Relation between the quality of the ultrasound image acquisition and the precision of the measurement of the crown-rump length in the late first trimester: what are the consequences?

Ferdinand Dhombres[a,b,*], Nathalie Roux[a], Stéphanie Friszer[a], Roger Bessis[a,b], Babak Khoshnood[c], Jean-Marie Jouannic[a,b]

[a] Sorbonne Universités, UPMC Univ Paris 06, Fetal Medicine Department in Armand Trousseau University Hospital, APHP, Paris, France
[b] French College of Fetal Echography (CFEF), France
[c] INSERM UMR 1153, Obstetrical, Perinatal and Pediatric Epidemiology Research Team (EPOPé), Center for Epidemiology and Statistics, Sorbonne Paris Cité (CRESS), DHU Risks in Pregnancy, Paris Descartes University, Paris, France

ARTICLE INFO

ABSTRACT

Objective: To assess the extent to which the distribution of crown-rump length (CRL) values may be correlated with different criteria for the quality of the CRL images.

Study design: This is a retrospective analysis of a series of 977 CRL images, by two independent observers, for the presence or the absence of 14 quality hallmarks. Inter-observer agreement for the hallmarks was assessed by the proportion of agreement and Cohen's kappa. The association between the quantiles of the CRL distribution and the presence or absence of the 14 quality hallmarks was modeled using quantile regression.

Results: The overall inter-observer agreement across the 14 hallmarks was 91.7%, kappa = 0.81, 95% CI [0.80–0.82]. Distribution of CRL measurements varied considerably as a function of image quality: when the fetus was in extension, the mean CRL was +5.7 mm (vs. not in extension, $p < 0.001$), when the fetus was in flexion (vs. not), the mean CRL was −4.7 mm ($p < 0.001$) and when the image magnification was <65% (vs. >65%), the mean CRL was −4.2 mm ($p < 0.001$). There was a global trend to over-estimate the CRL for the higher deciles and to under-estimate the CRL for the lower deciles when the sagittal quality hallmarks were absent. No significant impact on CRL distribution was observed in association with the precise placement of the calipers nor with the horizontal orientation of the fetus.

Conclusion: Distribution of CRL measurements was influenced by the quality of CRL images. In particular, inadequate position of the fetus (flexion/extension) and insufficient image magnification were associated with systematic changes in the values of CRL. Our results show that as the quality of CRL images decreases, the associated variations in the distribution of CRL can have an impact on the chromosomal risk assessment and may lead to inappropriate obstetrical decisions.

© 2016 Elsevier Ireland Ltd. All rights reserved.

Introduction

The crown-rump length (CRL) of the fetus is the gold standard for dating of pregnancy in the first trimester [1–3], facilitating the appropriate obstetrical care for fetal growth anomalies and postterm pregnancy [4,5]. The CRL is also the ultrasound biometry used for the assessment of the risk for aneuploidy screening during the first trimester in various calculation algorithms, in combination with the nuchal translucency (NT) thickness measurement and maternal serum testing [6–8].

In direct line with the quality assessment and training programs for NT measurement [9–11], recent focus on the CRL quality lead to new image scoring systems, dedicated to the CRL [12–14]. Among these scoring systems, the French College of Fetal Echography (CFEF) established the first scoring system covering both the CRL and the NT quality [11], and demonstrated the interest of the dedicated CRL evaluation, independent from the NT quality criteria [15]. Various overlapping sets of criteria and quality hallmarks have emerged as parts of all these scoring systems [12–17]. However, none of them has been evaluated on clinical data. Moreover, the relation between the quality of the CRL images based on these hallmarks and the distribution of CRL values has not

been established, hence clinical implications remains theoretical or based on simulation data [18].

The aim of this study is to investigate the relation between the quality of CRL images and the distribution of the CRL measurements. More precisely, we aim to quantify the extent to which the distribution of CRL values might be correlated with different aspects of the quality of CRL images.

Materials and methods

This is a retrospective analysis of images collected by the CFEF national first-trimester scan quality assessment program. We extracted from the database a series of 1000 CRL images collected during consecutive practice assessments of operators (MDs and midwives) enrolled in the program. As previously described [15], each of the participating operators submitted a set of 30 images of CRL measurement from consecutive routine scans for evaluation. Each of the images was anonymized, the images being cropped during the submission process to remove all the personally identifiable information. As part of the program, trained experts from the CFEF evaluated the global quality of the CRL images based on the 8th item of the CFEF-ISM for the CRL, ranging from 1/4 ("unacceptable") to 4/4 ("good") [15].

Two independent observers evaluated the series of CRL images for the presence or the absence of 14 quality hallmarks derived from published scoring systems for the CRL [12–17]. The observers were blind to the global quality evaluation previously established during the quality assessment program. The description of the quality hallmarks used in this study is presented in Table 1. These hallmarks were grouped under five criteria: (i) mid-sagittal section, (ii) position of the fetus, (iii) image magnification, (iv) orientation of the fetus and (v) placement of the caliper at the crown and at the rump of the fetus. Ultrasound images corresponding to the 14 quality hallmarks of the five criteria are presented in Fig. 1. The CRL distributions were then compared across the five criteria, by subgroups defined by the presence or the absence of each hallmark. When the CRL differences were statistically significant, the gestational age distributions based on Robinson formula [2] were also determined.

The inter-observer agreement for each hallmark was assessed by the proportion of agreement and the Cohen's kappa (κ) [19]. To compare the CRL distributions associated with the quality hallmarks, the difference of CRL was determined by quantile regression [20], when each quality hallmark was either present or absent. The comparison of the CRL distributions was also performed across the CFEF global quality groups. All CRL deciles (1st–9th) were considered in the regression model and the significance level (alpha) was set at 0.05. Welch Two Sample t-tests were used to examine the differences between mean values of CRL, when each quality hallmark was either present or absent. The images without the value of CRL or without the calipers were excluded from the study. The statistical analysis was performed in R calculation environment version 3.2.2 (A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015) with the quantile regression package quantreg version 5.19.

The CFEF has received approval from the French National Committee of Information and Liberty (Commission Nationale de l'Informatique et des Liberte's, loi n 2004-801 du 6 août, declaration no. 1321560), for its data collection, analysis, and medical communication.

Results

A total of 977 CRL images were evaluated by the two observers for the 14 quality hallmarks. The overall inter-observer agreement across the hallmarks was 91.7% with a $\kappa = 0.81$ (95% CI: 0.80–0.82). The proportion of agreement and the κ for each hallmarks are presented in Table 2.

Among the 1000 images extracted from the CFEF training program, 2.3% were not amenable for further analysis because the value of CRL was not available. Among the remaining images, the mean value of the CRL was 64.1 ± 0.5 mm and the global quality of the images was deemed "good" or "acceptable" in 81.9% based on the CFEF ISM (Fig. 2). The CRL distributions were equivalent in the "good or acceptable" global quality group versus the "insufficient or inacceptable" quality group, with no significant difference in mean value of CRL (respectively 64.0 mm and 64.3 mm, p = 0.69). The distribution of these differences is presented in Fig. 3.

The distributions of CRL values were significantly different for both observers and for all the deciles, when the position of the fetus was not neutral (flexion or extension) or when the image

Table 1
List of the 5 criteria and 14 quality hallmarks for crown-rump length (CRL) images at the first trimester scan evaluated in this study.

Quality criteria	Quality hallmarks
Caliper placement	The crown caliper is correct:
	1. correct position for the full CRL measurement
	2. visualization of the skin line (caliper at outer border)
	The rump caliper is correct:
	3. correct position for the full CRL measurement
	4. visualization of the skin line (caliper at outer border)
Fetal position	The fetus is in neutral position:
	5. Absence of flexion of the fetus
	6. Absence of extension of the fetus
Mid-sagittal section	The section is mid-sagittal:
	7. Visualization of the fetal spine
	8. Visualization of the facial profile (nasal bones)
	9. Visualization of the 4th cerebral ventricle
	10. Visualization of the genital tubercle
	11. Non visualization of the orbit
Image magnification	Image magnification is optimal:
	12. The fetus occupies more than 2/3 of the image
	13. The fetus image is not cropped (it fits within the image)
Fetal orientation	The fetus is horizontally oriented:
	14. the line of the CRL measurement is 90° $\pm$ 15 to the ultrasound beam