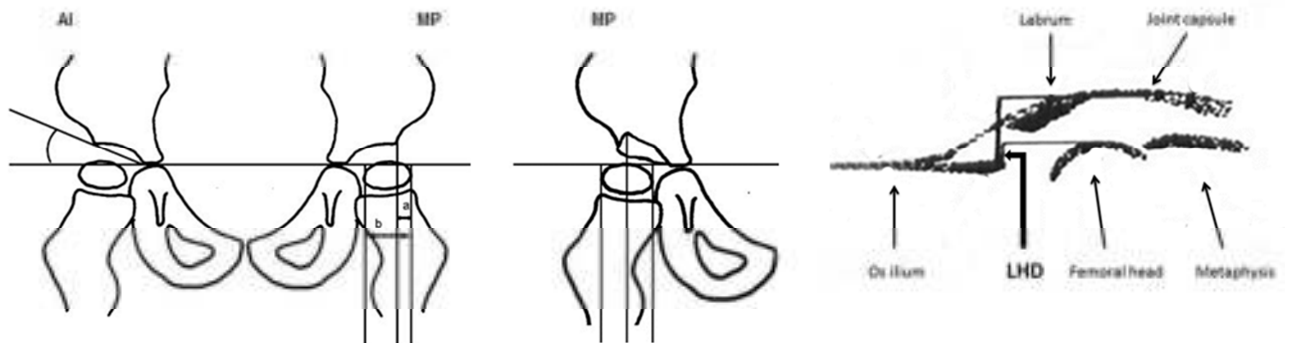


Fig. 2 Patients and hips in the present analysis. All X-ray and ultrasound (US) hip assessments were always performed bilaterally. A total of 36 children were included, 27 of whom underwent only one assessment and 9 patients underwent more than one assessment (6-8-month intervals). Once migration percentage (MP) $\geq 33\%$ had been detected (unilaterally or bilaterally), further radiological hip assessments for this particular child were not considered for the present analysis. Overall, 100 X-ray images and corresponding US hip scans were obtained.

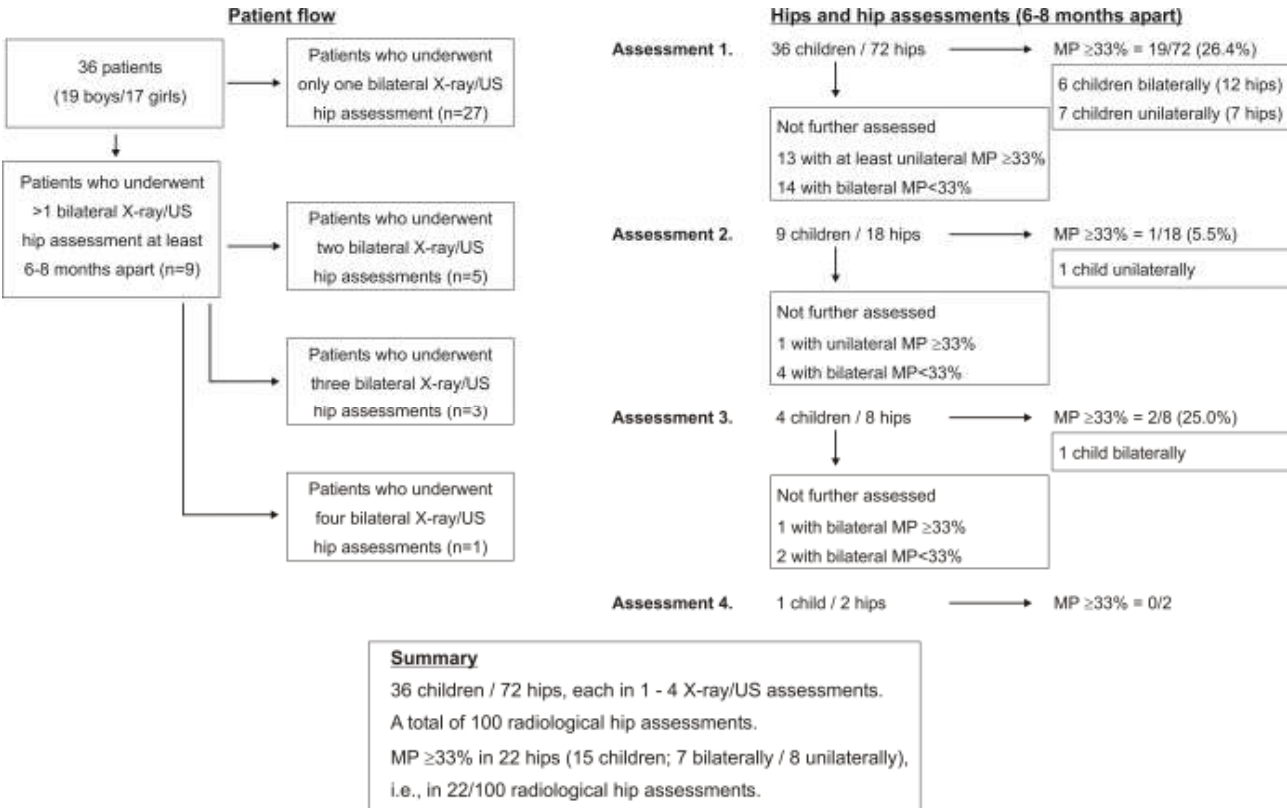


Fig. 3 Anteroposterior X-ray (left) of a marked bilateral hip decentration (migration percentage, MP, well above 33%) in a 39 months old girl with a severe bilateral spastic cerebral palsy (gross motor functions classification system level V) and a longitudinal lateral ultrasound scan of the right hip in the same patient (right) indicating lateral head distance (LHD) of 6.5 millimeters. In the clinical examination, passive abduction for both hips was possible to a maximum of 20 degrees. The girl underwent bilateral surgical preventive treatment.

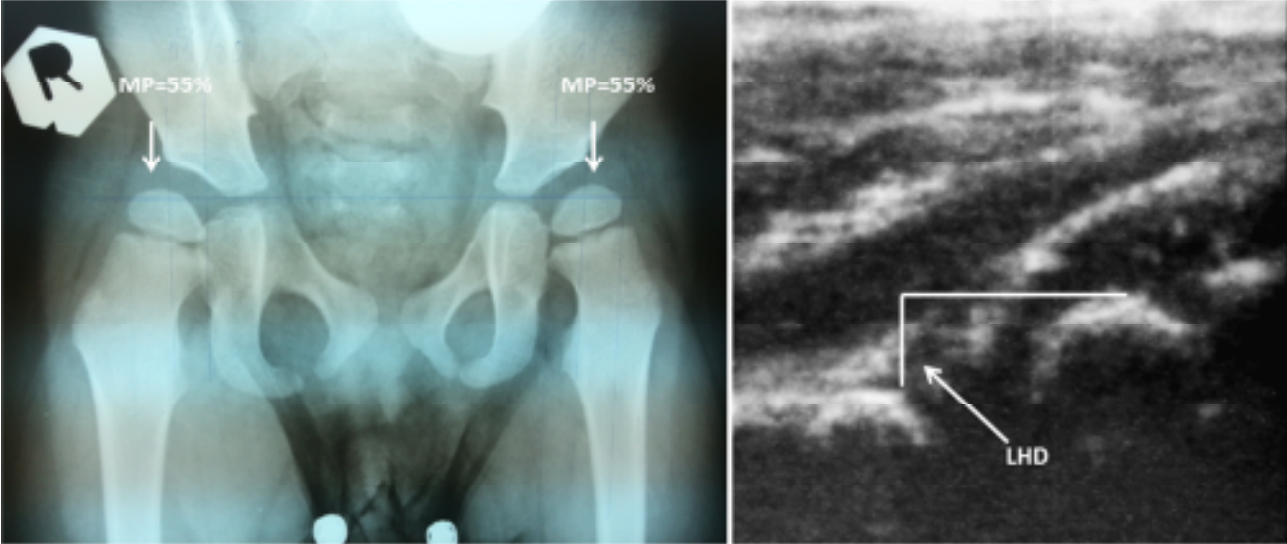


Fig. 4 Agreement in radiological measures (migration percentage, lateral head distance) between two hips in the same child **(a)** and between two assessments in the same hip **(b)**. Agreement plots with upper (UL) and lower (LL) 95% limit of agreement and mean difference (Δ).

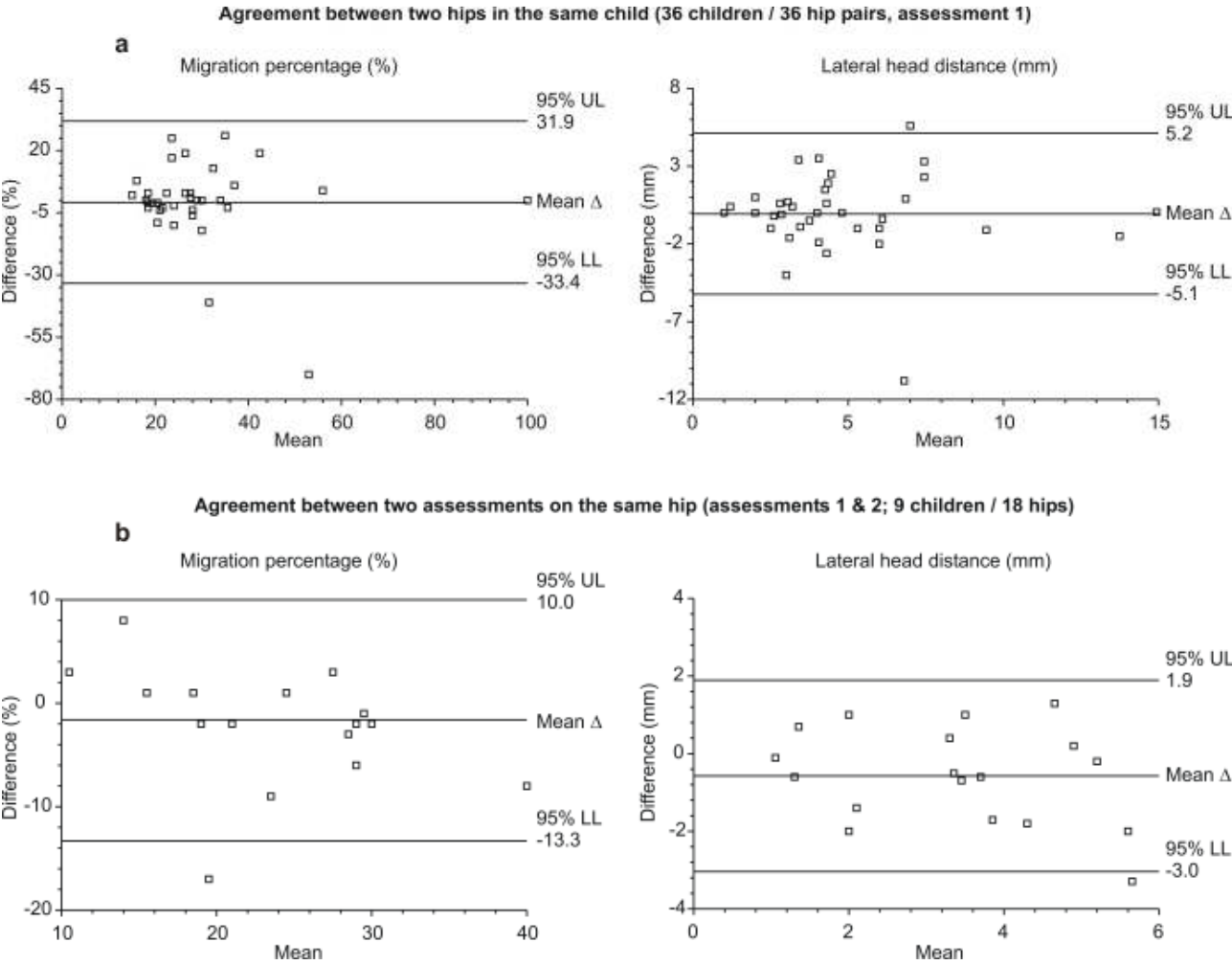


Fig. 5 Receiver-operating characteristics (ROC) curve (dashed lines) analysis for lateral head distance as a diagnostic test for migration percentage $\geq 33\%$ regarding all assessed hips and hips assessed at different age (months). Area under the curve (AUC) is given with 95 confidence intervals. Arrows depict points with an optimum relationship between sensitivity and specificity.

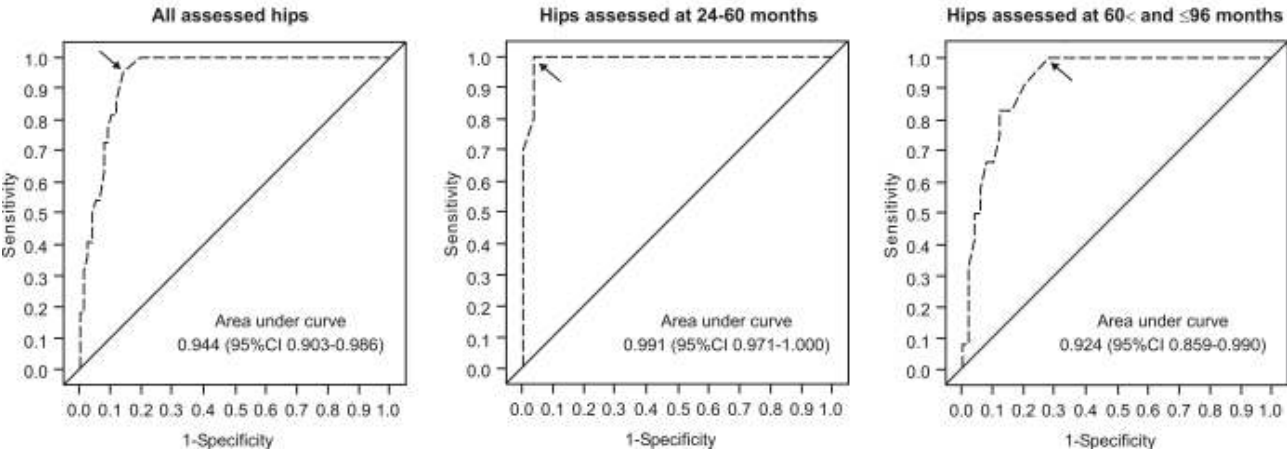


Fig. 6 Agreement plot for lateral head distance (in millimeters) measurement by ultrasound based on 50 hips assessed by two investigators with upper (UL) and lower (LL) 95% limits of agreement.

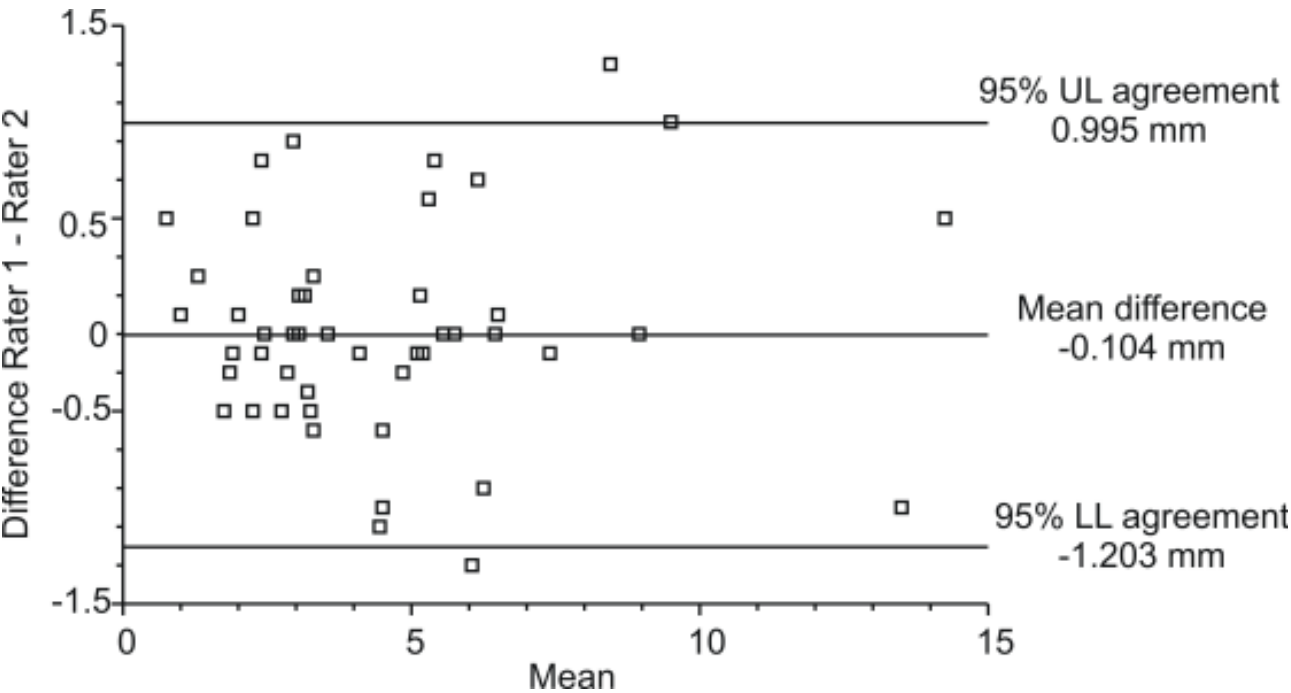


Table 1 Characteristics of the analyzed hips

	Overall	Age at assessment 24-60 months	Age at assessment 60< and ≤96 months
Number of assessed hips	100	38	62
Age at hip assessment (months)	69 (24-96)	47 (24-58)	78 (61-96)
<i>Main radiological findings</i>			
Migration percentage (MP) (%) (X-ray)	25 (9-100)	25.5 (10.0-100)	24.5 (9.0-58.0)
MP ≥33%	22 (22.0)	10 (26.3)	12 (19.4)
Lateral head distance (mm) (ultrasound)	3.9 (1.0-14.5)	3.1 (1.0-14.5)	4.0 (1.0-10.0)
<i>Other X-ray findings</i>			
MP=33%-99% (subluxation)	20 (20.0)	8 (21.1)	12 (19.4)
MP>60% (needs reconstruction)	3 (3.0)	3 (7.9)	0
MP=100% (luxation)	2 (2.0)	2 (5.3)	0
Acetabular index (AI) (°)	20 (10-35)	21 (10-35)	18 (12-34)
AI >30° (acetabular dysplasia)	4 (3.9)	3 (7.9)	1 (1.6)

Data are counts (%) or medians (range)

Table 2 Diagnostic properties of ultrasonographically measured lateral head distance (LHD) as a test for detecting hips with migration percentage $\geq 33\%$ (“event”) considering all hips and considering hips assessed at the age 24-60 months or at the age 60< and ≤ 96 months. Estimates are given with 95% confidence intervals.

	All assessed hips	Age 24-60 months	Age 60< and ≤ 96 months
Number of hips	100	38	62
Event prevalence [n (%)]	22 (22.0)	10 (26.3)	12 (19.4)
Optimal cut-off LHD value (mm) ^a	5.0	5.0	4.8
Sensitivity (true positives) (%)	95.5 (77.2-99.9)	100 (69.2-100) ^b	100 (73.5-100) ^b
Specificity (true negatives) (%)	85.9 (76.2-92.7)	96.4 (81.7-99.9)	72.0 (57.7-83.8)
1-specificity (false positives) (%)	14.1 (7.3-23.8)	3.6 (0.1-18.3)	28.0 (16.2-42.3)
1-sensitivity (false negatives) (%)	4.5 (0.1-22.8)	0 (0-30.8) ^b	0 (0-26.5) ^b
Correctly classified (%)	86.0	97.4	88.1
Positive predictive value (%)	65.6 (46.8-81.4)	90.9 (58.7-99.8)	46.2 (26.6-66.6)
Negative predictive value (%)	98.5 (92.1-100)	100 (87.2-100) ^b	100 (90.3-100) ^b
Event probability - negative test (%)	1.5 (0-7.9)	0 (0-12.8) ^b	0 (0-9.7) ^b
Likelihood ratio - positive test	6.8 (4.0-11.9)	28.0 (5.3-151)	3.6 (2.2-5.5)
Likelihood ratio - negative test	0.05 (0.01-0.26)	0 (0-0.29)	0 (0-0.34)
Diagnostic odds ratio	128 (16-5362)	19516- ∞	5895- ∞

^aCut-off value with an optimal level of sensitivity and specificity. LHD values \geq cut-off are considered a “positive test”. Values <cut-off are considered a “negative test”.

^bOne-sided 97.5 confidence interval

Table 3 Inter-rater reliability of lateral head distance (in millimeters) measurement by ultrasound based on 50 hips assessed by two investigators

	Rater 1	Rater 2
Mean	4.424	4.528
Standard deviation	2.827	2.785
Standard error	0.400	0.395
95% CI around mean	3.621-5.227	3.735-5.321
ICC= 0.980 (95% CI 0.969 to 0.991)		
WSV coefficient = 0.090 (95% CI 0.070 to 0.120)		

ICC, intraclass correlation coefficient (coefficient of reliability); WSV, within-subject variance