




## RESEARCH ARTICLE OPEN ACCESS

# The Performance of Selective Screening Ultrasound to Detect the Small-for-Gestational-Age Foetus: A Prospective Cohort Study Nested Within the DESiGN Randomised Control Trial

Charles Arcus<sup>1</sup>  | James Elhindi<sup>1</sup> | Chivon Winsloe<sup>2</sup>  | Andrew Copas<sup>3</sup> | Sophie Relph<sup>4</sup> | Annette Briley<sup>2,5</sup> | Kirstie Coxon<sup>6</sup> | Mark Johnson<sup>7</sup> | Asma Khalil<sup>8,9</sup> | Deborah A. Lawlor<sup>10,11</sup> | Christoph Lees<sup>12,13</sup> | Neil Marlow<sup>9</sup>  | Louise Page<sup>14</sup> | Andrew Shennan<sup>2</sup> | Jane Sandall<sup>2</sup> | Tanya Nippita<sup>1,15</sup> | Matias C. Vieira<sup>2,16</sup> | Dharmintra Pasupathy<sup>1,2</sup>

<sup>1</sup>Reproduction and Perinatal Centre, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia | <sup>2</sup>Department of Women and Children's Health, School of Life Course and Population Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK | <sup>3</sup>Centre for Pragmatic Global Health Trials, Institute for Global Health, University College London, London, UK | <sup>4</sup>Department of Obstetrics and Gynaecology, Royal London Hospital, Barts Health NHS Trust, London, UK | <sup>5</sup>Caring Futures Institute Flinders University and North Adelaide Local Health Network, Adelaide, Australia | <sup>6</sup>Faculty of Health, Social Care and Education, Kingston University and St. George's, University of London, London, UK | <sup>7</sup>Centre for Implementation Science and King's Health Economics, King's College London, London, UK | <sup>8</sup>UCL Institute for Women's Health, University College London, London, UK | <sup>9</sup>Fetal Medicine Unit, St George's University Hospitals NHS Foundation Trust, London, UK | <sup>10</sup>Medical Research Council Integrative Epidemiology Unit at, The University of Bristol, Bristol, UK | <sup>11</sup>Population Health Science, University of Bristol, Bristol, UK | <sup>12</sup>Department of Fetal Medicine Imperial College Healthcare NHS Trust, London, UK | <sup>13</sup>Centre for Fetal Care, Queen Charlotte's and Chelsea Hospital, London, UK | <sup>14</sup>West Middlesex University Hospital, Chelsea & Westminster Hospital NHS Foundation Trust, Isleworth, UK | <sup>15</sup>Department of Obstetrics, Royal North Shore Hospital, St Leonards, New South Wales, Australia | <sup>16</sup>Department of Obstetrics and Gynaecology, University of Campinas (UNICAMP), School of Medical Sciences, Campinas, São Paulo, Brazil

**Correspondence:** Dharmintra Pasupathy ([dharmintra.pasupathy@sydney.edu.au](mailto:dharmintra.pasupathy@sydney.edu.au))

**Received:** 10 June 2025 | **Revised:** 13 January 2026 | **Accepted:** 23 February 2026

**Keywords:** foetal growth ultrasound | selective screening performance | small-for-gestational-age

## ABSTRACT

**Objective:** Evidence on the performance of foetal growth ultrasound parameters to detect the Small-for-Gestational-Age (SGA) foetus for a selective cohort is limited. We report the performance of estimated foetal weight under the 10th percentile (EFW < 10th) at the last selective screening ultrasound to detect SGA based on gestational age of scan and scan to birth interval.

**Design:** We performed a prospective cohort study to analyse ultrasound data obtained from the DESiGN trial.


**Setting:** The DESiGN trial was the first cluster randomised controlled trial to evaluate the impact of the Growth Assessment Protocol (GAP) on antenatal detection of SGA.

**Population:** Ninety-four thousand, nine hundred twenty-five women included in DESiGN who had at least 1 ultrasound scan after 24 weeks' gestational age with complete ultrasound data underwent analysis.

**Methods:** Test performance characteristics of EFW (by Hadlock 1991) < 10th percentile to predict birthweight < 10th percentile (by 1990 UK Population Charts) were reported at different gestational ages of scan categories and scan to birth intervals.

**Results:** The performance of EFW < 10th percentile at the last selective screening ultrasound to correctly identify SGA declined from 34 weeks and with increasing scan to birth intervals. A scan between 34 and 37 + 6 weeks' gestation more than 4 weeks prior to birth had a particularly poor performance (sensitivity 4.0%, 95% CI 3.8–4.1).

Matias C. Vieira joint senior author.

This article has a Video Abstract presented by Arcus, Dr Charles .

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *BJOG: An International Journal of Obstetrics and Gynaecology* published by John Wiley & Sons Ltd.

**Conclusions:** In a selected population, the performance of late third-trimester ultrasound estimation of foetal size to detect SGA at birth is poor.

## 1 | Introduction

There are two current approaches to antenatal screening for the Small-For-Gestational-Age (SGA) foetus with the primary aim of stillbirth reduction [1–7]. Selective screening is based on antenatal risk-stratification which justifies serial ultrasound estimation of foetal size [8, 9]. Universal screening involves routine third-trimester ultrasound independent of risk-status. In both strategies, additional screening scans are performed if clinical concerns develop (e.g., small fundal height, reduced foetal movement or a new diagnosis of preeclampsia). SGA detection increases with routine third-trimester ultrasound for unselected women but an increase in false positive results has been described [10]. Recent evidence has shown that this cohort remains at higher risk of stillbirth than those correctly predicted to be appropriately grown [3]. Nevertheless, action upon false positive results can lead to unnecessary intervention and potential harm [11–13]. The clinical and economic value of routine third-trimester ultrasound is contentious and access limited across some settings [14]. Consequently, most guidelines recommend selective screening for SGA at present [15–20].

Foetal growth assessment and risk of stillbirth for the SGA foetus is largely reliant upon the predictive accuracy of biometric parameters that are measured at serial growth ultrasound. Factors that impact screening performance such as gestational age at scan and scan to birth interval must be considered when ultrasound estimation of foetal size is used to inform clinical management and prompt closer surveillance. Understanding the performance of estimated foetal weight under the 10th percentile (EFW < 10th) at the last selective screening ultrasound to predict SGA helps reduce the risk of perinatal mortality without added neonatal morbidity associated with false positive results and the timing of iatrogenic birth [2–4].

Evidence on the performance of EFW < 10th to detect SGA is primarily confined to studies that implement universal screening programmes [21]. Data on the performance of EFW < 10th to predict SGA from a large multicentre selective screening programme is lacking. The aim of this study is to determine the performance of EFW < 10th at the last selective screening ultrasound to detect SGA across gestational age of scan and scan to birth interval. We will examine and interpret performance trends to inform clinical management particularly in birth suite and define where research should focus to improve current methods of SGA detection when applied to a selective screening programme.

## 2 | Methods

### 2.1 | Study Design

This was a prospective cohort study which adhered to STROBE guidelines. Analyses were performed using clinical and ultrasound data obtained during the DESiGN trial which was the first cluster randomised control trial to evaluate the impact of

the Growth Assessment Protocol (GAP) on antenatal detection of SGA [22]. The DESiGN trial recruited women from 13 maternity units in England and all women who gave birth in maternity units during the trial period between November 2016 and February 2019 were included. Women with multiple pregnancy, antenatally detected major foetal anomalies and births before 24+1 weeks' gestation were excluded. Ethical approval for this trial was obtained from the Health Research Authority (HRA) through the London Bloomsbury Research Ethics Committee (Ref. 15/LO/1632) and the Confidentiality Advisory Group (Ref. 15/CAG/0195).

### 2.2 | Outcomes

The outcome reported was SGA at birth as defined by birthweight under the 10th percentile by UK population standards [23] to remain consistent with the DESiGN trial and the location of its study sites. We reported the outcome at different gestational age of scan categories and within different scan to birth intervals within 2, 3, 4 and beyond 4 weeks.

Clinical comparisons were performed between those who had at least 1 selective screening ultrasound from those that did not to differentiate the population that underwent statistical analysis. Characteristics that were observed and reported included: age at estimated conception, ethnicity (White, Black, Asian, Mixed, Other), Index of multiple deprivation (IMD Class 1–5), body mass index (BMI < 18, 18–25, 25–30, 30–35, 35–40, > 40), parity (0, 1+), smoking in pregnancy, pre-existing comorbidities (hypertension, diabetes mellitus) and antenatal complications (gestational diabetes mellitus, pre-eclampsia, eclampsia, gestational hypertension). Maternal outcomes of these two groups that were reported included onset of labour (spontaneous, induction, prelabour caesarean) and mode of birth (unassisted vaginal, instrumental birth, vaginal breech, emergency caesarean section). Perinatal outcomes that were reported included: gestational age at birth, preterm delivery, birthweight with corresponding percentiles defined by 1990 UK population charts [23], SGA at birth by population chart [23], Apgar scores < 7 at 5 min, admission to a neonatal unit, minor neonatal morbidity (which included hypothermia, hypoglycaemia or nasogastric feeding requirements), major neonatal morbidity (which included any neonatal brain injury such as intraventricular haemorrhage or hypoxic ischaemic encephalopathy, supplementary oxygen requirements beyond 28 days post-partum, necrotising enterocolitis, sepsis or retinopathy of prematurity) and stillbirth.

### 2.3 | Exposures

The performance of EFW < 10th percentile (as measured by the Hadlock 1991 formula using head circumference, femur length and abdominal circumference [24]) at the last selective screening ultrasound was reported to reflect current practice and a commonly accepted foetal growth standard. The last selective

screening ultrasound was defined in this study as the last ultrasound performed at or after 24 + 1 weeks of gestational age with complete foetal biometric data provided.

We also repeated the analysis when the 3rd percentile was used as the EFW threshold to suspect foetal growth restriction (FGR) and provided supplementary evidence on the performance of EFW < 10th percentile and abnormal umbilical artery Doppler parameters (defined as any of: pulsatility index > 95%, intermittent absent, absent or reversed end-diastolic flow) at the last selective screening ultrasound to predict SGA at birth.

## 2.4 | Statistical Analysis

Clinical characteristics and birth outcomes were reported using summary statistics and comparisons between groups with and without selective screening ultrasound were performed using Chi-square and Mann–Whitney tests as appropriate. Complete data analysis was performed using StataMP 17 on women in the DESiGN cohort who had at least 1 selective screening ultrasound after 24 weeks' gestational age with complete data for EFW, gestational age at birth and birthweight. Complete ultrasound data to enable test performance analysis was available from 12 of the 13 sites in the DESiGN trial which means that for the purpose of this analysis, only data from these sites were used. Multiple imputation and inverse probability weighting were not performed because the reason for missingness could not be deduced from the primary dataset. Funnel plots were also produced as Tables S6–S7 to inform the degree of heterogeneity in test performance between different sites.

The cohort was stratified into categories by different gestational age at last scan centred upon the mode for highest ultrasound activity which was observed by frequency histogram (Figure S1). Event rates (where event of SGA was defined by birthweight under the 10th percentile for population charts) were reported for each gestational age at last scan category. Area-under-the-curve (AUC) calculations were reported for each gestational age category to reflect the performance of EFW to predict SGA, and partial area-under-the-curves (pAUC) were reported for each gestational category to correspond to the predictive performance of the EFW < 10th percentile. Calculation of partial AUC was deliberately performed to restrict test performance analysis to the range of percentiles beneath a fixed threshold consistently used in clinical settings.

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were reported with 95% confidence intervals to reflect overall test performance of the EFW < 10th at the last scan to predict SGA at birth. These measures of test performance were reported by gestational age categories and scan to birth intervals as described above. All performance measures were reported to inform the precision and overall effectiveness of the screening test to delineate true positive, true negative, false positive and false negative cases. Positive and negative likelihood ratios were also reported as [Supporting Information](#) to provide an alternative measure of performance.

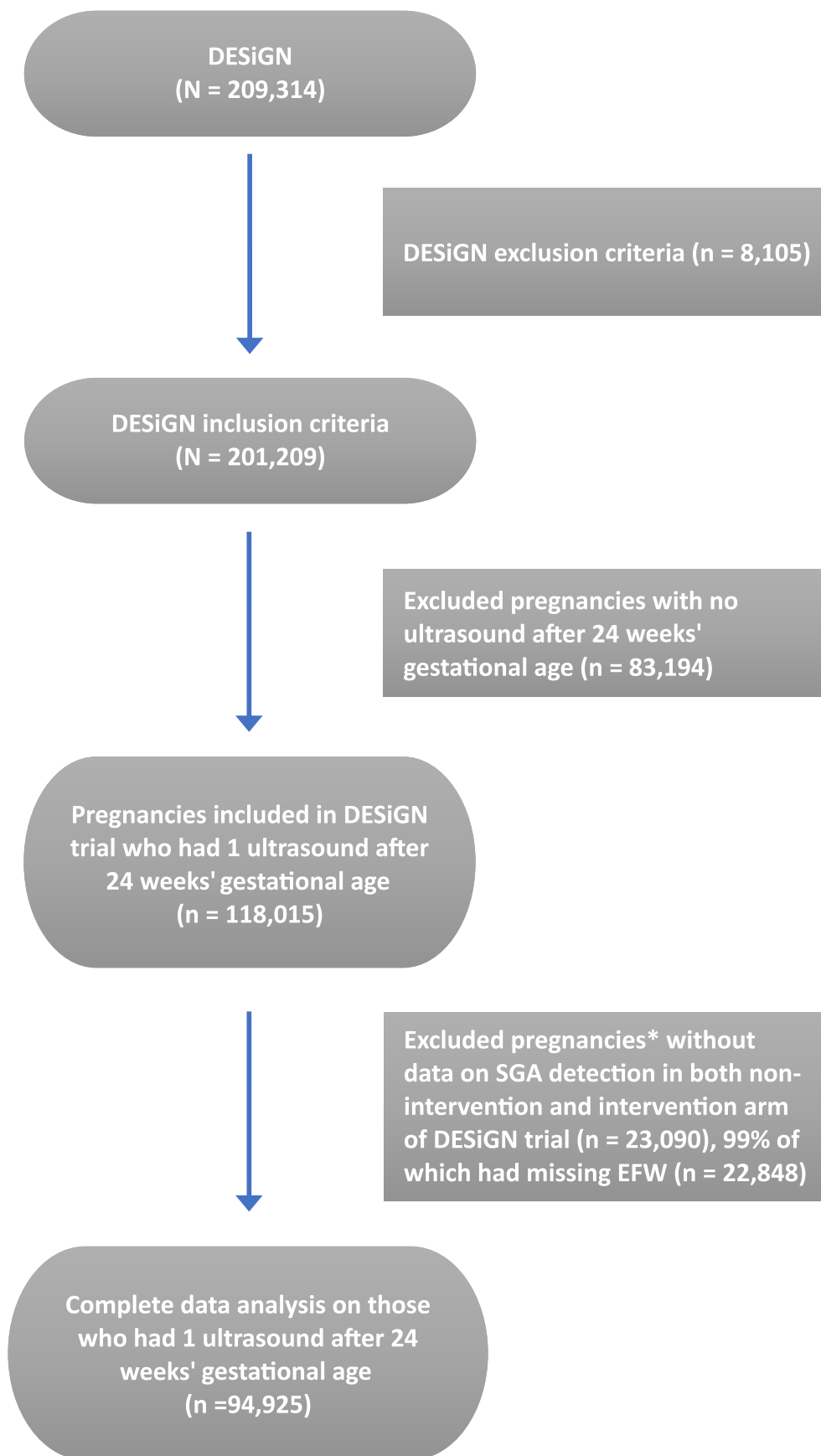
## 3 | Results

There were 201 209 pregnancies included in the DESiGN trial after exclusion of women with multiple pregnancy or antenatally detected foetal anomaly ( $n=8105$ ). Over half ( $n=118015$ ) of the women who met the DESiGN inclusion criteria had at least one ultrasound after 24 weeks' gestation. Complete data analysis was performed on 94 925 women who had an ultrasound after 24 weeks after those with missing data on SGA detection in both the intervention and non-intervention arm of the DESiGN trial were excluded ( $n=23090$ ) where missing data on SGA detection included EFW at last scan (99% of missing scan data) and gestational age at birth and birthweight (1% of missing scan data) (Figure 1).

Women who had a selective screening ultrasound were more likely older than 35 than those who did not have a scan, more likely from both least (Class 1 IMD 12.85 vs. 11.3%) and most deprived backgrounds (Class 5 IMD 21.9% vs. 20.2%), of white ethnicity (52.0% vs. 47.8%), to be multiparous (45.7% vs. 41.8%) and to smoke in pregnancy (5.5% vs. 4.3%). They were also more likely to be underweight (BMI < 18 kg/m<sup>2</sup> 2.7% vs. 1.5%) or overweight or with obesity (BMI 30–35 kg/m<sup>2</sup> 10.7% vs. 7.9%, BMI 35–40 kg/m<sup>2</sup> 4.5% vs. 1.7%, BMI > 40 kg/m<sup>2</sup> 2.1% vs. 0.5%), and have pre-existing (hypertension 1.4% v 0.8%, diabetes 1.8% vs. 0.8%), other antenatal (gestational diabetes 6.8% vs. 1.3%, hypertensive disorders of pregnancy 2.8% vs. 1.6%) comorbidities, to have labour induced (29.1% vs. 20.4%) or have caesarean birth (prelabour 19.3% vs. 11.4%, emergency 16.7% vs. 15.0%). Women who had a selective screening ultrasound gave birth to newborn infants with a lower birthweight (3315 g vs. 3400 g) and who were more commonly SGA (10.1% vs. 7.2%). Rates of preterm birth (6.4% vs. 4.9%), admission to a neonatal unit (13.3% vs. 9.2%) and neonatal morbidity (major morbidity 6.5% vs. 5.2%, minor morbidity 3.9% vs. 2.7%) were all higher amongst women who had a selective screening ultrasound. However, stillbirth was less common for women who had a selective screening ultrasound (0.3% vs. 1.0%) (Table 1).

Amongst women who had at least one selective screening ultrasound after 24 weeks, the prevalence of SGA was between 8.6% and 10.4%. However, for those who had their last scan between 38 and 41 + 6, it increased to 14.3% (Table S1). For each gestational age category, the rate of SGA increased with widening scan to birth intervals. Lower rates of birth within 4 weeks were observed when the scan was performed before 34 weeks. However, the rates of SGA amongst these births were higher compared to births from women who had their last scan after 34 weeks (Table S2).

The value of EFW at the last scan to predict SGA at birth as represented by AUC increased with gestational age of scan, but the value of the EFW < 10th threshold to predict SGA at birth as represented by partial AUC was comparable across gestational age (Table S1). Using EFW < 10th to correctly identify SGA, sensitivity largely decreased with increasing gestational age (Table 2). For most gestational age categories, a small reduction in sensitivities was observed across scan to birth intervals within 4 weeks. However, a greater decline in sensitivities across widening scan to birth intervals within 4 weeks was observed when the last scan was performed between 34 + 0 and 37 + 6 weeks when compared with



**FIGURE 1** | Population Flow Diagram. \*Exclusions are not mutually exclusive.

**TABLE 1** | A comparison of those who have a clinically indicated foetal growth scan with those who do not.

<b>Characteristic and Outcome</b>	<b>No Scans 83 194/201209 (41.3%)</b>	<b>1 or more indicated scans 118 015/201209 (58.7%)</b>	<b><i>p</i></b>
Maternal Age (Cont.)	31.4 (27.5–34.8)	32.2 (28.2–35.8)	< 0.01
Maternal Age (Cat.)			
< 20	1527 (2.2%)	2763 (2.5%)	< 0.01
20–24	8211 (11.8%)	11 401 (10.2%)	
25–34	43 536 (62.4%)	64 081 (57.2%)	
35–39	14 408 (20.6%)	26 351 (23.5%)	
< 40	2091 (3.0%)	7387 (6.6%)	
IMD			
Class 1 (least deprived)	9342 (11.4%)	15 071 (12.9%)	< 0.01
Class 2	10 452 (12.7%)	14 732 (12.6%)	
Class 3	18 295 (22.3%)	24 703 (20.6%)	
Class 4	27 303 (33.2%)	37 235 (31.8%)	
Class 5 (most deprived)	16 800 (20.4%)	25 807 (22.1%)	
Ethnicity			
White	39 745 (53.2%)	61 322 (56.1%)	< 0.01
Black	10 106 (13.5%)	15 010 (13.7%)	
Asian	15 592 (20.9%)	21 689 (10.8%)	
Mixed	1126 (1.5%)	2306 (2.1%)	
Other	8095 (10.8%)	9063 (8.3%)	
BMI (Cat.)			
1 (BMI < 18)	1217 (2.0%)	3221 (3.3%)	< 0.01
2 (18–25)	34 081 (54.9%)	46 347 (47.7%)	
3 (25–30)	18 335 (29.6%)	27 132 (28.0%)	
4 (30–35)	6583 (10.6%)	12 585 (13.0%)	
5 (35–40)	1406 (2.3%)	5266 (5.4%)	
6 (> 40)	417 (0.7%)	2534 (2.6%)	
Parity			
0	36 155 (51.0%)	51 575 (48.9%)	< 0.01
1+	34 746 (49.0%)	53 929 (51.1%)	
Smoking in pregnancy	3593 (4.3%)	6468 (5.5%)	< 0.01
Comorbidities			
Pre-existing HTN	688 (0.8%)	1597 (1.4%)	< 0.01
Pre-existing Diabetes	642 (0.8%)	2132 (1.8%)	
Gestational Diabetes	1042 (1.3%)	8044 (6.8%)	
HTN Dx of Pregnancy	1304 (1.6%)	3288 (2.8%)	
Onset of Labour			
Spontaneous	54 789 (67.4%)	58 321 (50.5%)	< 0.01

(Continues)

TABLE 1 | (Continued)

Characteristic and Outcome	No Scans 83 194/201 209 (41.3%)	1 or more indicated scans 118 015/201 209 (58.7%)	<i>p</i>
Induction	16 970 (20.9%)	34 365 (29.8%)	
Prelabour Caesarean	9 482 (11.7%)	22 796 (19.7%)	
Mode of Birth			
Unassisted Vaginal	50 134 (61.0%)	61 905 (53.1%)	< 0.01
Instrumental Birth	12 179 (14.8%)	16 489 (14.1%)	
Vaginal Breech	334 (0.4%)	251 (0.2%)	
Emergency Caesarean	12 450 (15.2%)	19 728 (15.6%)	
Neonatal Outcomes			
Gestational age at birth	40.0 (39.0–40.7)	39.6 (38.6–40.4)	< 0.01
Preterm birth	4 090 (4.9%)	7 507 (6.4%)	
BW (g)	3 400 (3 080–3 700)	3 315 (2 984–3 650)	
BW %	47.1 (26.2–69.6)	44.3 (22.7–68.7)	
SGA (BW < 10th %)	5 996 (7.7%)	11 926 (10.1%)	
APGAR < 7 at 5 mins	1 868 (2.3%)	1 938 (1.6%)	
Admission to NU	7 662 (9.2%)	15 736 (13.3%)	
Major NNM	4 297 (5.2%)	7 699 (6.5%)	
Minor NNM	2 225 (2.7%)	4 661 (4.0%)	
Stillbirth	801 (1.0%)	367 (0.3%)	

Note: Missing data (*missing variable, % of cohort without US vs % cohort with US*)—IMD 1.2% vs. 0.9%, ethnicity 10.3% vs. 7.3%, BMI 25.4% vs. 17.7%, parity 14.8% vs. 10.6%, onset of labour 2.3% vs. 2.1%, mode of birth 1.3% vs. 1.2%, SGA 6.5% vs. 0.3%, stillbirth 0.3% vs. 0.1%. Abbreviations: BW, birthweight (Cole et al. [23]); GDM, Gestational Diabetes; HTN Dx of Pregnancy, Hypertensive disorders of pregnancy; HTN, Hypertension; IMD, Indices of multiple deprivation; NNM, neonatal morbidity; NU, neonatal unit; SGA, small for gestational age infant (Cole et al. [23]).

other gestational age categories. At scan to birth intervals beyond 4 weeks, sensitivities for detecting SGA were much lower, at or below 22%, regardless of gestational age. When using EFW < 10th percentile to predict SGA at birth, specificity remained above 90% across all scan to birth intervals and gestational age categories (Table 2).

The positive predictive value of EFW < 10th percentile was between 55% and 74% across gestational age categories and scan to birth intervals within 4 weeks (Table 3). The negative predictive value of EFW < 10th minimally decreased with gestational age categories and widening scan to birth intervals within 4 weeks, but remained consistently above 90% (Table 3). When abnormal umbilical artery Doppler indices were added to EFW < 10th to predict SGA, sensitivity was poor and declined from 34 weeks' gestation across all scan to birth intervals, but specificity was stable and consistently above 98% (Table S3). When the 3rd percentile was used instead of the 10th percentile as the EFW threshold to antenatally suspect FGR, most patterns recurred but the absolute values differed. Notably, both PPV and sensitivity declined more sharply towards term (Table S4). For all exposures performed on the study cohort, selective screening ultrasound performed within 4 weeks of birth up until term gestation had strong rule-in capacity and moderate rule-out capacity for SGA (Table S5).

## 4 | Discussion

### 4.1 | Main Findings

We observed women from a selective screening programme who had ultrasound estimation of foetal size beyond 24 weeks were more likely associated with SGA and adverse outcomes than those who did not, consistent with high-risk status. The sensitivity and specificity of EFW < 10th to predict SGA declined when the last scan was performed closer to term. The sensitivity of EFW < 10th to predict SGA also deteriorated at increasing scan to birth intervals, particularly when ultrasound was performed between 34 and 37 + 6 and at scan to birth intervals beyond 4 weeks. The addition of abnormal umbilical Doppler indices maintained the rate of false positives but increased the rate of false negatives. The same performance limitations were observed when using EFW < 3rd to screen for FGR.

### 4.2 | Comparison to Previous Literature

The universal screening performance of routine third-trimester ultrasound to detect SGA improves with gestational age [21]. Our analysis relevant to the performance of selective screening

**TABLE 2** | Sensitivity (%) and specificity (%) of EFW < 10th at last scan to predict SGA across gestational age and scan to birth interval for those with ≥ 1 scan (n = 118 015).

GA at last scan	Sensitivity < 2 weeks	Specificity < 2 weeks	Sensitivity < 3 weeks	Specificity < 3 weeks	Sensitivity < 4 weeks	Specificity < 4 weeks	Sensitivity > 4 weeks	Specificity > 4 weeks
24 + 0–25 + 6	62.5 (59.7–65.3)	99.1 (96.2–100.0)	59.1 (56.2–62.0)	99.2 (96.6–100.0)	57.8 (54.9–60.7)	99.2 (96.6–100.0)	0*	96.9 (92.5–100.0)
26 + 0–29 + 6	75.0 (73.7–76.3)	98.9 (97.2–100.0)	74.0 (72.7–75.3)	98.9 (97.3–100.0)	71.2 (69.8–72.5)	99.0 (97.4–100.0)	2.5 (2.0–3.0)	96.4 (94.4–98.5)
30 + 6–33 + 6	75.8 (75.0–76.6)	98.6 (97.1–100.0)	70.1 (69.2–70.9)	98.7 (97.3–100.0)	65.0 (64.1–65.9)	98.7 (97.5–100.0)	4.8 (4.4–5.2)	97.1 (95.8–98.4)
34 + 0–37 + 6	54.1 (53.7–54.5)	98.4 (97.8–98.9)	41.6 (41.2–42.0)	98.5 (98.1–99.0)	32.7 (32.2–33.0)	98.6 (98.3–99.0)	4.0 (3.8–4.1)	96.5 (95.6–97.4)
38 + 0–41 + 6	37.7 (36.9–38.5)	97.1 (96.4–97.8)	35.4 (34.6–36.2)	97.2 (96.5–97.8)	34.8 (34.0–35.6)	97.2 (96.5–97.9)	22.2 (21.6–22.8)	92.7 (80.6–100.0)

Note: EFW references are derived from Hadlock et al. [24]. Birthweight references are derived from Cole et al. [23].  
\*0 = no cases.

**TABLE 3** | Positive (%) and Negative Predictive (%) Value of EFW < 10th at last scan to predict SGA across gestational age and scan to birth interval for those with ≥ 1 scan (n = 118 015).

GA at last scan	PPV < 2 weeks	NPV < 2 weeks	PPV < 3 weeks	NPV < 3 weeks	PPV < 4 weeks	NPV < 4 weeks	PPV > 4 weeks	NPV > 4 weeks
24 + 0–25 + 6	71.4 (68.8–74.1)	98.7 (98.0–99.3)	74.3 (71.7–76.8)	98.4 (97.7–99.1)	74.3 (71.7–76.8)	98.3 (97.6–99.1)	0*	94.6 (93.3–96.0)
26 + 0–29 + 6	68.9 (67.6–70.3)	99.2 (98.9–99.4)	70.8 (69.5–72.1)	99.1 (98.8–99.4)	72.1 (70.7–73.4)	98.9 (98.6–99.2)	5.0 (4.3–5.6)	93.0 (92.3–93.8)
30 + 6–33 + 6	54.6 (53.6–55.5)	99.5 (99.3–99.6)	56.9 (55.9–57.8)	99.3 (99.1–99.4)	58.9 (58.0–59.8)	99.0 (98.8–99.2)	9.4 (8.9–10.0)	94.2 (93.8–94.6)
34 + 0–37 + 6	55.5 (55.1–55.9)	98.3 (98.2–98.4)	60.9 (60.5–61.3)	96.9 (96.7–97.0)	63.6 (63.2–64.0)	95.2 (95.0–95.4)	2.7 (2.6–2.9)	97.6 (97.4–97.7)
38 + 0–41 + 6	65.9 (65.1–66.6)	91.4 (90.9–91.8)	66.9 (66.2–67.7)	90.3 (89.9–90.8)	67.2 (66.5–68.0)	90.0 (89.6–90.5)	0.3 (0.2–0.4)	99.9 (99.9–99.9)

Note: EFW references are derived from Hadlock et al. [24]. Birthweight references are derived from Cole et al. [23].  
\*0 = no cases.

ultrasound demonstrated the opposite trend. It can be shown from our data that the value of EFW as represented by AUC improved with gestational age but the value of EFW percentiles lower than the 10th threshold as represented by partial AUC did not. A previous retrospective analysis has found that when performed for a high-risk cohort, sonographic estimation of birthweight is more accurate prior to 37 weeks' gestation [25]. We have now verified these findings in a significantly larger cohort.

### 4.3 | Interpretation

Decline in test performance trends towards term for a selective cohort may simply reflect differences in the scanned populations such as differences in the prevalence of iatrogenic preterm birth for FGR or comorbid complications such as early preeclampsia. The prevalence of SGA at birth was highest amongst those who had their last scan after 38 weeks presumably because these women gave birth at later gestation with more time for slowing growth velocity. However, amongst births within 4 weeks of the last scan, rates of SGA were higher when the last scan was performed before 34 weeks. We speculate the greater severity of growth pathology at earlier gestational age contributes to the improved performance observed before 34 weeks.

### 4.4 | Clinical Implications

Understanding when screening ultrasound underperforms for a select cohort is clinically important. In particular, our analysis may encourage clinicians to remain cautious of the reliability of EFW < 10th percentile to accurately predict SGA when ultrasound estimation of foetal size is performed after 34 weeks. Moreover, it may also influence clinical suspicion of screen-negative results towards term and may be used to advise clinicians who counsel women based on findings of 'appropriate' growth to consider the duration for which such foetal size estimates remain valid. We observed that increasing scan to birth intervals beyond 4 weeks most negatively impacted test performance characteristics. This may have been because numbers were too small for reliable assessment under those conditions. Nevertheless, the weak correlation between EFW and birthweight at scan to birth intervals beyond 4 weeks challenges our capacity to rely on these projections and instigate appropriate management.

Data on the screening performance of foetal growth assessment is important but clinicians should understand the potential harm associated with action upon inaccurate results. Previous work from the DESIGN group has demonstrated that there are clinically relevant differences in perinatal complications between the SGA and non-SGA foetus that are based on screening outcomes. It was recently published that a foetus with a false negative result is at increased risk of stillbirth (aOR 1.2 95% CI 1.07–1.31), but decreased risk of major (aOR 0.83 95% CI 0.79–0.87) and minor neonatal morbidity (aOR 0.52 95% CI 0.50–0.54) when compared to a foetus with a true positive result [3]. The same paper demonstrated that a foetus with a false positive result is at increased risk of stillbirth (aOR 2.2 95% CI 1.8–2.6) and increased risk of major (aOR 1.0 95% CI 1.0–1.1)

and minor neonatal morbidity (aOR 1.7 95% CI 1.6–1.8) when compared to a foetus with a true negative result [3]. In our study, we observed that implementation of current antenatal screen positive ultrasound criteria for SGA at the last scan confers increased risk of false negative results, and hence stillbirth but not necessarily morbidity, for a select cohort towards term.

### 4.5 | Future Research

Primarily, our findings undermine confidence in ultrasound estimates of foetal size performed after 34 weeks' gestation, particularly for those who are born at term, and expose limitations of selective screening pathways to predict late-onset SGA. More frequent late preterm surveillance may improve detection but will likely have negative economic ramifications and may lead to unnecessary intervention. Future research should instead explore what individual and combined predictive accuracy additional sonographic parameters such as modified biometric thresholds or a measure of growth velocity have to sustain selective foetal growth screening performance.

### 4.6 | Strengths and Limitations

The strength of this study lies in evidence based on data obtained from a large multicentre selective screening programme in England which confers generalisability to a clinical context that conforms to RCOG screening recommendations [15]. The scale of this dataset also affords unique opportunity to meet our objective which was to primarily observe subtle trends in current selective screening performance that are shaped by the timing of ultrasound and its proximity to birth.

However, the study has some limitations. Firstly, the data from this study relevant to the last scan cannot be used to infer the selective screening performance of EFW < 10th to predict SGA at the next scan, or be used to inform the timing of the next scan. The value of our paper was not to inform antenatal surveillance and the timing of the next ultrasound but to instead aid decision making when planning birth and managing the intrapartum period.

Secondly, we observed greater heterogeneity in the sensitivity of EFW < 10th to predict SGA within 2 and 4 weeks of birth when the ultrasound was performed at earlier gestational age windows and speculate this is driven by fewer events (defined by scan and birth within the prespecified interval) related to site resources to manage preterm infants and the risk profile of its underlying serviced population.

Thirdly, we did not have access to data on the clinical indication for scans which may influence the strength of their estimation of foetal size. We speculate that natural cohort differences between scanned populations confer an observation bias on test performance trends that vary across gestational age of scan and scan to birth interval. We acknowledge that missing data on the indication for foetal growth ultrasound may introduce spectrum bias and limit generalisability of our findings to clinical samples selected by the same screening protocols. Nonetheless, we provide a pragmatic interpretation of the capacity of foetal growth

ultrasound to detect pathology as it is distributed asymmetrically across the course of pregnancy for a selective cohort.

Finally, our findings are not applicable to other foetal and neonatal birthweight classifications beyond those utilised in the DESiGN trial. This limits interpretability of our findings because biometric thresholds associated with adverse perinatal outcomes differ for separate charts as previous research has shown [26]. Use of Hadlock and UK population charts, in particular, also complicate interpretation of test performance of EFW < 10th to predict SGA because of potential unadjusted discordance between antenatal and postnatal standards, particularly towards term. Hadlock's formula is limited to a small homogenous population [24] and its accuracy diminishes in late gestation [27, 28], particularly with extremes of weight [27, 29]. Importantly, our study was bound to the charts used and percentiles calculated by the hospitals recruited in the DESiGN trial. Despite these limitations, the study remains relevant to a wide clinical setting because it measures the accuracy of a selective ultrasound screening programme to predict SGA when it conforms to foetal and neonatal growth standards that are commonly accepted by a large sample of maternity care providers in a public healthcare service accessed by most of the population [30].

## 5 | Conclusion

Understanding the temporal patterns of selective screening ultrasound performance across the course of pregnancy is clinically important. Without knowledge of where limitations are imposed on the selective screening performance of EFW < 10th, practitioners who look back at most recent size estimates cannot justify decisions aimed at reducing the risk of adverse perinatal outcomes associated with SGA. Our analysis reveals the need for research to improve the performance of current methods to predict late onset SGA when applied to a selective screening programme. Sonographic parameters to optimise detection of late-onset SGA may include lower biometric thresholds for EFW, Doppler velocimetry or a measure of foetal growth velocity.

### Author Contributions

D.P. was the Chief Investigator of the DESiGN trial and its secondary analyses. C.A., M.C.V., J.E. and D.P. conceived the analysis and were involved in its development, interpretation and review. C.A. and J.E. oversaw data management. C.A. drafted the manuscript. D.P., M.C.V., J.E., S.R., A.C. and C.W. were responsible for immediate and sustained critical manuscript reviews. All other authors were involved in subsequent edits.

### Acknowledgements

We wish to thank the other members of the DESiGN Trial Team: Andrew Healey, Alessandro Alanna, Donald Peebles and Baskaran Thilaganathan who contributed to the development and reporting of the DESiGN trial and its main secondary analyses. We would like to thank the members of the DESiGN Collaborative Group (site principal investigators, GAP clinical leads and clinicians or IT professionals) who assisted with data collection. DESiGN was funded by Guy's and St Thomas' Charity (MAJ150704), Tommy's Charity (the trial sponsor is the Tommy's London research centre) and SANDS (RG1011/16).

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australasian University Librarians

### Funding

DESiGN was funded by Guy's and St Thomas' Charity (MAJ150704), Tommy's Charity (the trial sponsor is the Tommy's London research centre) and SANDS (RG1011/16).

### Ethics Statement

Ethical approval for the DESiGN trial and its secondary analyses was obtained from the Health Research Authority (HRA) through the London Bloomsbury Research Ethics Committee (Ref. 15/LO/1632) and the Confidentiality Advisory Group (Ref. 15/CAG/0195).

### Conflicts of Interest

Jane Sandall is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust. The views expressed are those of the author[s] and not necessarily those of the NIHR or the Department of Health and Social Care.

### Data Availability Statement

Data availability is limited by third party restrictions (Confidentiality Advisory Group Ref 15/CAG/0195).

### References

1. V. M. Flenady, P. M. P. H. Middleton, G. C. P. Smith, et al., "Stillbirths: The Way Forward in High-Income Countries," *Lancet* 377, no. 9778 (2011): 1703–1717.
2. J. Gardosi, V. Madurasinghe, M. Williams, A. Malik, and A. Francis, "Maternal and Fetal Risk Factors for Stillbirth: Population Based Study," *BMJ (Online)* 346, no. 7893 (2013): 15.
3. C. Winsloe, J. Elhindi, D. Pasupathy, et al., "Perinatal Outcomes After Selective Third-Trimester Ultrasound Screening for Small-For-Gestational Age: Prospective Cohort Study Nested Within DESiGN Randomized Controlled Trial," *Ultrasound in Obstetrics & Gynecology* 65 (2024): 30–38.
4. P. G. Lindqvist and J. Molin, "Does Antenatal Identification of Small-For-Gestational Age Fetuses Significantly Improve Their Outcome?," *Ultrasound in Obstetrics & Gynecology* 25, no. 3 (2005): 258–264.
5. O. Hugh, M. Williams, S. Turner, and J. Gardosi, "Reduction of Stillbirths in England From 2008 to 2017 According to Uptake of the Growth Assessment Protocol: 10-Year Population-Based Cohort Study," *Ultrasound in Obstetrics & Gynecology* 57, no. 3 (2021): 401–408.
6. F. J. Cowan, C. J. D. McKinlay, R. S. Taylor, et al., "Detection of Small for Gestational Age Babies and Perinatal Outcomes Following Implementation of the Growth Assessment Protocol at a New Zealand Tertiary Facility: An Observational Intervention Study," *Australian and New Zealand Journal of Obstetrics and Gynaecology* 61, no. 3 (2021): 339–346.
7. J. Gardosi, S. Giddings, S. Clifford, L. Wood, and A. Francis, "Association Between Reduced Stillbirth Rates in England and Regional Uptake of Accreditation Training in Customised Fetal Growth Assessment," *BMJ Open* 3, no. 12 (2013): e003942.
8. F. Figueras and G. Eduard, "Stage-Based Approach to the Management of Fetal Growth Restriction: Update on FGR," *Prenatal Diagnosis* 34 (2014): 655–659.

9. S. Relph, M. C. Vieira, D. Pasupathy, et al., "Characteristics Associated With Antenatally Unidentified Small-For-Gestational-Age Fetuses: Prospective Cohort Study Nested Within DESiGN Randomized Controlled Trial," *Ultrasound in Obstetrics & Gynecology* 61, no. 3 (2023): 356–366.
10. U. Sovio, I. R. White, A. Dacey, D. Pasupathy, and G. C. S. Smith, "Screening for Fetal Growth Restriction With Universal Third Trimester Ultrasonography in Nulliparous Women in the Pregnancy Outcome Prediction (POP) Study: A Prospective Cohort Study," *Obstetrical & Gynecological Survey* 71, no. 3 (2016): 133–134.
11. G. C. S. Smith, "Should We Implement Universal Screening With Late Pregnancy Ultrasound to Prevent Stillbirth?," *BJOG: An International Journal of Obstetrics and Gynaecology* 125, no. 2 (2018): 101–103.
12. I. Monier, B. Blondel, A. Ego, M. Kaminiski, F. Goffinet, and J. Zeitlin, "Poor Effectiveness of Antenatal Detection of Fetal Growth Restriction and Consequences for Obstetric Management and Neonatal Outcomes: A French National Study," *BJOG: An International Journal of Obstetrics and Gynaecology* 122, no. 4 (2015): 518–527.
13. J. Henrichs, V. Verfaillie, P. Jellema, et al., "Effectiveness of Routine Third Trimester Ultrasonography to Reduce Adverse Perinatal Outcomes in Low Risk Pregnancy (The IRIS Study): Nationwide, Pragmatic, Multicentre, Stepped Wedge Cluster Randomised Trial," *BMJ (Online)* 367 (2019): 1–13.
14. G. C. Smith, A. A. Moraitis, D. Wastlund, et al., "Universal Late Pregnancy Ultrasound Screening to Predict Adverse Outcomes in Nulliparous Women: A Systematic Review and Cost-Effectiveness Analysis," *Health Technology Assessment (Winchester)* 25, no. 15 (2021): 1–190.
15. R. K. Morris, E. Johnstone, C. Lees, V. Morton, and G. Smith, "Investigation and Care of a Small-For-Gestational-Age Fetus and a Growth Restricted Fetus (Green-Top Guideline No. 31)," *BJOG: An International Journal of Obstetrics and Gynaecology* 131, no. 9 (2024): 31–90.
16. L. M. McCowan, F. Figueras, and N. H. Anderson, "Evidence-Based National Guidelines for the Management of Suspected Fetal Growth Restriction: Comparison, Consensus, and Controversy," *American Journal of Obstetrics and Gynecology* 218, no. 2 (2018): S855–S868.
17. C. C. Lees, T. Stampalija, A. Baschat, et al., "ISUOG Practice Guidelines: Diagnosis and Management of Small-For-Gestational-Age Fetus and Fetal Growth Restriction," *Ultrasound in Obstetrics & Gynecology* 56, no. 2 (2020): 298–312.
18. Institute of Obstetricians and Gynaecologists, Royal College of Physicians of Ireland and Directorate of Clinical Strategy and Programmes, Health Service Executive, *Guideline No 28. Fetal Growth Restriction—Recognition, Diagnosis & Management* (Institute of Obstetricians and Gynaecologists, Royal College of Physicians of Ireland and Directorate of Clinical Strategy and Programmes, Health Service Executive, 2017).
19. Perinatal Society of Australia and New Zealand and Centre of Research Excellence Stillbirth, *Position Statement: Detection and Management of Fetal Growth Restriction in Singleton Pregnancies* (Centre of Research Excellence in Stillbirth, 2019).
20. J. Kingdom, E. Ashwal, A. Lausman, et al., "Guideline No. 442: Fetal Growth Restriction: Screening, Diagnosis, and Management in Singleton Pregnancies," *Journal of Obstetrics and Gynaecology Canada* 45, no. 10 (2023): 102154.
21. J. Caradeux, R. J. Martinez-Portilla, A. Peguero, A. Sotiriadis, and F. Figueras, "Diagnostic Performance of Third-Trimester Ultrasound for the Prediction of Late-Onset Fetal Growth Restriction: A Systematic Review and Meta-Analysis," *American Journal of Obstetrics and Gynecology* 220, no. 5 (2019): 449–459.e19.
22. M. C. Vieira, S. Relph, D. Pasupathy, et al., "Evaluation of the Growth Assessment Protocol (GAP) for Antenatal Detection of Small for Gestational Age: The DESiGN Cluster Randomised Trial," *PLoS Medicine* 19, no. 6 (2022): e1004004-e.
23. T. J. Cole, J. V. Freeman, and M. A. Preece, "British 1990 Growth Reference Centiles for Weight, Height, Body Mass Index and Head Circumference Fitted by Maximum Penalized Likelihood," *Statistics in Medicine* 17, no. 4 (1998): 407–429.
24. F. P. Hadlock, R. B. Harrist, and J. Martinez-Poyer, "In Utero Analysis of Fetal Growth: A Sonographic Weight Standard," *Radiology* 181, no. 1 (1991): 129–133.
25. S. P. Chahaun, D. Parker, D. Shields, M. Sanderson, J. H. Cole, and J. A. Scardo, "Sonographic Estimate of Birth Weight Among High-Risk Patients: Feasibility and Factors Influencing Accuracy," *American Journal of Obstetrics and Gynecology* 195, no. 2 (2006): 601–606.
26. M. V. Vieira, S. Relph, M. Persson, P. T. Seed, and D. Pasupathy, "Determination of Birth-Weight Centile Thresholds Associated With Adverse Perinatal Outcomes Using Population, Customised, and Inter-growth Charts: A Swedish Population-Based Cohort Study," *PLoS Medicine* 16, no. 9 (2019): e1002902-e.
27. K. Dimassi, F. Douik, M. Ajroudi, A. Triki, and M. F. Gara, "Ultrasound Fetal Weight Estimation: How Accurate Are we Now Under Emergency Conditions?," *Ultrasound in Medicine & Biology* 41, no. 10 (2015): 2562–2566.
28. M. Oliver, G. McNally, and L. Leader, "Accuracy of Sonographic Prediction of Birth Weight," *Australian & New Zealand Journal of Obstetrics & Gynaecology* 53, no. 6 (2013): 584–588.
29. N. G. Anderson, I. J. Jolley, and J. E. Wells, "Sonographic Estimation of Fetal Weight: Comparison of Bias, Precision and Consistency Using 12 Different Formulae," *Ultrasound in Obstetrics & Gynecology* 30, no. 2 (2007): 173–179.
30. NMPA Project Team, "National Maternity and Perinatal Audit: Clinical Report 2019," in *Based on Births in NHS Maternity Services Between 1 April 2016 and 31 March 2017* (RCOG, 2019).

### Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** The distribution of clinically indicated foetal growth ultrasound in a selective cohort. **Table S1:** The rate and prediction of SGA by gestational age of last scan for those with  $\geq 1$  scan ( $n = 118,015$ ). **Table S2:** A comparison of birth rate with birth rate of those SGA within scan to birth intervals of 2–4 weeks for those with  $\geq 1$  scan ( $n = 118,015$ ). **Table S3:** Sensitivity (%) and Specificity (%) of EFW  $< 10$ th and abnormal umbilical doppler indices at the last scan to predict SGA within different scan to birth intervals for those with  $\geq 1$  scan ( $n = 118,015$ ). **Table S4:** Test performance characteristics (%) of EFW  $< 3$ rd at last scan to predict fetal growth restriction (birthweight  $< 3$ rd) across gestational age and scan to birth interval for those with  $\geq 1$  scan ( $n = 118,015$ ). **Table S5:** Positive and Negative Likelihood Ratio of each exposure at last scan to predict SGA at birth across gestational age of scan and scan to birth interval for those with  $\geq 1$  scan ( $n = 118,015$ ). **Table S6:** Funnel Plots to demonstrate heterogeneity across sites in sensitivity of EFW  $< 10$ th to predict SGA at birth within 2 weeks across gestational age of scan. **Table S7:** Funnel Plots to demonstrate heterogeneity across sites in sensitivity of EFW  $< 10$ th to predict SGA at birth within 4 weeks across gestational age of scan.