Basic heart examination: feasibility study of first-trimester systematic simplified fetal echocardiography

E. QUARELLO*†, A. LAFOUGE‡, N. FRIES§, L. J. SALOMON¶ and the CFEF§

*Unité d’Échographies Obstétricales et de Diagnostic Anténatal, Hôpital Saint Joseph, Marseille, France; †Institut de Médecine de la Reproduction, Marseille, France; ‡Cabinet de Gynécologie et Obstétrique, Hyères, France; §Collège Français d’Échographie Fœtale, Chateaubriant, France; ¶Service d’Obstétrique et de Médecine Fœtale, Hôpital Necker-Enfants Malades, AP-HP, Université Paris Descartes, Paris, France

KEYWORDS: congenital heart disease; echocardiography; first trimester; low risk; screening

ABSTRACT

Objective First-trimester fetal cardiac screening examinations in low-risk populations should not have to meet the specifications required for high-risk populations. Our aim was to evaluate a simplified fetal echocardiographic (‘basic heart’) examination for early detection of severe congenital heart defects in a low-risk population.

Methods This was a first-trimester national ‘flash study’, performed over a 2-week period. Each observer was requested to perform simplified echocardiography without modifying the time and methods deemed necessary for the routine first-trimester ultrasound examination, in fetuses with crown–rump length between 45 and 84 mm. This basic heart assessment used targeted cross-sections of the four-chamber view (4CV) and of the three vessels and trachea (3VT) view, using color and/or directional power Doppler. All examinations were then reviewed offline and scored for quality by a qualified expert.

Results Sixty observers performed a total of 597 first-trimester ultrasound examinations, each performing an average of 10 (range, 1–26) procedures. Examinations were conducted transabdominally (79%; 472/597), transvaginally (3%; 17/597) or both (18%; 108/597). In 13% (78/597) it was technically infeasible and in 8% (47/597) it was deemed feasible but atypical, which may have been due to the presence of an abnormality or to poor quality of the image. The 3VT view was obtained successfully and was normal in 79% (472/597) of the patients, in 13% (78/597) it was technically infeasible and in 8% (47/597) it was deemed feasible but atypical. Both the 4CV and the 3VT view were obtained successfully and were normal in 73% (435/597) of patients.

Conclusion It is possible for a trained operator to perform simplified fetal echocardiography during the routine first-trimester ultrasound examination in a low-risk population. In most cases, our basic heart examination can be used to reassure parents or identify potential problems to be clarified as early as possible in the second trimester. Copyright © 2016 ISUOG. Published by John Wiley & Sons Ltd.

INTRODUCTION

Congenital heart defects (CHDs) are a leading cause of death during the first year of life, affecting between five and nine per 1000 live births in the general population. Prenatal ultrasound is the gold standard for screening and diagnosing these malformations; fetal echocardiography has the ability to detect 60–100% of cases of severe CHD1,2 during the second trimester. However, almost all CHDs are already established by the first trimester and, over the past 15 years, there has been an increase in screening and diagnosis during this period. Several groups have reported the possibility and relevance of first-trimester fetal echocardiography, whether transabdominal3, transvaginal4 or both, particularly in high-risk populations5. Most first-trimester screening is in these high-risk populations, carried out by experts in tertiary centers; it has been shown that these
Echocardiographic examinations are particularly effective at reassuring couples by excluding a CHD in apparently normal cases with a very high negative predictive value (98.9% (95% CI, 98.1–99.4%))⁶. In fact, most CHDs occur in low-risk populations⁷, but few studies have focused on first-trimester fetal heart examination in these groups⁴,⁸,⁹. Our national recommendations¹⁰, as well as those of the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG)¹¹, propose that low-risk populations receive a heart-rate evaluation during the first trimester. In addition, ISUOG proposes optional evaluation of the symmetry of the four cardiac chambers.

We believe that first-trimester fetal cardiac screening examinations in low-risk populations do not need to meet the specifications required for high-risk populations. We therefore propose a simplified or ‘basic’ fetal echocardiographic examination to achieve early detection of severe CHD. We anticipated that such CHDs would be of the functionally univentricular type, amenable to skilled palliative surgery, such as mitral and tricuspid atresia, some complete AVSD, and some abnormalities of the main vessels such as aortic and pulmonary atresia, truncus arteriosus, and some transpositions of the great arteries or double outlet ventricle. Major defects, such as complete obstruction of any of the four cardiac valves, would also be evident from this type of screening. Such defects are associated with less successful surgical repair compared with other forms of CHD and termination of pregnancy may be an appropriate form of management. Data suggest that psychological recovery after termination of pregnancy for fetal malformation is better the earlier the diagnosis is made¹².

Here we present our ‘flash study’ on the feasibility of a simplified fetal echocardiographic examination performed by trained observers in low-risk, first-trimester populations.

**METHODS**

The concept of the flash survey, a short-term study with wide coverage, resulting in neither modification of obstetric and fetal care nor additional cost, and with an underlying educational message, was introduced by L.J.S. at the College of French Fetal Ultrasound (CFEF) 4ᵗʰ Scientific Meeting (Port-en-Bessin-Huppain, France, 1–3 October 2010). Our 15-day study was performed by CFEF between 22 April and 6 May 2014. CFEF members from private and/or public centers were recruited by email; those agreeing to participate were registered via a web portal and received a background questionnaire to complete, as well as an explanatory protocol. Briefly, during the study period, these observers were to record consecutively all patients on whom they were carrying out routine first-trimester ultrasound screening, in fetuses with crown–rump length (CRL) between 45 and 84 mm, and, in addition, attempt to perform a simplified echocardiographic examination; this involved acquisition of only two views rather than the usual five or six obtained during the routine second- or third-trimester scan, and they were required to do so without modifying the scan time and using only those methods that they deemed necessary for the routine first-trimester ultrasound examination.

First, the mother’s birth date, weight and height were recorded. Fetal ultrasound was performed either transabdominally, transvaginally or both. Position of the fetal back during the cardiac examination (anterior, left lateral, posterior or right lateral) was documented, along with the CRL. Nuchal translucency thickness (NT) was measured and a digital ultrasound image of the measurement was stored in an electronic database for later evaluation by a qualified expert assessor (E.Q.). The observer then attempted to obtain and store in the database two digital ultrasound images for the simplified echocardiographic examination, referred to as ‘basic heart’. This exam targeted, via color Doppler and/or directional power Doppler, particular cross-sections of the four-chamber view (4CV) and three vessels and trachea (3VT)¹³ view (Figure 1) that met specific criteria which had been given to the observers in advance (Table 1). We had identified these criteria in order to reflect the quality of the information in the images, and they were also used by the expert assessor to establish scores for each image. However, although the intention originally was to score images for quality, it became clear during scoring that, unlike the Herman quality score for NT measurement, it was difficult to assess the quality of the 4CV and 3VT views without taking into account normality of the heart itself. The qualified expert assessor (E.Q.) established ad-hoc scores, both for the NT images, using the Herman score for NT measurement (range, 0–9)¹⁴,¹⁵, for the 4CV and 3VT view cross-sections as defined in Table 1 (range, 0–8; a score ≤2 was defined as non-informative or technically infeasible, a score of 3–7 was defined as suboptimal or unusual and a score of 8 was defined as optimal and informative). Additionally, in order to study the inter- and intraobserver reproducibility of these scores for the basic heart examination, 50 cross-sections from both 4CV and 3VT view were selected arbitrarily from the image bank and two reviewers (E.Q. and A.L.) were asked to score each of these 100 images. For the intraobserver reproducibility, these scores were compared with the earlier scores by E.Q. of the same set of images. The scores of 4CV and 3VT view images were analyzed to assess their relation to maternal body mass index (BMI) category (< 20, 20–25, > 25), fetal CRL category (45–54 mm, 55–64 mm, 65–74 mm, 75–84 mm), and fetal back position (anterior, posterior, left lateral, right lateral). We also studied the relationship between basic heart score and NT Herman score (< 5, 5–7, > 7).

Scores for 4CV and 3VT views were analyzed by Kruskal–Wallis ANOVA test and/or Mann–Whitney U-test. Inter- and intraobserver scoring variability was assessed by percentages of the scoring agreement. The Friedman test was used for the significance of variability in inter- or intraobserver scores. Agreement coefficients (AC) for each criterion were computed as described.
by Gwet\textsuperscript{16} to test the reproducibility of each criterion. AC was calculated as follows: $P_a - P_e(\gamma)/(1 - P_e(\gamma))$, where $P_e(\gamma)$ is given by: $P_e(\gamma) = 2P_+/P_+ + 1$. An AC $<0.20$ was considered to indicate poor agreement, $0.21-0.40$ was fair, $0.41-0.60$ was moderate, $0.61-0.80$ was good and $>0.81$ was excellent agreement\textsuperscript{16}. Statistical analyses were performed using Stata 11 for Windows (StataCorp LP, TX, USA) and Statistica (Stat Soft, France), with $P < 0.05$ considered statistically significant for all tests.

RESULTS

Study population

In this flash study, 60 observers (12 midwives and 48 physicians) performed a total of 597 first-trimester ultrasound examinations. The background questionnaires completed by each observer revealed that 6.7\% (4/60) performed fewer than 500 ultrasound examinations, 16.7\% (10/60) performed between 500 and 1000 and 76.6\% (46/60) performed more than 1000 per year. Furthermore, 25\% (15/60) of the observers were already performing first-trimester fetal cardiac examinations (without specifying the nature of their examination) routinely in the low-risk general population, 55\% (33/60) were doing so in cases at high risk for CHD and 48\% (29/60) initiated second-trimester cardiac ultrasound examinations in high-risk CHD fetuses without systematically consulting a fetal cardiologist, irrespective of whether any early cardiac examination had been performed.

On average, each observer performed 10 (range, 1–26) examinations during the 15 days of the study. The median ± SD maternal BMI was 22.75 ± 4.58. There were 472 (79\%) transabdominal examinations, 17 (3\%) transvaginal examinations and 108 (18\%) using both approaches. The CRL was between 45 and 54 mm in 11\% (65/597) of cases, between 55 and 64 mm in 46\% (273/597), between 65 and 74 mm in 35\% (212/597) and between 75 and 84 mm in 8\% (47/597) of cases. During the simplified cardiac examination, the back of the fetus was anterior in 8\% (45/597) of cases, left-lateral in 18\% (108/597), posterior in 63\% (377/597) and right-lateral in 11\% (67/597) of cases. The Herman score for NT quality was $<5$ in 2.5\% (15/597) of cases, 5–7 in 6.5\% (39/597) and $>7$ in 91\% (543/597) of cases.
Basic heart examination

Analysis of the scores established by the expert revealed that the 4CV was obtained successfully and was deemed normal (score 8/8) in 86% (512/597) of cases, the image was described as technically infeasible (score ≤ 2/8) in 7% (41/597) of cases, and it was described as feasible but atypical, which may have been due to the presence of an abnormality or to poor quality of the image (score 3–7) in 7% (44/597) of cases. A 3VT view cross-section was obtained successfully and was normal in 79% (472/597) of cases, the image was described as technically infeasible in 13% (78/597) of cases, and it was deemed as feasible but atypical in 8% (47/597) of cases. Both the 4CV and the 3VT view were obtained successfully and were normal in 73% (435/597) of patients.

Factors potentially influencing success of basic heart examination

Figure 2 shows the relationship between image scores, as determined by E.Q., and maternal BMI, position of the fetal spine, fetal CRL and Herman scores for corresponding NT images, respectively. The scores for the 4CV and for the 3VT view cross-sections did not differ significantly between patients grouped according to BMI < 20, BMI 20–25 or BMI > 25 (P = 0.17 and P = 0.1, respectively). However, patients with BMI < 20 had significantly higher 4CV and 3VT view scores than did those with BMI > 25 (P = 0.02 and P = 0.03) (Figure 2a). While the scores for the cross-sections of the 4CV were not significantly different according to the position of the fetal spine (P = 0.06), scores for the 3VT view did differ significantly when the position of the fetal spine was anterior (P = 0.0003) (Figure 2b). The scores for the 4CV and for the 3VT view cross-sections did not differ significantly when patients were grouped according to CRL (P = 0.12 and P = 0.08, respectively). However, patients with CRL < 55 mm had significantly lower 4CV cross-sectional scores (P = 0.01) than did those with CRL ≥ 75 mm; this was not the case for the 3VT view cross-sections (P = 0.17) (Figure 2c). The scores for both the 4CV and the 3VT view cross-sections differed significantly according to the NT Herman scores (P = 0.0001 for both) (Figure 2d):
Reproducibility of basic heart examination scores

Examination of the AC according to Gwet\(^\text{16}\) indicated excellent intra- and interobserver reproducibility for all items used to rate the cross-sections of the 4CV and the 3VT view (Table 2), with intraobserver ACs of 0.9–1 and 0.94–1, respectively, and interobserver ACs of 0.89–1 and 0.81–1, respectively.

Table 2  Intra- and interobserver reproducibility for first-trimester basic heart image scoring criteria for four-chamber view (4CV) and three vessels and trachea (3VT) view

<table>
<thead>
<tr>
<th>Criterion†</th>
<th>Agreement coefficient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>4CV</td>
<td></td>
</tr>
<tr>
<td>Symmetrical axial chest view</td>
<td>1</td>
</tr>
<tr>
<td>Visualization of spine and initial part of one rib to each side</td>
<td>0.95</td>
</tr>
<tr>
<td>Two atrioventricular filling flow patterns</td>
<td>0.93</td>
</tr>
<tr>
<td>Two separate atrioventricular filling flow patterns</td>
<td>0.97</td>
</tr>
<tr>
<td>Two symmetrical atrioventricular filling flow patterns</td>
<td>0.90</td>
</tr>
<tr>
<td>3VT view</td>
<td></td>
</tr>
<tr>
<td>Symmetrical axial chest view</td>
<td>1</td>
</tr>
<tr>
<td>Visualization of spine and initial part of one rib to each side</td>
<td>0.97</td>
</tr>
<tr>
<td>Two separate filling flow patterns</td>
<td>0.97</td>
</tr>
<tr>
<td>Convergence of the two filling flow patterns</td>
<td>0.97</td>
</tr>
<tr>
<td>Two symmetrical filling flow patterns</td>
<td>0.94</td>
</tr>
<tr>
<td>Two antegrade filling flow patterns</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*Agreement coefficients calculated according to Gwet\(^\text{16}\). †Scoring criteria as described in Table 1.

DISCUSSION

We propose a first-trimester cardiac examination that is easy and quick to perform, without requiring any change to the methods used for routine ultrasound screening. We have demonstrated that trained observers can perform this examination in a low-risk CHD population by obtaining color and/or directional power Doppler cross-sections of the 4CV and 3VT view, using predefined quality criteria.

It is now well established that fetuses at high risk for CHD should undergo fetal echocardiography as soon as possible, ideally during the late first trimester\(^\text{17,18}\), and it has been shown that echocardiography before 15 weeks has high specificity and a good negative predictive value for the absence of CHD\(^\text{6,15}\). However, often, the integrity of the interventricular septum and the pulmonary veins may be analyzed only suboptimally this early in gestation, and patients are therefore reviewed systematically later in the second trimester.

Assessing whether the 4CV and 3VT view cross-sections are unusual in appearance is more difficult using just gray-scale ultrasound, particularly when using a transabdominal approach, than it is with color and/or directional power Doppler modes; this is due to the probes’ limited resolution, even when using high-frequency probes, as well as to the depth of the structures under analysis. We found that a transabdominal examination could be performed quite easily in most cases. A transvaginal examination could, however, be performed in cases of poor visualization transabdominally or to complement the transabdominal exam in high-risk populations.

Despite being asked not to alter their examination time and methods, most observers who were familiar with the cross-sectional protocol criteria were able to achieve the additional views (in 86% of patients for the 4CV, in 79% for the 3VT view and in 73% for both). Maternal BMI, fetal spine position and CRL apparently influenced the feasibility of obtaining the cross-sections and/or affected whether they were informative, and when there was no statistically significant difference associated with one of these variables, there was nonetheless a trend; the lack of significance might be explained by the number of operators participating in the study and especially the number of fetuses assessed. Our results are in agreement with previous findings that CRL, and in particular CRL ≥ 75 mm, greatly influences whether an informative first-trimester fetal echocardiogram is achieved\(^\text{20,21}\). As expected, there was an association between obtaining a good NT Herman score (≥ 5) and obtaining satisfactory cardiac cross-sections (Figure 2d).

Among our 597 cases, 7% (n = 41) of 4CV images and 13% (n = 78) of 3VT views were deemed not technically feasible, and 7% (n = 44) and 8% (n = 47), respectively, were viewed as feasible but atypical. In our view, most of these cases were likely due to a lack of operator experience in setting the Doppler modes correctly during a first-trimester examination; in our experience, once a learning curve has been achieved, few images are technically impossible to obtain, and therefore few are categorized as ‘infeasible or atypical’. We elected to review these cases with a second-trimester scan so as not to generate unnecessary potential anxiety or add to the department’s workload.

First-trimester screening relies on proper settings for color and/or directional power Doppler in the same way that screening later in pregnancy does. Standard examination recommendations for the use of color and power Doppler should always be followed carefully and these modes should be used sparingly and appropriately. Operators should first optimize their gray-scale settings and then adjust the pulse repetition frequency, vascular filter, color gain, energy gain, power and transmission frequencies; these modifications will generally render informative images (Figures 3 and 4).

Our study has several strengths: this is the first such study of the feasibility of a first-trimester basic heart examination; the study was multicentric as opposed to previous studies in individual centers; the majority of observers were not echocardiography experts; and we...
First-trimester basic heart ultrasound examination

Figure 3 Examples of unusual first-trimester basic heart four-chamber views. In (a), color fill is not set correctly, so two separate atrioventricular blood flows are not visualized (score: 1/1/2/0/0 = 4). Image (b) shows asymmetry between two atrioventricular blood flows, but it is not possible to determine whether this asymmetry is related to abnormality or to a suboptimal setting (score: 1/1/2/2/0 = 6). Image (c) identifies only two intraventricular blood flows instead of two atrioventricular blood flows, due to a suboptimal setting (score: 1/1/0/0/2 = 4). Ant, anterior; L, left; Post, posterior; R, right.

Figure 4 Examples of unusual first-trimester basic heart three vessels and trachea views. Image (a), probably due to suboptimal Doppler adjustment, does not identify two distinct blood flows converging, but assesses only grossly a single antegrade vascular structure (score: 1/1/0/0/0 = 2). Level of image (b) is too low, corresponding to three-vessel view defined by Yoo et al. instead of three vessels and trachea view of Yagel et al.; there are two vascular structures visible, but because of absence of visualization of ‘V’-shaped converging blood flows, it is not possible to exclude interruption of aortic arch (score: 1/1/2/1/1 = 6). In image (c), because of lateral position of ultrasound beam, there is poor visualization of vascular flow in aorta. On imaging without Doppler, aortic arch was visible, but because of chest position this arch is nearly perpendicular to ultrasound beam, so blood flow within arch could not be displayed with color Doppler (score: 1/1/2/0/1 = 5). Ant, anterior; L, left; Post, posterior; R, right.

present new criteria for assuring and assessing the quality of fetal echocardiography when obtaining 4CV and 3VT view cross-sections through the use of color and/or power Doppler modes.

A major limitation is the possible bias in recruitment of observers. It is likely that those choosing to participate were responsive to such studies and wished to improve their ultrasound proficiency. The distribution of observer experience (6.7% performing fewer than 500 ultrasound examinations/year and 76.6% performing more than 1000 examinations/year) and the fact that 91% of NT Herman scores were > 7/9 demonstrate that our population of observers was well trained; these results might not have been duplicated if we had used less well-trained observers. Another limitation is that our work did not address the relevance of CHD detection via a simplified cardiac examination; feasibility and relevance are two different issues. However, Wiechec et al. recently demonstrated that a single-center study conducted by three trained observers had very good sensitivity and specificity in detecting severe CHD by simplified cardiac examination; this issue should be the subject of future study.

Another limitation of this study is the use of the Herman score, which assesses the quality of the NT measurement regardless of whether it is thick (> 95th percentile for the CRL) or thin, at the same time as applying our criteria to obtain 4CV and 3VT view cross-sections via color and/or power Doppler modes. Our criteria were originally intended, similar to the Herman score, to assess quality of the 4CV and 3VT view regardless of their normality; however, it proved very difficult to specify purely quality criteria, and the score was more a reflection of whether an image was unusual in nature, relative to its expected...
appearance. Such atypicality may be due either to the presence of an abnormality or to poor quality of the image.

In conclusion, it appears important to encourage development of the concept of a simplified first-trimester echocardiographic examination in populations at low risk of CHD. We have shown that this is possible to achieve using color and/or power Doppler modes. In most cases our basic heart examination can be used to reassure parents or identify potential problems to be clarified as early as possible in the second trimester. The appropriateness of using this new strategy for screening for major CHD should be evaluated on a large scale and its value compared with that of current obstetric ultrasound screening in the second and third trimesters of pregnancy.

ACKNOWLEDGMENT

We thank all colleagues, doctors and midwives for their enthusiasm in participating in this study and making it possible.

REFERENCES


