Ultrasound Obstet Gynecol 2016; **47**: 81–88
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.15740

Contingent *versus* routine third-trimester screening for late fetal growth restriction

S. TRIUNFO*†, F. CROVETTO*†‡, E. SCAZZOCCHIO§, M. PARRA-SAAVEDRA¶, E. GRATACOS*† and F. FIGUERAS*†

**Fetal i+D Fetal Medicine Research Center, BCNatal - Barcelona Center for Maternal-Fetal and Neonatal Medicine (Hospital Clínic and Hospital Sant Joan de Deu), IDIBAPS, University of Barcelona, Barcelona, Spain; †Centre for Biomedical Research on Rare Diseases (CIBER-ER), Valencia, Spain; ‡Fondazione Cà Granda, Ospedale Maggiore Policlinico, Dipartimento Ostetricia e Ginecologia, Università degli Studi di Milano, Milan, Italy; §Obstetrics, Gynecology and Reproductive Medicine Department, Institut Universitari Dexeus, Barcelona, Spain; ¶Maternal-Fetal Unit, CEDIFETAL, Centro de Diagnóstico de Ultrasonido e Imágenes, CEDIUL, Barranquilla, Colombia*

KEYWORDS: estimated fetal weight; fetal growth restriction; neonatal complications; ultrasonography; uterine artery Doppler

ABSTRACT

Objective *To evaluate the use of third-trimester ultrasound screening for late fetal growth restriction (FGR) on a contingent basis, according to risk accrued in the second trimester, in an unselected population.*

Methods *Maternal characteristics, fetal biometry and second-trimester uterine artery (UtA) Doppler were included in logistic regression analysis to estimate risk for late FGR (birth weight $< 3^{rd}$ percentile, or 3^{rd}–10^{th} percentile plus abnormal cerebroplacental ratio or UtA Doppler, with delivery ≥ 34 weeks). Based on the second-trimester risk, strategies for performing contingent third-trimester ultrasound examinations in 10%, 25% or 50% of the cohort were tested against a strategy of routine ultrasound scanning in the entire population at 32 + 0 to 33 + 6 weeks.*

Results *Models were constructed based on 1393 patients and validated in 1303 patients, including 73 (5.2%) and 82 late FGR (6.3%) cases, respectively. At the second-trimester scan, the* a-posteriori *second-trimester risk (*a-posteriori *first-trimester risk (baseline* a-priori *risk and mean arterial blood pressure) combined with second-trimester abdominal circumference and UtA Doppler) yielded an area under the receiver–operating characteristics curve (AUC) of 0.81 (95% CI, 0.74–0.87) (detection rate (DR), 43.1% for a 10% false-positive rate (FPR)). The combination of* a-posteriori *second-trimester risk plus third-trimester estimated fetal weight (full model) yielded an AUC of 0.92 (95% CI, 0.88–0.96) (DR, 74% for a 10% FPR). Subjecting 10%, 25% or 50% of the study population to third-trimester ultrasound, based on* a-posteriori *second-trimester risk, gave AUCs of 0.81 (95% CI, 0.75–0.88), 0.84 (95% CI, 0.78–0.91) and 0.89 (95% CI, 0.84–0.94), respectively. Only the 50% contingent model proved statistically equivalent to performing routine third-trimester ultrasound scans (AUC, 0.92 (95% CI, 0.88–0.96),* P = *0.11).*

Conclusion *A strategy of selecting 50% of the study population to undergo third-trimester ultrasound examination, based on accrued risk in the second trimester, proved equivalent to routine third-trimester ultrasound scanning in predicting late FGR.* *Copyright © 2015 ISUOG. Published by John Wiley & Sons Ltd.*

INTRODUCTION

Annually, approximately 400 000 pregnancies in Europe alone can be complicated by fetal growth restriction (FGR). However, antenatal detection falls short in clinical practice, failing to recognize up to 75% of babies at risk of FGR late in pregnancy[1]. Such poor performance takes a toll in terms of perinatal health, given that small babies who are overlooked have higher risk of adverse perinatal outcome and stillbirth[2,3].

There are several baseline risk factors that should raise suspicion of FGR[4]. Nevertheless, the predictive performance of these factors is poor, with respective detection rates (DRs) of 50% and 20% for early- and late-onset FGR[5]. Uterine artery (UtA) Doppler (to assess trophoblastic invasion) has been proposed as a screening tool, yielding DRs of about 25% and 75% in the first and second trimesters, respectively[6,7]. By combining UtA Doppler findings and baseline maternal characteristics, DRs for early-onset FGR reach clinically acceptable levels[5,8]. Unfortunately, however, FGR developing late

Correspondence to: Dr F. Figueras, Maternal-Fetal Medicine Department, Hospital Clinic, University of Barcelona, Sabino de Arana 1, 08028 Barcelona, Spain (e-mail: ffiguera@clinic.ub.es)

Accepted: 2 September 2015

in pregnancy still goes largely undetected[9–12], despite representing the largest contributing factor to adverse perinatal outcome and stillbirth[10].

Current growth screening strategies involve measuring the symphysis–fundus height, but fewer than 25% of small-for-gestational-age (SGA) infants will be detected using this methodology in a low-risk population[13]. Routine third-trimester ultrasound examination (US) for fetal growth assessment, in place in many countries, has DRs ranging from 50% to 80%[2,11], but the beneficial impact on perinatal outcome is unclear. Eight trials were included in a recently updated meta-analysis[12], which concluded that routine US in late pregnancy in low-risk or unselected populations does not confer benefit on mother or baby. Furthermore, serial third-trimester US has not been demonstrated to improve this performance[14,15].

Because there is no evidence to support routine performance of third-trimester US, it seems reasonable to select women who are at the highest risk, for whom third-trimester growth assessment may be most effective.

The aim of this study was to evaluate the use of third-trimester US, performed on a contingent basis (according to risk accrued in the second trimester) in an unselected population, as a means of predicting late FGR.

METHODS

Study population

Between January 2010 and December 2012, a prospective cohort of consecutive singleton pregnancies was recruited from those attending the Department of Maternal–Fetal Medicine in the Hospital Clinic of Barcelona for routine first-trimester aneuploidy screening (8 + 0 to 13 + 6 weeks' gestation). Calculated gestational age was based on crown–rump length at first-trimester US[16]. We excluded any pregnancy with aneuploidy or major fetal abnormality, and those involving termination, miscarriage or fetal death, suspected FGR (estimated fetal weight (EFW)[17] < 10th percentile according to local standards[18]) or pre-eclampsia before 32 weeks. The local ethics committee approved the study protocol, and each patient gave written informed consent to participate (IRB 2010/5736).

Sample size estimation

We estimated the sample size necessary for pairwise comparison of areas under the receiver–operating characteristics curve (AUC), given a non-inferiority hypothesis and assuming a rank correlation coefficient of 0.2 for predictive variables in both late FGR and non-FGR states. Based on prior evidence[11], an AUC value of 0.9 was expected for routine third-trimester US imaging. Differences in AUC > 15% were considered relevant. For a Type-I error rate of 5% and a target power of 90%, the number of required cases of late FGR was 65 (for rank correlation coefficient of 0.2). Assuming a 5% prevalence of late FGR, it was estimated that a maximum total sample size of 1300 was necessary[19]. Because a validation cohort of the same size as the construction cohort was required, the final sample size was set at 2600.

Definition of outcome measures and clinical management

Late FGR was defined as birth weight < 3rd percentile according to local standards[18], or 3rd–10th percentile with prenatal abnormalities of cerebroplacental ratio (< 5th percentile)[20] or UtA Doppler pulsatility index (PI) (> 95th percentile)[21], with delivery from 34 weeks onwards. In pregnancies complicated by FGR, a specific clinic protocol was applied[22]. Pre-eclampsia was defined according to the International Society for the Study of Hypertension in Pregnancy[23] guidelines. All women underwent continuous fetal monitoring during labor, using a three-tiered classification for heart tracings[24]. Non-reassuring fetal status during labor was defined as a pathological fetal heart rate[23] or a suspicious tracing with fetal blood scalp sampling pH < 7.15 or < 7.20 in two samples 30 min apart. Neonatal metabolic acidosis at birth was defined as umbilical arterial pH < 10th percentile (< 7.15) and base excess > 90th percentile (12 mEq/L)[25].

Predictive variables

1. *Maternal characteristics*

Data on maternal characteristics, including age, ethnicity, low socioeconomic status (routine occupation, long-term unemployment or never worked), body mass index (BMI), nulliparity (no prior deliveries after 22 weeks' gestation), smoking status, method of conception (spontaneous or assisted reproductive technique (ART) including ovulation induction, *in-vitro* fertilization and oocyte donation), medical history (known chronic disease, such as hypertension, diabetes mellitus, renal disease, autoimmune disorder and congenital or acquired thrombophilic conditions) and obstetric history (including prior pregnancy complicated by stillbirth, FGR or pre-eclampsia) were recorded in the hospital database at study inclusion. In addition, all data pertaining to follow-up, developing complications, US evaluations and perinatal conditions were collected prospectively.

2. *Biophysical parameters*

At first-trimester screening, blood pressure was recorded by a nurse in our outpatient clinic, in accordance with standard procedure, with subjects seated after a 5-min rest. An automated calibrated device (M6 Comfort; Omron Corp., Kyoto, Japan) was used, selecting either right or left arm arbitrarily. Mean arterial pressure (MAP) was calculated as: diastolic BP + (systolic BP – diastolic BP)/3.

3. *Ultrasound evaluation*

First-trimester screening for chromosomal abnormalities by a combination of fetal nuchal translucency thickness and biochemical measurements (the latter obtained previously, at 8 + 0 to 9 + 6 weeks) was offered at 11 + 0 to 13 + 6 weeks' gestation.

Second- and third-trimester US (generally performed at 19 + 0 to 21 + 6 and at 32 + 0 to 33 + 6 weeks, respectively) included measurement of the following biometric parameters: biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC) and femur length (FL). All measurements were obtained by one of seven trained sonographers, adhering to recommended techniques[26]. In the third trimester, EFW was also calculated, using the Hadlock formula[17].

4. *Doppler measurements*

Prenatal Doppler studies were conducted using either a Siemens Sonoline Antares (Siemens Medical Systems, Malvern, PA, USA) or a GE Voluson E8 (GE Medical Systems, Zipf, Austria) US machine, each equipped with a 6–2-MHz linear curved-array transducer. UtA Doppler evaluations were performed using a transabdominal approach in the second trimester, according to recommended methodology[21]. Mean PI was calculated as the average of right and left arterial values.

Statistical analysis

The cohort was subdivided arbitrarily (1:1 ratio) into construction and validation sub-cohorts. Student's *t*-test, Pearson's chi-square test or Fisher's exact test were performed to make univariate comparisons of quantitative and qualitative variables, as appropriate. A binomial distribution model was used to determine the 95% CIs of proportions.

The following steps (also see Appendix S1) were applied to the construction cohort to develop the full model for predicting late FGR:

1. Logarithmic transformations to normalize mean MAP, UtA-PI and all fetal biometric parameters.
2. Multiple linear regression (forward stepwise selection, with *P*-value cut-off points of 0.05 for inclusion/exclusion) formulae were calculated to derive first-trimester expected log MAP; second-trimester expected log UtA-PI, BPD, HC, AC and FL; and third-trimester expected log BPD, HC, AC, FL and EFW.
3. Individual expected log values were then predicted for both affected and unaffected pregnancies.
4. From expected log values calculated in steps 2 & 3, multiples of the median (MoM) were calculated for each subject as follows: observed log value – expected log value.
5. Receiver–operating characteristics (ROC) curves to predict late FGR were constructed for each predictor (log MoM values).

Multiple logistic regression analysis (forward stepwise selection with *P*-value cut-offs for inclusion/exclusion of 0.05) was performed for individual estimates of the following late FGR risks:

6. *A-priori* risk (covariables: maternal age, BMI, ethnicity, low socioeconomic status, nulliparity, smoking, ART, previous obstetric complications, medical diseases).
7. *A-posteriori* first-trimester risk (covariables: nuchal translucency, MAP, mean UtA-PI, *a-priori* risk).
8. *A-posteriori* second-trimester risk (covariables: biometric measures, mean UtA-PI, *a-posteriori* first-trimester risk).
9. *A-posteriori* third-trimester risk (full model) (covariables: biometric measures, EFW, *a-posteriori* second-trimester risk).

In the validation cohort, predictive performance of the various strategies stipulated for contingent third-trimester scanning (based on second-trimester *a-posteriori* risk and conducted in 10%, 25% or 50% of the study population) were delineated in ROC curves, and each ROC curve was tested[27] against the full model (i.e. routine third-trimester US).

With highly correlated measurements ($R > 0.75$), such as US biometric measures, only a variable with the largest individual area under the ROC curve was entered into a model for regression.

Assumptions for regression were checked in each model, and goodness-of-fit was assessed by calculating R^2 or Nagelkerke R^2 for linear and logistic models, respectively.

Statistical software R, version 2.15.1 (The R Foundation for Statistical Computing), with the package pROC version 1.7.2, was used for all statistical analysis and graph construction.

RESULTS

Of 2775 women recruited in the first trimester of pregnancy, 79 (2.8%) were excluded for the following reasons: lost to follow-up ($n = 20$); miscarriage ($n = 13$); congenital malformation/chromosomal abnormality ($n = 20$); termination of pregnancy without medical indication ($n = 5$); stillbirth (< 32 weeks) ($n = 6$); early-onset (delivery < 32 weeks) FGR/pre-eclampsia ($n = 15$). A total of 155 (5.7%) of the 2696 women analyzed met the criteria for late FGR. Women were divided arbitrarily into construction ($n = 1393$, including 73 late-FGR fetuses and 1320 controls) and validation ($n = 1303$, including 82 late-FGR fetuses and 1221 controls) sub-cohorts. Table S1 details the epidemiological and perinatal characteristics of each sub-cohort. Of note, significant differences between construction and validation sub-cohorts were found only in proportions of those with white ethnicity (63.1% *vs* 59.3%), those with chronic hypertension (0.9% *vs* 2.1%) and those with gestational diabetes (6.8% *vs* 4.9%).

Epidemiological, clinical characteristics and perinatal outcomes of the validation cohort are summarized

Table 1 Epidemiological and clinical characteristics and maternal and neonatal outcomes of the validation population according to late fetal growth restriction (FGR) status

Characteristic	*No FGR (n = 1221)*	*Late FGR (n = 82)*	P*
Age (years)	31.7 ± 5.3	31.9 ± 5.6	0.816
BMI (kg/m^2)	23.9 ± 3.7	24.2 ± 4.7	0.551
Ethnicity			0.51
White	715 (58.6)	57 (69.5)	
Latin-American	113 (9.3)	11 (13.4)	
Other	393 (32.4)	14 (17.1)	
Low socioeconomic status†	524 (42.9)	47 (57.3)	0.428
Smoking status			< 0.001
None	1077 (88.2)	53 (64.6)	
1–10 cigarettes/day	121 (9.9)	17 (20.7)	
11–20 cigarettes/day	23 (1.9)	12 (14.6)	
Medical history			
Chronic hypertension	23 (1.9)	4 (4.9)	0.062
Diabetes mellitus	30 (2.5)	0 (0)	0.154
Renal disease	11 (0.9)	2 (2.4)	0.169
Autoimmune disease	20 (1.6)	4 (4.9)	0.032
Coagulation disorder	18 (1.5)	5 (6.1)	0.002
Obstetric history			
Nulliparous	718 (58.8)	57 (69.5)	< 0.001
Previous FGR	10 (0.8)	8 (9.8)	< 0.001
Previous pre-eclampsia	10 (0.8)	7 (8.5)	< 0.001
Previous stillbirth	13 (1)	2 (2.4)	0.251
Previous preterm birth	15 (1.2)	3 (3.7)	0.064
ART	54 (4.4)	4 (4.9)	0.826
Gestational hypertension	25 (2.1)	12 (14.6)	< 0.001
Pre-eclampsia (late onset)	33 (2.7)	8 (9.8)	< 0.001
HELLP syndrome	3 (0.3)	0 (0)	0.655
Gestational diabetes	60 (4.9)	4 (4.9)	0.991
GA at delivery (weeks)	39.8 ± 1.5	38.2 ± 2.7	< 0.001
Cesarean section	326 (26.7)	26 (31.7)	0.151
Cesarean section for NRFS	99 (8.1)	8 (9.8)	0.365
Birth weight (g)	3370 ± 416.3	2289.5 ± 460.2	< 0.001
Birth-weight percentile	51.2 ± 29.5	1.6 ± 1.7	< 0.001
Male gender	661 (54.1)	24 (29.3)	< 0.001
5-min Apgar < 7	11 (0.9)	1 (1.2)	0.760
Neonatal metabolic acidosis	83 (6.8)	11 (13.4)	0.020
NICU admission	2 (0.2)	0 (0)	0.716
Perinatal death	1 (0.1)	0 (0)	0.797

Data are expressed as mean ± SD or *n* (%). *Student's *t*-test, Pearson's chi-square test or Fisher's exact test. †Routine occupation, long-term unemployment or never worked. ART, assisted reproductive technique; BMI, body mass index; GA, gestational age; HELLP, hemolysis elevated liver enzymes, and low platelets; NICU, neonatal intensive care unit; NRFS, non-reassuring fetal status.

in Table 1. Maternal demographics (smoking status, autoimmune disease, coagulation disorder, nulliparity, previous FGR and previous pre-eclampsia) differed significantly between pregnancies with late FGR and those without FGR. As expected, perinatal outcomes in terms of late pre-eclampsia, birth weight and neonatal metabolic acidosis were poorer in FGR pregnancies.

Biophysical and US variables of the validation cohort in first, second and third trimesters are reported in Table 2, in which AUCs are detailed individually for each parameter.

The following models best fit specific risks for late FGR:

A-priori risk

Logit = − 1.613 − 0.064 × BMI + 0.133 if Latin-American ethnicity − 0.746 if nulliparous + 1.506 if smoker + 1.036 if previous adverse obstetric outcome; $R^2 = 10\%$

A-posteriori first-trimester risk

Calculation of Log MAP (mmHg): Expected Log MAP (mmHg) = 1.84 + 0.00153 × BMI + 0.000602 × maternal age

Logit = 0.421 + 2.491 × Log *a-priori* risk + 11.884 × Log MoM MAP; $R^2 = 12.4\%$

A-posteriori second-trimester risk

Calculation of AC MoM: Expected Log AC = 2.006 − 0.0058 × GA (weeks) + 0.00074 × GA^2 + 0.0061 if male

Calculation of UtA-PI MoM: Expected Log UtA-PI = 1.946 − 0.176 × GA + 0.0039 × GA^2 + 0.0125 if male

Logit = 0.199 + 2.796 × Log *a-posteriori* first-trimester risk − 31.614 × Log MoM AC + 3.456 × Log MoM UtA-PI; $R^2 = 22.4\%$

A-posteriori third-trimester risk

Calculation of EFW MoM: Expected Log EFW = −1.891 + 0.28 × GA − 0.00368 × GA^2 + 0.00604 if male

Logit = −1.591 + 1.841 × Log *a-posteriori* second-trimester risk − 29.1 × Log MoM EFW; R^2 47.5%

Table 3 and Figure 1 present the predictive performance of each derived model. In the first trimester, a combination of *a-priori risk* and MAP (i.e. *a-posteriori* first-trimester risk) yielded an AUC of 0.71 (95% CI, 0.65–0.77) (DR, 36.6% for a 10% FPR). In the second trimester, *a-posteriori* first-trimester risk combined with second-trimester AC and UtA Doppler (i.e. *a-posteriori* second-trimester risk) yielded an AUC of 0.81 (95% CI, 0.74–0.87) (DR, 43.1% for a 10% FPR). Finally, the combination of *a-posteriori* second-trimester risk plus third-trimester EFW (full model) yielded an AUC of 0.92 (95% CI, 0.88–0.96) (DR, 74% for a 10% FPR). All formulae for calculating these risks are given in Appendix S1.

Cut-offs for second-trimester risk, as the basis for contingent third-trimester US scanning of 10%, 25% and 50% of subjects, were 12.3%, 5% and 1.7%, respectively. In other words, third-trimester US was warranted at *a-posteriori* second-trimester risks above these cut-off points. Respective AUC (95% CI) values for contingent models involving 10%, 25% and 50% of subjects were 0.81 (0.75–0.88), 0.84 (0.78–0.91) and 0.89 (0.84–0.94), respectively. Only the 50% contingent model proved statistically equivalent to the full model ($P = 0.11$). Table 4 and Figure 2 show ROC curves for each contingent model and the full model.

Table 2 Biophysical and ultrasound variables of the validation population in first, second and third trimesters, according to fetal growth restriction (FGR) status

	No FGR (n = 1221)	*Late FGR (n = 82)*	P	*AUC (95% CI)*
First-trimester US				
GA at scan (weeks)	12.7 ± 0.63	12.8 ± 0.6	0.026	—
CRL (mm)	64.6 ± 9.8	64.8 ± 8.3	0.89	—
NT (mm)	1.5 ± 0.45	1.44 ± 0.38	0.19	0.51 (0.46–0.56)
MAP (mmHg)	79.2 ± 6.3	82.6 ± 8.7	< 0.001	—
MAP (MoM)	1 ± 0.08	1.035 ± 0.1	< 0.001	0.59 (0.52–0.66)
Second-trimester US				
GA at scan (weeks)	20.9 ± 0.79	21 ± 0.72	0.49	—
BPD (mm)	49.6 ± 3.2	48.7 ± 2.9	0.016	—
BPD (MoM)	1 ± 0.05	0.99 ± 0.04	0.007	0.59 (0.51–0.67)
HC (mm)	185.5 ± 11.1	180.9 ± 14.8	< 0.001	—
HC (MoM)	1 ± 0.04	0.98 ± 0.07	< 0.001	0.68 (0.60–0.76)
AC (mm)	163.5 ± 11.7	158 ± 11.1	< 0.001	—
AC (MoM)	1 ± 0.05	0.97 ± 0.05	< 0.001	0.71 (0.64–0.79)
FL (mm)	34.6 ± 2.8	33.7 ± 2.5	0.007	—
FL (MoM)	1 ± 0.06	0.97 ± 0.05	< 0.001	0.70 (0.63–0.78)
Mean UtA-PI	0.99 ± 0.29	1.26 ± 0.47	< 0.001	—
Mean UtA (MoM)	1.03 ± 0.3	1.32 ± 0.49	< 0.001	0.69 (0.60–0.78)
Third-trimester US				
GA at scan (weeks)	32.9 ± 0.67	32.7 ± 0.86	0.019	—
BPD (mm)	82.6 ± 3.7	78.9 ± 4.1	< 0.001	—
BPD (MoM)	1 ± 0.04	0.96 ± 0.04	< 0.001	0.74 (0.68–0.80)
HC (mm)	303 ± 13.2	290.1 ± 12.5	< 0.001	—
HC (MoM)	1 ± 0.04	0.97 ± 0.03	< 0.001	0.78 (0.73–0.83)
AC (mm)	290.9 ± 15	267.8 ± 16.6	< 0.001	—
AC (MoM)	1 ± 0.05	0.93 ± 0.04	< 0.001	0.88 (0.84–0.92)
FL (mm)	62.9 ± 2.8	60.1 ± 3.9	< 0.001	—
FL (MoM)	1 ± 0.04	0.96 ± 0.05	< 0.001	0.75 (0.69–0.82)
EFW (g)	2123.9 ± 256	1751 ± 266	< 0.001	—
EFW (MoM)	1 ± 0.1	0.84 ± 0.09	< 0.001	0.90 (0.86–0.93)

Data expressed as mean ± SD. AC, abdominal circumference; AUC, area under the receiver–operating characteristics curve; BPD, biparietal diameter; CRL, crown–rump length; EFW, estimated fetal weight; FL, femur length; GA, gestational age; HC, head circumference; MAP, mean arterial pressure; MoM, multiples of the median; NT, nuchal translucency; PI, pulsatility index; US, ultrasound examination; UtA, uterine artery.

Table 3 Detection rates for late fetal growth restriction, combining first-, second- and third-trimester models

		5% FPR			*10% FPR*			*20% FPR*		
	AUC (95% CI)	*DR (%)*	*LR+*	*LR–*	*DR (%)*	*LR+*	*LR–*	*DR (%)*	*LR+*	*LR–*
A-priori risk	0.69 (0.63–0.75)	21	4.2	0.83	30.5	3.05	0.77	40.2	2.01	0.75
A-posteriori risk										
First trimester	0.71 (0.65–0.77)	25.6	5.12	0.78	36.6	3.66	0.70	50	2.5	0.63
Second trimester	0.81 (0.74–0.87)	35.3	7.06	0.68	43.1	4.31	0.63	62.7	3.135	0.47
Third trimester (full model)	0.92 (0.88–0.96)	68	13.6	0.34	74	7.4	0.29	86	4.3	0.18

AUC, area under receiver–operating characteristics curve; DR, detection rate; FPR, false-positive rate; LR+/–, positive/negative likelihood ratio.

Table 4 Comparison of full model with various strategies of contingent third-trimester scanning (based on accrued risk at second trimester), selecting 10%, 25% or 50% of study population (popn)

			5% FPR			*10% FPR*			*20% FPR*		
	AUC (95% CI)	P *(against full model)*	*DR (%)*	*LR+*	*LR–*	*DR (%)*	*LR+*	*LR–*	*DR (%)*	*LR+*	*LR–*
Full model	0.92 (0.88–0.96)	—	68	13.6	0.34	74	7.4	0.29	86	4.3	0.18
Strategy											
50% of popn	0.89 (0.84–0.94)	0.11	68	13.6	0.34	72	7.2	0.31	80	4	0.25
25% of popn	0.84 (0.78–0.91)	< 0.001	57.1	11.42	0.45	59.2	5.92	0.45	73.5	3.675	0.33
10% of popn	0.81 (0.75–0.88)	< 0.001	40.8	8.16	0.62	46.9	4.69	0.59	63.3	3.165	0.46

AUC, area under receiver–operating characteristics curve; DR, detection rate; FPR, false-positive rate; LR+/–, positive/negative likelihood ratio.

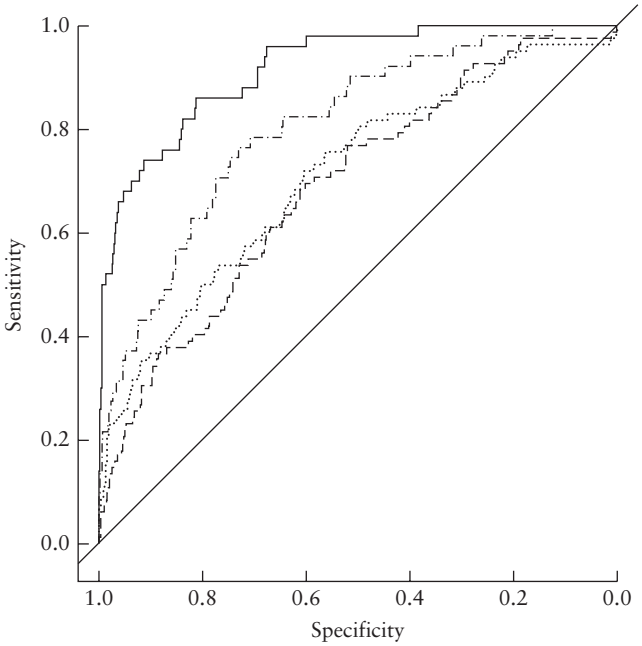

Figure 1 Performance (receiver–operating characteristics curves) of *a-priori* (- - - -) and first-trimester (·····), second-trimester (–·–·) and third-trimester (full model) (——) models for prediction of late fetal growth restriction.

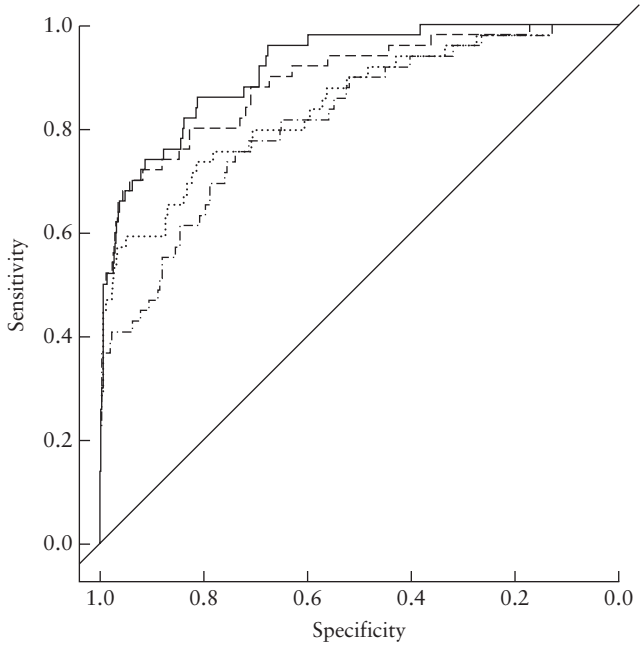

Figure 2 Performance (receiver–operating characteristics curves) of different strategies of contingent third-trimester ultrasound examination in prediction of late fetal growth restriction: full model (100% of population scanned) (——); 50% of population scanned (- - - -); 25% of population scanned (·····) and 10% of population scanned (–·–·).

DISCUSSION

In predicting late FGR, our study shows that a strategy of third-trimester US in 50% of the population, based on the combined first- and second-trimester risks, is equivalent to scanning routinely the whole population of pregnant women.

Although evidence from randomized trials has failed to demonstrate any real benefit from routine third-trimester scanning[7,12], it may be argued that the results of the meta-analysis[12] have limited contemporary validity as it included studies using outdated surrogates of fetal growth, such as BPD[28], or protocols in which diagnosis of FGR elicited no change in management. Subsequently, a trial to investigate the impact of third-trimester US should find willing participants[29], but a large sample size would be mandatory to assess the effects on hard outcomes, such as perinatal mortality related to birth-weight percentile outside the normal range.

The effectiveness of using third-trimester US biometry routinely in the diagnosis of FGR is also unclear. The DR of AC for a birth weight $< 10^{th}$ percentile ranges from 48% to 87%, with a specificity of 69–85%[30]. For EFW, DRs of 25–100% and specificities of 69–97% have been reported[14,30]. A recent large study shows that improvements in the detection of SGA can be obtained combining maternal characteristics and fetal biometry at 30–34 weeks, achieving DRs of 80% for neonates delivering within 5 weeks after assessment[31]. Recently, Lesmes *et al.*[32] demonstrated the potential value of UtA-PI at 19–24 weeks' gestation, in combination with maternal characteristics, medical history and fetal biometry, for prediction of delivery of SGA babies, achieving DRs of 66% and 43% for those delivering at 32–36 and ≥ 37 weeks' gestation, respectively. By refining our definition of FGR to include only those pregnancies with evidence of placental insufficiency before delivery or those that were very small, we tried to minimize the contamination of our cohort with constitutionally small babies, who represent the end of the spectrum of the normal population and do not in themselves represent a predictive target.

Souka *et al.*[11] evaluated a contingent strategy for SGA, rescanning 50% of an unselected cohort ($n = 2310$) according to first-trimester risk. DRs were 50–60%, for a 5–10% FPR. Our higher rate of detection is attributable to our incorporation of second-trimester UtA Doppler, which is known to be the single best predictor of FGR[12].

With the goal of improving healthcare systems, economic evaluation of intervention as compared with hard-to-justify scans is warranted. Very few studies have addressed the cost-effectiveness of US. Estimations of cost from the payer's perspective for US in the USA included 200 $ in 1998[33] and 43–74 $ (year unreported) for Medicaid reimbursement[34]. In Europe, a policy of routine third-trimester US is in place in most countries[35,36], resulting in an exceedingly large number of scans. At an estimated cost of 29–39 €[37] per US, this policy can generate annual costs of over 95–125 million €[35,36]. Furthermore, when scans are performed by physicians, rather than sonographers or technicians, the costs are likely to be much higher, up to 20% more[38]. Our study highlights the potential benefit of research into new strategies for improving the economic balance in health costs. We found that a contingent strategy

could lead to a 50% reduction in the number of third-trimester US scans, with no reduction in overall predictive capacity for late FGR. Women most likely to benefit from third-trimester US can be identified and managed, according to protocols integrating current evidence that classifies stages of deterioration in FGR, in order to establish both appropriate follow-up intervals and optimal timing of delivery[22].

Our study had some limitations. First, we had no information on Doppler findings in the third trimester. There is currently no conclusive evidence that the use of routine umbilical artery Doppler, or a combination of umbilical artery and UtA Doppler[39] or a combination of mean arterial pressure and UtA Doppler benefits either mother or baby in low-risk or unselected populations[40]. Recent evidence from observational studies suggests that UtA Doppler in the third trimester, as a proxy for trophoblastic invasion, provides information additional to biometry in the detection of FGR[41,42]. It might be argued that UtA Doppler earlier in pregnancy could identify even earlier these cases of defective trophoblastic invasion; however, it is interesting to note that in a recently reported longitudinal series[43], approximately one-third of abnormal third-trimester UtA Doppler studies occurred in women with normal scans during the second trimester, suggesting that, in a proportion of cases, placental compromise occurs late in pregnancy. Second, as opposed to scanning routinely before 34 weeks (as was done in this study), it could be argued that, because lagging growth is accentuated as term approaches, routine third-trimester US studies nearer term may improve detection of severe FGR. Indeed, a recent study[44] showed that US at 35–37 weeks predicted 89% of SGA neonates (birth weight < 5th percentile) delivering within 2 weeks following assessment. Third, the addition of cerebral Doppler in term pregnancies was suggested recently to improve the detection of placental insufficiency[45]. Our findings of a similar performance of risk-based screening and routine third-trimester US could not be translated into other strategies of scanning late in the third trimester.

In conclusion, a policy of performing third-trimester US in 50% of the population based on combined first- and second-trimester risks achieves a diagnostic performance for late FGR that is equivalent to that resulting from third-trimester US performed routinely in the entire population of pregnant women.

ACKNOWLEDGMENTS

This work was supported by grants from The Cerebra Foundation for the Brain Injured Child (Carmarthen, Wales, UK), Obra Social 'la Caixa' and AGAUR 2014 SGR grant n° 928. This publication was funded with support from the Erasmus + Programme of the European Union (Framework Agreement number: 2013–0040). This publication reflects the views only of the author and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

1. Hepburn M, Rosenberg K. An audit of the detection and management of small-for-gestational age babies. *Br J Obstet Gynaecol* 1986; **93**: 212–216.
2. Lindqvist PG, Molin J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome? *Ultrasound Obstet Gynecol* 2005; **25**: 258–264.
3. Gardosi J, Madurasinghe V, Williams M, Malik A, Francis A. Maternal and fetal risk factors for stillbirth: population based study. *BMJ* 2013; **346**: f108.
4. McCowan LM, Roberts CT, Dekker GA, Taylor RS, Chan EH, Kenny LC, Baker PN, Moss-Morris R, Chappell LC, North RA; SCOPE consortium. Risk factors for small-for-gestational-age infants by customised birthweight centiles: data from an international prospective cohort study. *BJOG* 2010; **117**: 1599–1607.
5. Crovetto F, Crispi F, Scazzocchio E, Mercade I, Meler E, Figueras F, Gratacos E. First-trimester screening for early and late small-for-gestational-age neonates using maternal serum biochemistry, blood pressure and uterine artery Doppler. *Ultrasound Obstet Gynecol* 2014; **43**: 34–40.
6. Papageorghiou AT, Yu CK, Bindra R, Pandis G, Nicolaides KH. Multicenter screening for pre-eclampsia and fetal growth restriction by transvaginal uterine artery Doppler at 23 weeks of gestation. *Ultrasound Obstet Gynecol* 2001; **18**: 441–449.
7. Martin AM, Bindra R, Curcio P, Cicero S, Nicolaides KH. Screening for pre-eclampsia and fetal growth restriction by uterine artery Doppler at 11–14 weeks of gestation. *Ultrasound Obstet Gynecol* 2001; **18**: 583–586.
8. Cnossen JS, Morris RK, ter Riet G, Mol BW, van der Post JA, Coomarasamy A, Zwinderman AH, Robson SC, Bindels PJ, Kleijnen J, Khan KS. Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. *CMAJ* 2008; **178**: 701–711.
9. Karagiannis G, Akolekar R, Sarquis R, Wright D, Nicolaides KH. Prediction of small-for-gestation neonates from biophysical and biochemical markers at 11–13 weeks. *Fetal Diagn Ther* 2011; **29**: 148–154.
10. Clausson B, Gardosi J, Francis A, Cnattingius S. Perinatal outcome in SGA births defined by customised versus population-based birthweight standards. *BJOG* 2001; **108**: 830–834.
11. Souka AP, Papastefanou I, Pilalis A, Michalitsi V, Kassanos D. Performance of third-trimester ultrasound for prediction of small-for-gestational-age neonates and evaluation of contingency screening policies. *Ultrasound Obstet Gynecol* 2012; **39**: 535–542.
12. Bricker L, Neilson JP, Dowswell T. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *Cochrane Database Syst Rev* 2008: CD001451.
13. Kean L, Liu D. Antenatal care as a screening tool for the detection of small for gestational age babies in the low risk population. *J Obstet Gynaecol* 1996; **16**: 77–82.
14. Hedriana HL, Moore TR. A comparison of single versus multiple growth ultrasonographic examinations in predicting birth weight. *Am J Obstet Gynecol* 1994; **170**: 1600–1604.
15. McKenna D, Tharmaratnam S, Mahsud S, Bailie C, Harper A, Dornan J. A randomized trial using ultrasound to identify the high-risk fetus in a low-risk population. *Obstet Gynecol* 2003; **101**: 626–632.
16. Robinson HP, Sweet EM, Adam AH. The accuracy of radiological estimates of gestational age using early fetal crown–rump length measurements by ultrasound as a basis for comparison. *Br J Obstet Gynaecol* 1979; **86**: 525–528.
17. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements-a prospective study. *Am J Obstet Gynecol* 1985; **151**: 333–337.
18. Figueras F, Meler E, Iraola A, Eixarch E, Coll O, Figueras J, Francis A, Gratacos E, Gardosi J. Customized birthweight standards for a Spanish population. *Eur J Obstet Gynecol Reprod Biol* 2008; **136**: 20–24.
19. Jin H, Lu Y. A non-inferiority test of areas under two parametric ROC curves. *Contemp Clin Trials* 2009; **30**: 375–379.
20. Baschat AA, Gembruch U. The cerebroplacental Doppler ratio revisited. *Ultrasound Obstet Gynecol* 2003; **21**: 124–127.
21. Gomez O, Figueras F, Fernandez S, Bennasar M, Martinez JM, Puerto B, Gratacos E. Reference ranges for uterine artery mean pulsatility index at 11–41 weeks of gestation. *Ultrasound Obstet Gynecol* 2008; **32**: 128–132.
22. Figueras F, Gratacós E. Update on the diagnosis and classification of fetal growth restriction and proposal of a stage-based management protocol. *Fetal Diagn Ther* 2014; **36**: 86–98.
23. Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy* 2001; **20**: IX–XIV.
24. National Institute for Health and Clinical Excellence (NICE). Intrapartum care: care of healthy women and their babies during childbirth. NICE Clinical Guideline *55*. 2012.
25. Gregg AR, Weiner CP. "Normal" umbilical arterial and venous acid–base and blood gas values. *Clin Obstet Gynecol* 1993; **36**: 24–32.
26. Figueras F, Torrents M, Munoz A, Comas C, Antolin E, Echevarria M, Mallafré J, Carrera JM. References intervals for fetal biometrical parameters. *Eur J Obstet Gynecol Reprod Biol* 2002; **105**: 25–30.
27. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–845.
28. Bakketeig LS, Eik-Nes SH, Jacobsen G, Ulstein MK, Brodtkorb CJ, Balstad P, Eriksen BC, Jörgensen NP. Randomised controlled trial of ultrasonographic screening in pregnancy. *Lancet* 1984; **2**: 207–211.
29. Le Ray C, Morin L. Routine versus indicated third trimester ultrasound: is a randomized trial feasible? *J Obstet Gynaecol Can* 2009; **31**: 113–119.

30. Figueras F, Gardosi J. Intrauterine growth restriction: new concepts in antenatal surveillance, diagnosis, and management. *Am J Obstet Gynecol* 2011; **204**: 288–300.
31. Bakalis S, Silva M, Akolekar R, Poon LC, Nicolaides KH. Prediction of small-for-gestational age neonates: screening by fetal biometry at 30–34 weeks. *Ultrasound Obstet Gynecol* 2015; **45**: 551–558.
32. Lesmes C, Gallo DM, Saiid Y, Poon LC, Nicolaides KH. Prediction of small-for-gestational age neonates: screening by uterine artery Doppler and mean arterial pressure at 19–24 weeks. *Ultrasound Obstet Gynecol* 2015; **46**: 332–340.
33. Vintzileos AM, Ananth CV, Smulian JC, Beazoglou T, Knuppel RA. Routine second-trimester ultrasonography in the United States: a cost-benefit analysis. *Am J Obstet Gynecol* 2000; **182**: 655–660.
34. Cahill AG, Odibo AO, Caughey AB, Stamilio DM, Hassan SS, Macones GA, Romero R. Universal cervical length screening and treatment with vaginal progesterone to prevent preterm birth: a decision and economic analysis. *Am J Obstet Gynecol* 2010; **202**: 548.e1–8.
35. Prenatal Screening Policies in Europe: EUROCAT. 2010.
36. Boyd PA, Devigan C, Khoshnood B, Loane M, Garne E, Dolk H. Survey of prenatal screening policies in Europe for structural malformations and chromosome anomalies, and their impact on detection and termination rates for neural tube defects and Down's syndrome. *BJOG* 2008; **115**: 689–696.
37. Henderson J, Bricker L, Roberts T, Mugford M, Garcia J, Neilson J. British National Health Service's and women's costs of antenatal ultrasound screening and follow-up tests. *Ultrasound Obstet Gynecol* 2002; **20**: 154–162.
38. Society of Diagnostic Medical Sonography. https://www.sdms.org/docs/default-source/Resources/2014-SDMS-Annual-Report.pdf?sfvrsn=2.
39. Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. *Cochrane Database Syst Rev* 2013; **11**: CD007529.
40. Bakalis S. Prediction of small-for-gestational-age neonates: screening by uterine artery Doppler and mean arterial pressure at 30–34 weeks. *Ultrasound Obstet Gynecol* 2015; **45**: 707–714.
41. Di Lorenzo G, Monasta L, Ceccarello M, Cecotti V, D'Ottavio G. Third trimester abdominal circumference, estimated fetal weight and uterine artery Doppler for the identification of newborns small and large for gestational age. *Eur J Obstet Gynecol Reprod Biol* 2013; **166**: 133–138.
42. Maroni E, Youssef A, Arcangeli T, Nanni M, De Musso F, Contro E, Kuleva M, Bellussi F, Pilu G, Rizzo N, Ghi T. Increased uterine artery pulsatility index at 34 weeks and outcome of pregnancy. *Ultrasound Obstet Gynecol* 2011; **38**: 395–399.
43. Llurba E, Turan O, Kasdaglis T, Harman CR, Baschat AA. Emergence of late-onset placental dysfunction: relationship to the change in uterine artery blood flow resistance between the first and third trimesters. *Am J Perinatol* 2013; **30**: 505–512.
44. Fadigas C, Saiid Y, Gonzalez R, Poon LC, Nicolaides KH. Prediction of small-for-gestational-age neonates: screening by fetal biometry at 35–37 weeks. *Ultrasound Obstet Gynecol* 2015; **45**: 559–565.
45. Morales-Rosello J, Khalil A, Morlando M, Papageorghiou A, Bhide A, Thilaganathan B. Changes in fetal Doppler indices as a marker of failure to reach growth potential at term. *Ultrasound Obstet Gynecol* 2014; **43**: 303–310.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Table S1 Epidemiological characteristics and perinatal outcome of the sub-cohorts (construction and validation)

Appendix S1 Formulae for MoM and for risk calculations

