A standardized approach for the assessment of the lower uterine segment at first trimester by transvaginal ultrasound: a flash study

Marina Kuleva, Olivier Castaing, Nicolas Fries, Jean-Pierre Bernard, Laurence Bussières, Marianne Fontanges, Daniel Moeglin & Laurent J. Salomon

To cite this article: Marina Kuleva, Olivier Castaing, Nicolas Fries, Jean-Pierre Bernard, Laurence Bussières, Marianne Fontanges, Daniel Moeglin & Laurent J. Salomon (2016) A standardized approach for the assessment of the lower uterine segment at first trimester by transvaginal ultrasound: a flash study, The Journal of Maternal-Fetal & Neonatal Medicine, 29:9, 1376-1381, DOI: 10.3109/14767058.2015.1051956

To link to this article: https://doi.org/10.3109/14767058.2015.1051956

Published online: 05 Jun 2015.
A standardized approach for the assessment of the lower uterine segment at first trimester by transvaginal ultrasound: a flash study

Marina Kuleva¹, Olivier Castaing², Nicolas Fries², Jean-Pierre Bernard¹, Laurence Bussières², Marianne Fontanges², Daniel Moeglin², and Laurent J. Salomon¹

¹Hôpital Necker-Enfants Malades, Maternité, Paris, France and ²Collège Français d’Echographie Fœtale (CFEF), AP-HP, Université Paris Descartes, Paris, France

Abstract

Objectives: To evaluate the reproducibility of a standardized approach to lower uterine segment (LUS) imaging by transvaginal ultrasound at 11–14 weeks.

Methods: This was a “flash” study lasting for 1 month. Obstetrician-sonographers performing more than 50 first trimester ultrasounds per year participated. All consecutive women attending for their 11–14 weeks scan were included. A standardized, transvaginal approach to the imaging of LUS was defined. The sonographers recorded one or two images of the LUS. The quality of the images was assessed by sonographers and reviewed by an independent fetal medicine specialist using the same scoring system. Inter and intra-reviewer variability was assessed.

Results: Seventy-one sonographers and 851 pregnant women participated. The mean (±SD) and medium (IQR) scores attributed by sonographer versus reviewer were 5.01 (±0.92) and 5 [4–6] versus 4.68 (±1.14) and 5 [4–5.24], p = 0.08. The mean [95% CI] difference of −0.33 [−2.6;2] was recorded. There was good, moderate and poor agreement in 74.4%, 16.7% and 8.9% cases, respectively. Variability in inter-reviewer and intra-reviewer was low with the mean [95% CI] difference of 0.1 [−1.6;1.4] and −0.1 [−1.4;1.2] respectively.

Conclusions: A standardized approach to LUS imaging at 11–14 weeks is feasible and highly reproducible in a large population.

Keywords

Cesarean scar, first trimester ultrasound, lower uterine segment

Introduction

Obstetric hemorrhage remains the leading cause of maternal morbidity and one of the top three causes of maternal mortality. In around half of all cases, obstetric hemorrhage is related to an anomaly in placental location or insertion [1]. Cesarean section (CS) has become the most common major surgical procedure in many parts of the world [2,3], and between 20% and 50% of babies are delivered by CS depending on the country. Consequently, more women are becoming pregnant with a scar on their uterus. This is associated with an increased risk of life threatening complications such as uterine rupture, low-lying placenta and placenta accreta [4,5].

The 11–14 weeks scan, which until recently was mostly used to assess the risk of Down’s syndrome, has lately been used to identify women at increased risk of adverse pregnancy outcomes such as placenta previa [6] as well as to evaluate the uterine cervix [7] or the morphology of CS scar [8,9].

Several studies have suggested that transvaginal ultrasound, as early as at the first trimester, is a useful technique to assess cesarean scarring and placental localization [8,10,11]. However, at present, there is no established standardized approach for the assessment of the lower uterine segment (LUS) in the first trimester. In this study, we introduced a standardized approach to LUS imaging by transvaginal ultrasound at 11–14 weeks. This approach involved a scoring system of image quality. We evaluated the reproducibility of this technique in an unselected population of women and sonographers.

Population and methods

This was a “flash” study carried out by College Français D’Echographie Fœtale (CFEF) starting in November 2012 and lasting for 1 month. The concept of a “flash study” was proposed by Salomon LJ at the CFEF 4th Scientific Meeting, 1–3 October 2010, Port-en-Bessin, France. A flash study is characterized by a short duration, a wide coverage, no modification to obstetric management and no additional costs. A standardized, transvaginal approach to the imaging of low uterine segment was defined (Figure 1) using the criteria listed below. These criteria were selected on the basis of
established standards and publications regarding low-uterine segment imaging [9]

1) The bladder is not visible (empty).
2) The internal cervical os (IO), (defined as the most internal point of the cervico-isthmic canal reaching the gestational sac) is clearly visible.
3) The external os is visible.
4) The cervico-isthmic canal is visible.
5) The internal os (IO) is situated in the median third of the image (this ensures the correct visualization of the lower segment of the uterus).
6) The posterior aspect of the posterior labia of the cervix is in the deepest half of the image (zoom criterion).

Figure 1. Standardized approach for the imaging of LUS. (1) The bladder is not visible (empty). (2) The internal cervical os (IO), (defined as the most internal point of the cervico-isthmic canal reaching the gestational sac) is clearly visible. (3) The external os is visible. (4) The cervico-isthmic canal is visible. (5) The internal os (IO) is situated in the median third of the image (this ensures the correct visualization of the lower segment of the uterus). (6) The posterior aspect of the posterior labia of the cervix is in the deepest half of the image (zoom criterion).

Cervical length was measured by a single straight line from the lower edge of the gestational sac to the external OS [12]. IO–trophoblast distance was measured by a single straight line from the IO to the lower trophoblast edge. Uterine scars were classified according to Stirnemann et al.: type 1A, thin and within the cervicoisthmic canal; type 1B, thin and above the cervicoisthmic canal; type 2A, dehiscent and within the cervicoisthmic canal; type 2B, dehiscent and above the cervicoisthmic canal [8] (Figure 2).

One or two high quality ultrasound images of the LUS using the standardized approach were required for each woman. The sonographers were asked to assess the quality of their images, by scoring one or zero points for each of the six criteria mentioned above. The difficulty in obtaining a correct image was also scored (1 – very easy, 2 – relatively easy, 3 – difficult, 4 – impossible). The necessity to record more than one image to obtain a high quality image was also noted.

One hundred and eighty randomly selected images were then reviewed by an independent fetal medicine specialist (herein referred to as a “reviewer”) using the same scoring system of quality. Sixty randomly selected images were then used to assess the variability in scores obtained by two reviewers (inter-reviewer) or in scores obtained by the same reviewer on two consecutive evaluations (intra-reviewer). For these images, instead of providing an overall score, the reviewers scored each of the six criteria separately, to analyze the reproducibility of these individual measures. These sample sizes were sufficiently large to detect systematic differences in scoring of 0.5 points between sonographers and reviewers, and 0.5 points between different reviewers. In both cases, alpha was set at 0.05 and beta was set at 0.2. The power calculation made using the PS software (Power and Sample...
Size Calculations) of the Vanderbilt University, Nashville, TN and based on the mean score and standard deviation values obtained by sonographers was >0.90. Operator and reviewer scoring were compared using a Mann–Whitney U-test. The percentage of scoring agreement between ultrasound technicians and reviewers was also calculated. To simplify the categorization of scoring agreement, a difference of one point or less was considered good agreement, a difference between one and two points was considered moderate agreement and a difference of more than two points was considered poor agreement. Inter- and intra-reviewer variability in scoring was assessed by percentage of scoring agreement. The Friedman test was used to the significance of variability in inter- or intra-reviewer scores. Agreement coefficients (AC1) for each criterion were computed as described by Gwet [13] to test the reproducibility of each criterion. AC1 was calculated as follows:

\[ AC_1 = \frac{P_a - P_e(\gamma)}{1 - P_e(\gamma)}, \]

where \( P_e(\gamma) \) is given by:

\[ P_e(\gamma) = 2P_+(1 - P_+), \]

where \( P_+ \) is equal to:

\[ P_+ = \frac{(A_+ + B_+)}{2} \]

Statstica 11 for Windows (StataCorp LP, College Station, TX) and Statistica (Stat Soft, Paris, France) were used for statistical analyses. For all tests, \( p < 0.05 \) was considered statistically significant. AC1 values of <0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80 and >0.81 were considered as poor, fair, moderate, good and excellent agreement, respectively [13].

**Results**

Seventy-one sonographers and 851 pregnant women participated in this study [the median number of women per sonographer was seven (IQR) = 5–15]. The median gestation age at the time of the scan was 12.5 weeks (IQR = 12.1–12.9).

Ultrasound measurements obtained by sonographers are shown in Table 1. The number of women with no history of CS was 754 (88.6%). Seventy-nine women (9.3%) had had one lower segment SC and 18 (2.1%) had had two. No uterine scar was found in 765 cases, including 100% of women who had never had an SC, and 11 (14%) of the women who had had one SC. In 82 cases, only one scar was detected. In four cases, two scars were detected, corresponding to women with a history of two lower segment SC. The trophoblast was visible in 755 cases (88.8%), being anterior in 355 (47%) and posterior in 400 (53%) cases. In cases with a uterine scar the trophoblast was anterior in 40% of cases, and posterior in 60% of cases (\( p = 0.2 \)). Uterine scars were classified according to Stirnemann et al.: there were 54 (54%) type 1a, 22 (24.2%) type 2a, 13 (14.3%) type 1b and 2 (2.2%) type 2b scars.

The achieving acquisition of a correct LUS image was considered easy, relatively easy, difficult and impossible in 56.4%, 30.8%, 12.2% and 0.6% of cases, respectively (Table 2). More than one image was required in 26.8% of cases. The mean (+SD) and medium (IQR) scores attributed by sonographer versus reviewer were 5.01 (+0.92) and 4 [4–6] versus 4.68 (+1.14) and 5 [4–5.24], \( p = 0.08 \). The mean

<table>
<thead>
<tr>
<th>Table 1. Measurements obtained by sonographers.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>CRL</td>
</tr>
<tr>
<td>Cervical length (mm)</td>
</tr>
<tr>
<td>Internal os-scar distance (mm)</td>
</tr>
<tr>
<td>Internal os-trophoblast distance (mm)</td>
</tr>
</tbody>
</table>

CRL, crown rump length.
Table 2. Ultrasound scans classified according to the difficulty encountered by the sonographer to obtain a correct LUS image.

<table>
<thead>
<tr>
<th>Evaluation of LUS</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Easy (few seconds)</td>
<td>480</td>
<td>56.4</td>
</tr>
<tr>
<td>2 = Relatively easy (less than 1 min)</td>
<td>262</td>
<td>30.8</td>
</tr>
<tr>
<td>3 = Difficult (more than 1 min)</td>
<td>104</td>
<td>12.2</td>
</tr>
<tr>
<td>4 = Impossible</td>
<td>5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

[95% CI] difference of $-0.33 \pm 2.62$ was recorded (Figure 3). The score equal or $\geq 4$ was recorded in 85.2% and 87.1% by sonographers and reviewers, respectively. There was good, moderate and poor agreement in 74.4%, 16.7% and 8.9% cases, respectively.

Variability in inter-reviewer and intra-reviewer was low with the mean [95% CI] difference of $-0.1 [-1.61,1.4]$ and $-0.1 [-1.41,1.2]$, respectively. Good agreement was found in 89.8% and 95.1%. AC1 values for inter- and intra-reviewer scores for each criterion are illustrated in Table 3. Either good or excellent agreement was noted for all criteria for the intra-reviewer scores. For the inter-reviewer scores, good or excellent agreement was noted for all criteria except no. 5 (the position of the internal os in the image).

Discussion

In this study, we introduced a standardized approach to LUS imaging involving transvaginal ultrasound at 11–14 weeks. This is a preliminary but necessary step to analyze the potential value of such assessment during early pregnancy.

A large proportion (87.2%) of sonographers found the LUS imaging to be easy or relatively easy, suggesting that the implementation of LUS imaging during the first trimester is feasible. An ultrasound at 11–14 weeks ultrasound is now routinely offered to all pregnant women in many countries [14] mainly for screening for fetal aneuploidy and diagnosis of some severe fetal malformations. Attempts have been made to predict severe obstetrical complications that may develop later in pregnancy at the time of this first trimester scan. One report investigated whether cervical length at 11–13 weeks, in combination with maternal characteristics, is a good predictor for preterm delivery [15]. Using these criteria, the authors were able to predict preterm delivery in 54% of cases, with a false positive rate of 10%. We did not attempt to correlate cervical length with obstetrical outcome in this flash study. However, correct imaging of the LUS is necessary to obtain accurate cervical measurements during the first trimester. The method for obtaining these measurements also requires standardization because different methods have been reported to give different results [12]. Potential anomalies of placental location could also be investigated in the first trimester scan. Mustafá et al. examined 351 singleton pregnancies in a prospective study. They found a relationship between the probability of placenta previa and the distance between the lower placental edge and the internal cervical at 11–14 weeks [6]. The morphological appearance of cesarean scars has also been examined during the first trimester. Stirnemann et al. [8] found that the transvaginal ultrasound could detect cesarean scars with a sensitivity of 82% and a specificity of 100%. This high sensitivity and specificity is also reflected in our data, although sonographers (for pragmatic reasons) were aware of previous CS. However, cesarean scars were undetectable in up to 14% of women who had had a CS. The proportions of the various types of scars were similar to those described by Stirnemann et al. [8]. Other authors studied the reproducibility of cesarean scar measurements in all trimesters of pregnancy finding good reproducibility for the first trimester comparing with moderate agreement for the second and third trimesters [11]. Signs of placenta acreta the first trimester have been described in several reports [16–18], moreover screening at 11–13 in women at risk was evaluated by some authors, allowing identifying most cases and reassuring patients otherwise at risk based on demographic characteristics [19]. Lastly, screening for pre-eclampsia by uterine artery Doppler combined with maternal serum markers at 11–13 weeks is promising and already routinely used in some clinical settings together with an appropriate standardized approach and quality control [20,21].

A standardized approach to transvaginal ultrasound is required to reproducibly evaluate the following factors in a single image of the LUS; cervix length, trophoblast-internal os distance, and in the case of CS, type of the uterine scar and scar-trophoblast distance to compare studies carried out in various institutions. We developed a criterion based scoring system to analyze the quality of LUS images, and evaluated the feasibility and reproducibility of this scoring system in an unselected population of women and sonographers. Our study shows that the implementation of such a standardized approach is feasible. Our score-based method is comparable to the Herman score for the measurement of nucal translucency [22]. In both cases, these scoring systems allow a somewhat objective assessment of image quality. We found good agreement in scores between sonographers and reviewers in 74.4% of cases. We found excellent inter-reviewer agreement in 89.8% of cases and excellent intra-reviewer agreement in 95.1% of cases. These findings demonstrate the reproducibility of this scoring system. Also, we found either good or excellent intra-reviewer agreement for each of the six criteria of our scoring system, and good or excellent inter-reviewer agreement for five of six criteria. Inexplicably, the criterion showed that poor inter-reviewer agreement was the position of internal os (but not its visualization) in the medium third of the image. Nevertheless, the visualization of the internal which remains the key criterion for the application of a standard lower segment image if it is to relate to low lying placenta or position of a previous CS scar showed good inter and intra-reviewer agreement.

This study is based on the concept of a “flash study” [23], which is characterized by a short duration, wide coverage, no modifications to obstetric management and no additional costs. Therefore, a major strength of the study is that it accurately reflects “real life” conditions. However, given the short duration of flash studies, we did not aim to evaluate the predictive value of LUS imaging during the first trimester.

In conclusion, a standardized approach to LUS imaging at 11–14 weeks is feasible and highly reproducible in a large population. Yet, there are limits to the ability to predict late...
pregnancy complications based on observations during the first trimester, this study emphasizes that every effort should be made to obtain high quality images that will allow risk stratification and the early recognition of women at risk. We believe a standardized and widely accepted approach to first trimester LUS imaging would help to better predict life-threatening conditions as early as first trimester.

Declaration of interest
The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

References


