**Diagnostic accuracy of individual antenatal tools for the detection of the small for gestational age newborn.**

Borna POLJAK1.Umber AGARWAL, MD2. Richard JACKSON, PhD3.Zarko ALFIREVIC, MD1, Andrew SHARP, PhD1

**Affiliations:**

1. Department of Women’s and Children’s Health Research, University Department, Liverpool Women’s Hospital, Crown Street, Liverpool, L8 7SS, UK.
2. Fetal Medicine Unit, Liverpool Women’s Hospital, Crown Street, Liverpool, L8 7SS, UK.
3. Liverpool Cancer Trials Unit, 1-3 Brownlow St, Liverpool L69 3GL

**Correspondence:**

Dr Andrew Sharp

Department of Women’s and Children’s Health Research, Liverpool Women’s Hospital Crown Street, Liverpool, L8 7SS, United Kingdom.

Tel: +44 151 795 9560

Email: asharp@liv.ac.uk

**Abstract**

**Objective**

To define the accuracy of fetal and newborn growth charts for the detection of small for gestational age (SGA) fetuses (<10th centile).

**Design**Prospective cohort study.

**Setting**

UK specialist fetal growth clinic. **Population or Sample**

105 consecutive pregnant women referred with a fetus suspected of being SGA.  **Methods**All pregnancies were managed according to a standard protocol using estimated fetal weight (EFW) plotted on customised GROW charts. Last antenatal estimate of EFW (GROW, Mikolayczyk), abdominal circumference (AC Hadlock, Ig-21, AC Chitty) or change in AC over time (POP study) was compared to 4 birthweight charts (GROW, Ig-21, Mikolayczyk, WHO) and to APO.

**Main Outcome Measures**

Birthweight <10th centile**.
Results**

GROW defined 62 babies as having a birth weight <10th centile (59%), 57 babies using Mikolayczyk (54%), 55 babies using Ig-21 (52%) and 51 babies using WHO (48.6%). AC (Hadlock) had the best LR- (0.3-0.4) and best sensitivity for detecting SGA as defined by all 4 postnatal birthweight charts (74% - 82%). AC Ig-21 had the best LR+ (5.9-10.9) and specificity (94%-96%).

For adverse perinatal outcome (APO), AC (Hadlock) and EFW (GROW) had the best sensitivity (57% and 52%), whilst AC (POP) had the best LR+ (2.2) and best specificity (88%). APO detection increased to a sensitivity of 61% when AC (POP) and EFW <10th centile were combined but specificity was only 56%.

**Conclusions**We have identified a wide variation in the diagnostic accuracy of various tests for the detection of both SGA and APO dependent upon choice of chart. Suboptimal diagnostic accuracy of commonly used antenatal tests may lead to increasing medicalisation without preventing APO. Researchers should focus their attention on a combination of fetal biometry and biomarkers for better detection of SGA and prevention of APO.

**Keywords**

fetal growth restriction, small for gestational age, ultrasound prediction.

**Introduction**

There has been a growing drive to identify more accurately the fetus that is small for gestational age (SGA). Detection of the SGA fetus is intended to allow monitoring for pathological deterioration in fetal condition as part of intrauterine fetal growth restriction (IUGR) characterised by abnormal fetal Dopplers [1](#_ENREF_1), which may occur as a prelude to term stillbirth [2](#_ENREF_2). This is of particular importance in the UK, where stillbirth rates have recently been identified as being amongst the worst in developed countries [3](#_ENREF_3).

Within the UK screening for the SGA fetus relies upon measurement of the symphysial fundal height (SFH), plotted against population charts [4](#_ENREF_4) or customised charts in an attempt to increase the sensitivity of this technique [5](#_ENREF_5), however the detection rates remain poor [6](#_ENREF_6). If this primary screening raises a concerns about the possibility of an SGA fetus the woman will be referred for an ultrasound assessment. Screening using third trimester ultrasound is an obvious alternative, but the timing of the scan [7](#_ENREF_7) and its cost effectiveness remain controversial [8](#_ENREF_8). Also, it has been suggested that without the use of additional biomarkers these scans may be inefficient at predicting and preventing adverse perinatal outcome (APO) [9](#_ENREF_9), [10](#_ENREF_10).

Antenatal ultrasound detection of SGA based on abdominal circumference (AC) and estimated fetal weight (EFW) is far from straightforward, because these well defined parameters are estimated using a multitude of mathematical calculations that may give varying results for the same fetus. To further complicate matters, obtained results can then be plotted on a number of different antenatal reference charts generated from local, national or international cohorts some of which are customised for maternal factors, such as; parity, height, weight and ethnicity. These variations contribute to large differences in antenatal detection rates for SGA [5](#_ENREF_5), [7](#_ENREF_7), [10](#_ENREF_10), [11](#_ENREF_11).

The 10th centile has been studied most frequently, although lower centiles (3rd, 5th) have also been used in an attempt to grade the severity of SGA [12](#_ENREF_12), [13](#_ENREF_13). Hadlock et al. have suggested that the use of an EFW <10th centile may be superior to AC alone with good predictive capacity for the detection of SGA [14](#_ENREF_14), although even greater prediction occurs when both the AC and EFW are <10th centile [14](#_ENREF_14), [15](#_ENREF_15). It has recently been suggested that the change in velocity of growth across gestation may be a better predictor of APO than a single assessment [16](#_ENREF_16). Increasingly, fetal growth clinics are being introduced in order to optimise the diagnosis and management of suspected SGA fetuses.

The difficulty in accurate diagnosis and pregnancy planning is further exacerbated by the timing of the ultrasound assessment. Some fetuses may not be destined to continue to maintain the same growth velocity close to term. It may be these fetuses that are at greatest risk as they ‘cross’ centiles in late pregnancy, rather than being identified and remaining as SGA throughout gestation [16](#_ENREF_16). Following the confirmation of SGA labour induction is usually offered to eliminate the perceived risk of late stillbirth [12](#_ENREF_12).

Whilst the prevention of stillbirth is a laudable goal, the low sensitivity of the currently available screening tests and false positive labelling of SGA may lead to more interventions in pregnancy, which may increase maternal anxiety and has significant resource implications for maternity service providers.

The management of fetal growth restriction with abnormal Dopplers is very well established and fetal biometry is not critical for clinical decision making. The significant proportion of fetuses diagnosed as SGA without Doppler abnormalities represent a much greater clinical dilemma. Changes in the approach to managing these pregnancies have led to an increase in the demands and medicalisation of these pregnancies and it is these pregnancies that create the largest burden to health care services as a whole.

The difficulty in accurate assessment of growth restriction and any beneficial impact of screening programmes does not stop with birth. Whilst the birth weight of a baby is easily and routinely measured, the clinical relevance placed upon a birth weight centiles is determined by paediatric clinicians and a variety of postnatal/neonatal birth weight charts, which may be international, national, local or customised [17-19](#_ENREF_17).

There has been only a limited consideration of the impact of the variation in charts on the correlation between fetal biometry and birth weight [20](#_ENREF_20). We have, therefore, investigated the impact of different ultrasound assessment tools used in our high risk fetal growth clinic on the diagnostic accuracy of 3rd trimester ultrasound in predicting an SGA newborn. Our aim was to identify which antenatal assessment tools perform the best at detecting a birth weight <10th centile for gestational age when compared to a variety of postnatal charts. Ideally, one would expect a clinically useful test to achieve both sensitivity and specificity above 90%.

**METHODS**

We identified all consecutive cases of suspected SGA referred, after primary screening with symphysial fundal height measurement and ultrasonographer assessment, to specialist fetal growth clinic between January and October 2014. The indication for referral in all cases was ultrasound evidence of EFW <10 centile on the Gestation Related Optimal Weight (GROW) chart. These initial scans were performed by sonographers or specialists midwifes as part of the local screening programme.

A review in the specialist fetal growth clinic included a detailed assessment of fetal biometry, liquor and fetal Doppler. Further management depended upon the EFW using the Hadlock calculation [15](#_ENREF_15) which was plotted on a customised antenatal fetal weight chart (GROW) [18](#_ENREF_18). Fetuses confirmed as SGA (<10th centile) were managed in accordance with the RCOG green top guideline [5](#_ENREF_5) with regular ultrasound (every 2-3 weeks), umbilical and middle cerebral artery (MCA) Dopplers. Delivery was indicated if there was clinical concern about fetal or maternal wellbeing, or alternatively induction of labour was offered after 37 weeks. Fetuses that were identified as having an EFW >10th centile (AGA) and were referred back to routine antenatal care, although their outcomes were still included in the analysis.

For the purposes of this study, we have excluded all fetuses with abnormal Doppler, oligohydramnios or structural anomalies identified at any point in their care. This approach was taken as the purpose of the study was to identify variation in fetal growth assessment in a SGA population without critical IUGR.

On completion of pregnancy, fetal biometry parameters were assessed from the last ultrasound scan prior to delivery. AC velocity was calculated from 2 scans between 18-22 and 34-38 weeks of gestation. 23 cases had to be excluded from AC velocity calculations due to missing data (n=7), or delivery prior to 34 weeks (n=16).

All measurements were evaluated using the following parameters:

*Estimated fetal weight*

* GROW is a customised growth chart using fetal biometry and maternal characteristics to create an individualised growth curve for that pregnancy [18](#_ENREF_18).
* Hadlock produced an estimated fetal weight chart in 1991 based upon biometric assessment [21](#_ENREF_21).
* Mikolayczyk produced an international population with the express purpose of its being used in developing countries [19](#_ENREF_19).

*Abdominal circumference*

* Hadlock is a population chart based upon AC measurement [22](#_ENREF_22).
* Chitty is a population chart based upon AC measurement [23](#_ENREF_23).
* The Intergrowth 21st project (Ig-21) is a multiethnic population chart developed to provide an international standard for fetal growth [17](#_ENREF_17).
* AC velocity has been calculated using a tool provided by the Pregnancy Outcome Prediction (POP) study [16](#_ENREF_16).

*Birth weight charts*

* GROW is a postnatal growth chart which utilises fetal biometry and maternal characteristics to create an individualised growth curve for that pregnancy [18](#_ENREF_18).
* The intergrowth 21st project (Ig-21) is a multiethnic population chart developed to provide an international standard for neonatal growth [24](#_ENREF_24).
* The WHO neonatal and infant growth chart is a UK population based birthweight chart [25](#_ENREF_25). For this chart we have used the 9th centile as there is no 10th centile on postnatal charts.
* Mikolayczyk produced an international population with the express purpose of being used in developing countries [19](#_ENREF_19).

**Diagnostic accuracy**

The main outcome of interest was birthweight <10th centile for gestational age. Data for analysis were collected in terms of the number of true positives, true negatives, false positives and false negatives for each prenatal and antenatal assessment tool. Analyses are presented in terms of sensitivity, specificity, positive likelihood ratio and negative likelihood ratio for each prenatal assessment tool in relation to each postnatal birthweight assessment tool.

In order to calculate diagnostic accuracy for each of the 5 antenatal assessment tools we have used following definitions:

|  |  |
| --- | --- |
| **True positive** | EFW/AC <10th centile by ultrasound at the last scan before delivery that triggered iatrogenic delivery (induction of labour or elective CS) with a birth weight <10th centile |
| **True negative** | EFW/AC >10th centile by ultrasound at the last scan before delivery and did not undergo iatrogenic delivery for suspected IUGR/SGA followed by birth weight >10th centile |
| **False positive** | EFW/AC<10thcentile by ultrasound at the last scan before delivery that triggered an iatrogenic delivery (induction of labour or elective CS) followed by birth weight >10th centile |
| **False negative** | EFW/AC > 10th centile by ultrasound at the last scan before delivery and did not undergo iatrogenic delivery for suspected IUGR/SGA followed by birth weight <10 centile |

We also assessed the ability of each antenatal test to predict those babies which would go on to have an APO. We have used a composite of adverse perinatal outcome as defined by the POP study [16](#_ENREF_16) i.e. 5 min Apgar score of less than 7 OR metabolic acidosis (cord blood pH <7.1 and base deficit >10 mmol/L) OR term admission to the neonatal unit (admission <48 h after birth at ≥37 weeks’ gestational age and discharge ≥48 h after admission) OR stillbirth OR term livebirth associated with neonatal death OR hypoxic ischaemic encephalopathy OR use of inotropes OR need for mechanical ventilation OR severe metabolic acidosis (cord blood pH ≤7.0 and base deficit >12 mmol/L).

Given the strong correlations between post-natal measurements, further assessment of antenatal assessments are modelled using a hierarchical summary receiver operator characteristics (HSROC) approach [26](#_ENREF_26). This approach calculates the position and shapes of the receiver operator curve for each antenatal test. Posterior summaries of the fitted model allow for average estimates of antenatal test performance with results presented in terms of the negative likelihood ratio and positive likelihood ratio. Results are presented in terms of posterior medians with associated 95% credibility intervals.

**RESULTS**

118 women were referred as SGA during the study period. 77 found to be SGA on initial assessment and 41 were AGA. 13 fetuses had abnormal Doppler assessments at least once in their pregnancy and were excluded; leaving 105 women for analysis. Demographic and pregnancy outcome data can be found in Table One.

Actual Birthweight was plotted against a variety of commonly used postnatal charts. GROW defined 62 babies as having a birth weight <10th centile (59%), 57 babies using Mikolayczyk (54%), 55 babies using Ig-21 (52%) and 51 babies using WHO (48.6%).

**Estimated fetal weight for prediction of birthweight <10th centile**

The use of EFW plotted on antenatal GROW charts demonstrated a positive likelihood ratio of 1.9-3.0, negative likelihood ratio of 0.5-0.7, sensitivity of 53-56% and specificity of 72-81% (Table Two). The greatest sensitivity was 56% against postnatal GROW. The greatest specificity was against postnatal GROW.

EFW plotted on antenatal Mikolayczyk charts demonstrated a positive likelihood ratio of 4.2-5.1, negative likelihood ratio of 0.4-0.5, sensitivity of 60-67% and specificity of 80-86%. (Table Two) The greatest sensitivity was 67% against postnatal WHO chart. The greatest specificity was against postnatal IG-21.

**Abdominal Circumference for prediction of birthweight <10th centile**

The use of AC plotted on antenatal Hadlock charts demonstrated a positive likelihood ratio of 2.5-2.6, negative likelihood ratio of 0.3-0.4, sensitivity of 74-82% and specificity of 69-98% (Table Two). The greatest sensitivity was 82% against postnatal WHO chart. The greatest specificity was 98% against postnatal IG-21.

The use of AC plotted on antenatal Chitty charts demonstrated a positive likelihood ratio of 4.7-6.1, negative likelihood ratio of 0.5-0.6, sensitivity of 44-51% and specificity of 90-92%. The greatest sensitivity was 51% against postnatal WHO chart. The greatest specificity was 92% against postnatal IG-21 chart.

The use of AC plotted on antenatal IG-21 charts demonstrated a positive likelihood ratio of 5.9-10.9, negative likelihood ratio of 0.6-0.7, sensitivity of 35-41% and specificity of 94-96% (Table Two). The greatest sensitivity was 41% against postnatal WHO chart. The greatest specificity was 96% against postnatal WHO chart.

**Abdominal circumference velocity for prediction of birthweight <10th centile**

Assessment of antenatal AC velocity for the detection of a birthweight <10th centile demonstrated a positive likelihood ratio of 2.5-6.9, negative likelihood ratio of 0.8-0.9, sensitivity of 21-23% and specificity of 91-96% (Table Three). The greatest sensitivity was 23% against postnatal GROW, WHO and IG-21 charts. The greatest specificity was 96% against postnatal GROW.

We have also calculated the diagnostic accuracy of AC and EFW combinations and found that all point estimates of sensitivity and specificity either worsened or remained unchanged (data not shown).

**Antenatal prediction of adverse perinatal outcomes**

Overall the median gestation at delivery for women in the cohort was 37 (28+4 – 41+2). 34 (32%) women had a caesarean delivery. There were 2 (1.9%) stillbirths within the study cohort, both of which were SGA with normal Doppler, one of which suffered a placental abruption and the other acute fetal thrombotic vasculopathy.

There were 23 (24%) babies affected by an APO (Table Three). Antenatal detection (sensitivity) ranged from 22% to 57%, the specificity from 46%-88%, with a positive likelihood ratio of 0.7-2.2 and negative likelihood ratio of 0.8-1.2. The most sensitive antenatal test for prediction of APO was AC Hadlock (57%), whilst the most specific was AC velocity (88%). In addition, we calculated the diagnostic accuracy of AC growth velocity combined with EFW GROW <10th centile as suggested by Sovio et al (15) – sensitivity was 61% and specificity 56%.

**DISCUSSION**

**Main findings:**

We have demonstrated that the choice of both antenatal and postnatal tools to define SGA has critical impact on the detection rates of SGA babies. The sensitivity of different charts for the antenatal detection of SGA varied considerably depending upon the methods used, with sensitivities ranging from 21% to 82% and specificities from 69 to 98%. Overall, AC IG-21 had the best LR+ (5.9-10.9), LR- (0.6-0.7) and specificity (94%-96%). However, AC (Hadlock) had the best sensitivity for detecting antenatal SGA, ranging from 74% to 82%. None of the tests achieved both sensitivity and specificity above 90% (Figure One). Several charts (AC IG-21, Chitty AC, AC velocity, Mikolayczyk EFW and Hadlock EFW) demonstrated a positive LR of 5 but no one chart was paramount.

Given that in the UK, birth weight is most commonly calculated either by WHO or GROW charts, our data suggest that there is no single antenatal tool with adequate diagnostic accuracy for SGA detection. Whilst antenatal assessment using AC IG-21 has the best overall performance, it has unacceptably low sensitivity (35%-41%). Sensitivity is highest with AC (Hadlock) (74%-82%) but with unacceptably low specificity for low birth weight by WHO (69%) and GROW (70%). AC growth velocity does not appear to improve diagnostic accuracy for low birth weight in our selected cohort of high risk women.

None of the antenatal tests used were effective at predicting those pregnancies at risk of APO, with sensitivities ranging from 22 to 57% and specificities from 46 to 88%, best LR+ (0.7-2.2) and best LR- (0.8-1.2). However, when AC growth velocity was combined with EFW<10th centile as suggested by Sovio et al. [16](#_ENREF_16) the sensitivity and specificity was increased, but still remains relatively poor at 61% and 56% respectively.

**Strength and limitations:**

It is important to stress that our study did not evaluate diagnostic accuracy of a universal SGA screening programme. Our focus was on the diagnostic accuracy of antenatal fetal biometry tests as applied to a high risk population. Our cohort was already screened using standard techniques (symphysial fundal height and ultrasonographer scan) and then reviewed and managed in a specialist fetal growth clinic. We excluded fetuses with IUGR and abnormal Dopplers as there are several nationally endorsed management pathways for these pregnancies. We also excluded those fetuses who were >10th centile on ultrasonographer assessment and therefore were never referred to the specialist growth clinic. By focussing on SGA fetuses without critical deterioration in fetal Doppler, unlike some other similar size studies[1](#_ENREF_1), we aimed to address the diagnostic accuracy for the majority of SGA pregnancies for which there is clinical uncertainty in management.

We have tested all commonly used antenatal assessment charts and added the recently proposed AC velocity curve [16](#_ENREF_16). We have taken the same approach with postnatal charts to highlight the differences that exist in current practice. This is particularly relevant when we consider that many manuscripts on fetal growth do not report exactly which postnatal charts were used, with the implication being that there is no difference [10](#_ENREF_10), [16](#_ENREF_16). It should also be remembered that the initial screening for SGA was performed by ultrasonographers using GROW EFW charts. This pre-selection may have favoured subsequent performance of GROW EFW when repeated in our specialist clinic.

We acknowledge that our study is of a relatively small size and limited to high risk UK pregnancies. Therefore extrapolating our findings to other clinical setting and particularly low risk populations should be undertaken with caution. It is worth noting that a study with twice as many SGA cases would have only minimal impact on the precision of the estimates. For example, assuming that the results remain the same, doubling our sample size would result in a change of 95% CI for AC (Hadlock) sensitivity to detect low birthweight (WHO) from 0.69-0.92 to 0.74-0.89.

**Interpretation:**

Clinicians attempting to identify and manage SGA, with the ultimate goal of reducing stillbirth, must strive to find a balance between highly sensitive tests that accurately detect most SGA fetuses, versus very specific tests that incorrectly label only very few AGA fetuses as SGA. Tests with low sensitivity could lead to more stillbirths and poor perinatal outcome, whilst a low specificity inevitably leads to more interventions, anxiety and medicalisation of pregnancy without necessarily improving perinatal outcomes. The interpretation of this data relies very much on the ‘value’ placed on over diagnosis of normal fetuses as SGA versus the potential risk of harm from ‘missing’ an SGA fetus.

**CONCLUSION**

We have demonstrated that a range of currently used antenatal biometry parameters, when used in selected cases of suspected SGA and managed by senior medical teams experienced in ultrasound, have inadequate test accuracy leading to low detection rates that may lead to unnecessary interventions and no impact on APO. We suggest that clinically important improvements in diagnostic accuracy will not come from further refinement of fetal biometry data generated by routine growth scans, but from a combination of fetal biometry and placental biomarkers.

**Acknowledgements**

We are grateful to Prof Gordon Smith who kindly gave us access to an unpublished AC velocity calculator and provided helpful comments on the manuscript.

**Disclosure of interests**

None of the authors has any interests to disclose.

**Contributions to authorship**

AS, UA and ZA conceived the idea for the manuscript. BP and AS performed the data analysis. RJ performed the statistical analysis. BP, AS and ZA wrote the manuscript.

**Details of ethics approval**

Ethical approval was not sought as this formed part of an ongoing audit into the outcomes from the specialist fetal growth clinic.

**Funding**

No funding was received for this study.

**REFERENCES**

1. Marconi AM, Ronzoni S, Bozzetti P, Vailati S, Morabito A, Battaglia FC. Comparison of fetal and neonatal growth curves in detecting growth restriction. Obstet Gynecol. 2008 Dec;112(6):1227-34.

2. Gardosi J, Kady SM, McGeown P, Francis A, Tonks A. Classification of stillbirth by relevant condition at death (ReCoDe): population based cohort study. BMJ. 2005 Nov 12;331(7525):1113-7.

3. Flenady V, Koopmans L, Middleton P, Froen JF, Smith GC, Gibbons K, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. Lancet. 2011 Apr 16;377(9774):1331-40.

4. NICE. Antenatal care: Routine care for the healthy pregnant woman. London: National Institute of Health and Clinical Excellence; 2008.

5. RCOG. The investigation and management of the small-for-gestational-age fetus. Green-top guideline No31. 2013.

6. Neilson JP. Symphysis-fundal height measurement in pregnancy. Cochrane Database Syst Rev. 2000(2):CD000944.

7. Roma E, Arnau A, Berdala R, Bergos C, Montesinos J, Figueras F. Ultrasound screening for fetal growth restriction at 36 vs 32 weeks' gestation: a randomized trial (ROUTE). Ultrasound Obstet Gynecol. 2015 Oct;46(4):391-7.

8. NICE. Antenatal care routine care for the healthy pregnant woman (CG62). National Institute for Clinical Excellence. London: NICE; 2008.

9. Triunfo S, Parra-Saavedra M, Rodriguez-Sureda V, Crovetto F, Dominguez C, Gratacos E, et al. Angiogenic Factors and Doppler Evaluation in Normally Growing Fetuses at Routine Third-Trimester Scan: Prediction of Subsequent Low Birth Weight. Fetal Diagn Ther. 2015 Oct 1.

10. Callec R, Lamy C, Perdriolle-Galet E, Patte C, Heude B, Morel O. Impact on obstetric outcome of third-trimester screening for small-for-gestational-age fetuses. Ultrasound Obstet Gynecol. 2015 Aug;46(2):216-20.

11. Chauhan SP, Beydoun H, Chang E, Sandlin AT, Dahlke JD, Igwe E, et al. Prenatal detection of fetal growth restriction in newborns classified as small for gestational age: correlates and risk of neonatal morbidity. Am J Perinatol. 2014 Mar;31(3):187-94.

12. RCOG. The Investigation and Management of the Small-for-Gestational-Age Fetus. Royal College of Obstetricians and Gynaecologists. 2013;Green-top Guideline No.31:1-34.

13. Unterscheider J, Daly S, Geary MP, Kennelly MM, McAuliffe FM, O'Donoghue K, et al. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. Am J Obstet Gynecol. 2013 Apr;208(4):290 e1-6.

14. Chauhan SP, Cole J, Sanderson M, Magann EF, Scardo JA. Suspicion of intrauterine growth restriction: Use of abdominal circumference alone or estimated fetal weight below 10%. J Matern Fetal Neonatal Med. 2006 Sep;19(9):557-62.

15. Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. Radiology. 1984 Feb;150(2):535-40.

16. Sovio U, White IR, Dacey A, Pasupathy D, Smith GC. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. Lancet. 2015 Sep 7.

17. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet. 2014 Sep 6;384(9946):869-79.

18. Gardosi J, Mongelli M, Wilcox M, Chang A. An adjustable fetal weight standard. Ultrasound Obstet Gynecol. 1995 Sep;6(3):168-74.

19. Mikolajczyk RT, Zhang J, Betran AP, Souza JP, Gulmezoglu AM, Merialdi M. A global reference for fetal-weight and birthweight percentiles. Lancet. 2011;377(9780):1855-61.

20. Norris T, Johnson W, Farrar D, Tuffnell D, Wright J, Cameron N. Small-for-gestational age and large-for-gestational age thresholds to predict infants at risk of adverse delivery and neonatal outcomes: are current charts adequate? An observational study from the Born in Bradford cohort. BMJ Open. 2015;5(3):e006743.

21. Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. Radiology. 1991;181(1):129-33.

22. Hadlock FP, Deter RL, Harrist RB, Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. Radiology. 1984 Aug;152(2):497-501.

23. Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 3. Abdominal measurements. Br J Obstet Gynaecol. 1994 Feb;101(2):125-31.

24. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet. 2014 Sep 6;384(9946):857-68.

25. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med. 1998 Feb 28;17(4):407-29.

26. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20(19):2865-84.

**Figure/Table Legends**

Figure One

ROC plot to show the estimated sensitivity and specificity of each prenatal test with each antenatal assessment and forrest plots to show the estimated negative likelihood ratios and positive likelihood ratios from the HSROC modelling.

Table One

Demographic data for SGA cohort. \*only 2 cases of prelabour CS were delivered due to abnormal fetal monitoring, all others were for obstetric reasons – breech, previous CS, etc.

Table Two

Diagnostic accuracy of estimated fetal weight (EFW) and abdominal circumference (AC) in predicting birthweight <10th centile. Sensitivity (Sn), Specificity (Sp), Positive Likelihood Ratio (LR+), Negative Likelihood Ratio (LR-), Gestation Related Optimal Weight (GROW), Intergrow 21st (IG-21), World Health Organisation (WHO), Pregnancy Outcome Prediction (POP) study.

Table Three

Diagnostic accuracy of estimated fetal weight (EFW) and abdominal circumference (AC) in predicting adverse perinatal outcome. Gestation Related Optimal Weight (GROW), Intergrow 21st (IG-21), Pregnancy Outcome Prediction (POP) study. \*5 min Apgar score of less than 7 OR metabolic acidosis (cord blood pH <7.1 and base deficit >10 mmol/L) OR term admission to the neonatal unit (admission <48 h after birth at ≥37 weeks’ gestational age and discharge ≥48 h after admission) OR stillbirth OR term livebirth associated with neonatal death OR hypoxic ischaemic encephalopathy OR use of inotropes OR need for mechanical ventilation OR severe metabolic acidosis (cord blood pH ≤7.0 and base deficit >12 mmol/L).