Best Practice & Research Clinical Obstetrics and Gynaecology 49 (2018) 16–28

Contents lists available at ScienceDirect

Best Practice & Research Clinical Obstetrics and Gynaecology

journal homepage: www.elsevier.com/locate/bpobgyn

2

Universal screening for foetal growth restriction

Prof Gordon C.S. Smith, DSc, FMedSci *

Department of Obstetrics and Gynaecology, University of Cambridge, NIHR Cambridge Comprehensive Biomedical Research Centre, Cambridge, UK

Keywords:
Foetal growth restriction
Ultrasound
Doppler
Placenta
Biomarker
Screening

A B S T R A C T

Foetal growth restriction (FGR) is a major cause of morbidity and mortality. Clinical methods for identifying women whose pregnancies are affected by FGR do not perform well. Despite this, the current approach to screening includes the clinical assessment of risk and targeted use of ultrasound. Universal screening of women using ultrasound has not been shown to improve outcomes in randomised controlled trials and, when implemented nationally in France, appeared mostly to change outcomes for the worse through the effect of iatrogenic prematurity on false positives. Research is currently focused on the development of screening tests with higher sensitivity and specificity, for example, by combining ultrasound with placental biomarkers. The diagnostic tests employed should be identified through high-quality research that investigates the diagnostic accuracy of the tests, and this will usually involve blinding of the results. Therefore, future trials of screening and intervention will require careful planning. Moreover, if trials are to be powered for perinatal death, large sample sizes will be required.

Introduction

Foetal growth restriction (FGR) is generally defined as the failure of a foetus to grow according to its genetically determined growth potential. The condition is believed to be a major determinant of perinatal and childhood morbidity and mortality [1]; FGR may also lead to predisposition to a range of diseases later in life, in particular cardiovascular and metabolic disease [2]. Screening for FGR is part of

* Department of Obstetrics and Gynaecology, University of Cambridge, The Rosie Hospital, Box 223, Cambridge, CB2 2SW, UK.
E-mail address: paoandghod@medschl.cam.ac.uk.

https://doi.org/10.1016/j.bpobgyn.2018.02.008
1521-6934/

normal antenatal care. However, the approach to screening differs between countries, and there is a general awareness that screening for the condition is problematic because of both failure of detection of severely affected individuals and inappropriate medicalisation of false positives. The aim of the current review is to outline the current relevant knowledge on universal screening for FGR, to summarise the evidence base that informs current practice and to identify future approaches to improve clinical detection of the condition.

Defining FGR

A major problem in studying FGR is that there is no gold standard for identifying the condition, as the genetically determined growth potential of a foetus is unknown. A range of different proxies have been described. As none of these is a perfect measure of the disease, distinguishing FGR pregnancies from those that are not FGR using any of the available methods will be associated with misclassification, i.e. in every analysis of a cohort of pregnancies, there will be cases defined as FGR where the baby was healthy and cases where the baby had FGR but was wrongly classified as normal (Table 1).

Small for gestational age (SGA)

One of the most commonly employed methods for classifying infants as FGR is that the baby's weight (either estimated foetal weight [EFW] or birth weight [BW]) should be below a given threshold of the normal range for the given gestational age. The results of biometry are usually expressed as a percentile for the given gestational age and different thresholds are applied, such as <3rd, <5th or <10th. Some authors use the terms small-for-gestational age (SGA) and FGR interchangeably, although the majority of SGA babies are healthy and constitutionally small. The lower the weight threshold used, the smaller the proportion of false positives but the higher the proportion of false negatives. Choice of the threshold depends on the question being asked. For example, if a researcher wants to study the placenta from cases of suspected FGR using a very expensive test, a very low threshold should be set to avoid studying large numbers of healthy samples. However, if screening of a population is being performed with the aim of preventing an important adverse outcome, a higher threshold may be chosen if the consequences of failing to identify cases of FGR are serious.

Assessing the appropriateness of foetal growth

It might seem relatively simple to define the weight percentiles describing normal growth; however, there are a number of complexities. First, these thresholds are gestational age specific. This is particularly relevant when assessing the growth of babies born preterm, as they do not represent a random sample of the population at that gestational age. Prospective studies involving the serial ultrasonic assessment of foetal growth have demonstrated that slow growth between 20 and 28 weeks of gestational age is associated with an increased subsequent risk of spontaneous preterm birth [3,4]. Moreover, biochemical markers of abnormal placentation associated with FGR, such as low first trimester levels of plasma protein A (PAPP-A) and high second trimester levels of alpha-foetoprotein (AFP) are associated with an increased risk of preterm birth [5,6]. It follows, therefore, that the population of babies born preterm are over-represented with cases of FGR, and the BW distribution in preterm births will be shifted toward lower values compared with that in the whole population of foetuses of on-going pregnancies at that given week. At 40 weeks, the observed weight distribution of the babies born at this gestational age is probably a reasonable approximation of the normal range of

Table 1
A 2 × 2 table for assessing the screening performance of a measurement against the gold standard.

	Gold standard	
Measurement	FGR	Not FGR
Positive	True positive (TP)	False positive (FP)
Negative	False negative (FN)	True negative (TN)

healthy foetuses. However, for the reasons explained above, this will not be the case with preterm births with the practical consequence that foetal growth should be assessed with reference to measurements from foetuses of on-going pregnancies at a given gestational age rather than the assessment being based on the actual BW of infants born at the given gestational age. Fig. 1 shows that an infant with an actual BW on the 50th percentile judged against a BW-based standard is actually <3rd percentile at 28 weeks but around the 50th percentile at 40 weeks.

A second issue is the effect of physiological determinants of variability in the size of the baby. Consider two healthy women with normal pregnancies: the first is 1.80 m tall and weighs 75 kg and the second is 1.50 m tall and weighs 52 kg. They both have the same body mass index (23.1 kg/m^2), but it is self-evident that we would expect the taller, heavier woman to have the larger baby. Hence, when considering an individual case, it seems intuitive to consider the mother's anthropometry. Methods have been developed to 'customise' the EFW or BW centile for the mother's characteristics. Although intuitively attractive, there are a number of concerns about this approach, in particular how to differentiate between physiological and pathological parental determinants of foetal growth. For example, nulliparity is associated with a lower BW, and it is also associated with some of the most serious associations with FGR, such as stillbirth and pre-eclampsia. If foetal growth lies on the causal pathway between nulliparity and adverse outcome, customising percentiles for this variable may result in a weaker association with disease and this has been observed using large-scale studies of perinatal death [7]. Similarly, another "physiological" determinant of foetal growth, maternal short stature, has also consistently been shown to be associated with preterm birth [8] and is also associated with stillbirth [9]. Finally, even if it is accepted that variation within the normal range of maternal weight is a physiological determinant of variation in the size of the foetus, adjustment across the whole range of weight would cease to be physiological at the extremes. Women weighing 40 kg or 140 kg would not usually be regarded as exhibiting physiological variation in weight. Hence, customisation involves deciding a point on the scale for a given variable where further adjustment should not be performed. Overall, it is currently unclear whether customisation is appropriate and the significant weaknesses in the evidence base around this method are discussed in more detail below.

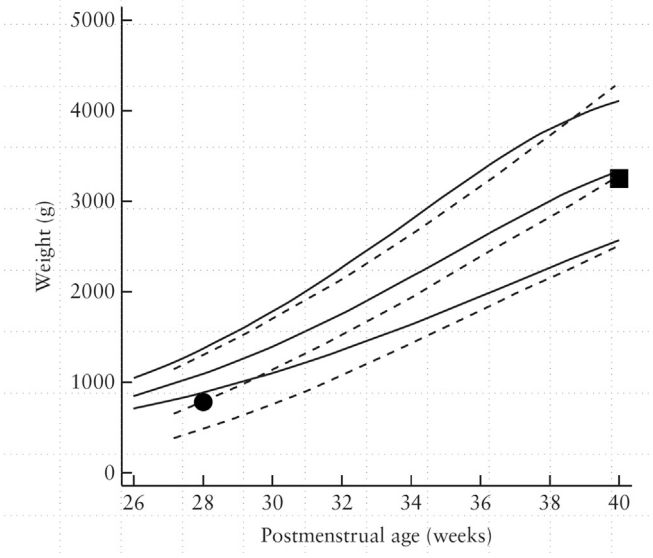

Fig. 1. Median, 3rd, and 97th percentiles for gestational age comparing actual birth weight (broken line) and foetal weight estimated using an ultrasound (solid line) from the InterGrowth-21st project. Symbols indicate a neonate with a birth weight on the 50th percentile as judged by the birth weight standard at 28 weeks (circle) and 40 weeks (square). The graph illustrates that the use of a normal range based on birth weight measurements will systematically underestimate the degree of FGR in preterm infants. Graph used (with modifications) with permission from Stirnemann et al., Ultrasound Obstet Gynecol 2017; 49: 478–486.

A third issue is whether the normal pattern of foetal growth differs according to the parental race and ethnicity. One problem in addressing this is that parental anthropometry varies between different populations, and this could lead to differences in foetal growth and BW. A second problem is that studying populations in different countries worldwide means studying women who also vary considerably in their exposure to disease and sub-optimal (or supra-optimal) nutrition. Hence, dissecting out differences in foetal growth between populations is problematic. This then raises an additional complication, namely, whether it is possible to define the normal range of growth for all women or whether each country requires its own growth norms. Two major studies addressed this question using similar methods and drew different conclusions; hence, the question of creating a globally appropriate chart for normal foetal growth remains unresolved [10,11].

Screening tests for FGR: ultrasonic

The primary method for assessing foetal growth is to estimate the weight of the foetus, to define the foetus as SGA on the basis of an EFW threshold for gestational age, and then to differentiate healthy SGA from FGR using other markers of pathological growth impairment. These include assessment of blood flow in the uteroplacental or foetal vessels, ultrasonic examination of the placenta, and serial assessment of foetal growth. These methods are briefly described below, as they are discussed in more detail elsewhere in this volume.

Doppler flow velocimetry

In Doppler flow velocimetry, a flow velocity waveform is plotted with the X axis as time and the Y axis as the calculated speed of flow. Based on these plots, a number of indices are calculated, which reflect the resistance to blood flow downstream of the site of measurement (reviewed in detail elsewhere) [12]. The most commonly used quantification in the UK is the pulsatility index.

Umbilical artery Doppler

High-resistance flow in the umbilical artery is associated with FGR. Along with quantitative indices, the qualitative appearance of the waveform is assessed on the basis of the presence, absence, or reversal of forward flow in the vessel at the end of diastole. High-resistance patterns of flow are associated with perinatal morbidity and reflect maldevelopment of the placental villous vascular tree [13]. Umbilical artery Doppler is probably the single most informative measurement made by ultrasound for the detection of early-onset FGR.

Uterine artery Doppler

High-resistance flow in the uterine arteries is associated with both pre-eclampsia and FGR. The physiological basis for this measurement is thought to be that high-resistance patterns of flow reflect a lesser degree of invasion of the maternal spiral arteries by the extravillous trophoblast (EVT). The non-pregnant uterine circulation has a high resistance owing to the presence of the maternal spiral arteries. Healthy pregnancy is associated with invasion of these vessels by the EVT and subsequent reduction in the resistance to flow. The measurement is informative from the first trimester but is most commonly measured in the middle of pregnancy [14]. However, uterine Doppler can also be measured during the later stages of pregnancy and is thought to be similarly informative for measuring adverse outcome.

Middle cerebral artery (MCA) Doppler

The foetal cerebral circulation usually exhibits a high-resistance pattern of flow in the MCA, although the measures progressively decline toward term. One component of the foetal arterial chemoreceptor-mediated response to hypoxia is to increase the cerebral proportion of the cardiac output. Hence, a low-resistance pattern of flow in the MCA is thought to be indicative of foetal hypoxia [15].

Ductus venosus Doppler

The ductus venosus is part of the system of shunt blood vessels in the foetus, which allows gaseous exchange to take place at the placenta *in utero*. The ductus connects the umbilical vein to the inferior

vena cava (IVC), bypassing the foetal liver. Flow is streamed in the IVC and the oxygenated blood from the ductus venosus is preferentially directed to the left atria by the crista dividens. Hence, the ductus venosus helps maintain the higher levels of oxygenation of blood pumped by the left ventricle. Increased pressure in the foetal venous system is thought to be a feature of foetal compromise and hence reflected in high-resistance patterns of ductus venosus flow [15].

Growth velocity and growth trajectory

It has been argued that, by definition, the optimal way of classifying FGR is by the analysis of serial measurements of foetal growth [16]. Deter has described methods for deriving the expected foetal size from third trimester measurements based on the magnitude of second trimester measurements [17]. Another approach is to calculate the deviation of a given measurement from the mean (using a gestational age-corrected z score) and then to assess the change in the deviation calculated from a second scan later in pregnancy [18]. Using these approaches, each foetus acts as its own control. The assumption is that the factor leading to growth restriction would have little or no effect at the earliest measurement but its effect would become manifested with advancing pregnancy. While there is evidence to suggest that this might be one of the best ways to identify foetuses at increased risk of morbidity [18], there is a large body of evidence to suggest that factors associated with FGR are apparent in the first trimester of pregnancy, i.e. well before the conventional first full assessment of foetal biometry at 20 weeks [19].

Other

One approach of reducing noise and simplifying measurements (e.g. avoiding the need for gestational age correction) is to use ratios of informative measurements. A Doppler ratio called the cerebroplacental ratio (CPR) is quite widely discussed in the recent literature. This is the ratio of the pulsatility index in the MCA to the pulsatility index in the umbilical artery. A recent systematic review concluded that the ratio may be more informative than the umbilical Doppler on its own, but the quality of the studies supporting this conclusion was not high [20]. A ratio of two biometric measurements (the head and abdominal circumferences) has also been proposed to identify FGR on the basis that true FGR is likely to exhibit evidence of sparing of brain growth over the growth of the foetal abdomen. While there seems to be an association between a high ratio and FGR, small foetuses with a normal ratio were still at increased risk of perinatal morbidity compared to foetuses with appropriate-for-gestational age biometry [18]. Other approaches include ultrasonic assessment of the placenta (thickness, presence of lesions, and patterns of calcification), assessment of amniotic fluid (amniotic fluid index or deepest vertical pool) and assessment of the foetal biophysical status (movement, tone, presence of breathing movements and analysis of the foetal heart rate trace).

Screening tests for FGR: biochemical

First trimester biochemical predictors

The placenta is thought to have a key role in the aetiology of FGR and this reflects primarily its transport function, mediating materno–foetal exchange of nutrients, gases and waste products. It is increasingly appreciated that placental dysfunction has its origins, at least in part, in the first trimester of pregnancy [19]. The placenta releases multiple factors into the maternal circulation from the very earliest stages of pregnancy, and first trimester maternal serum levels of a number of these factors have been shown to be associated with the risk of complications during the later stages of pregnancy. There is a substantial body of evidence associating maternal serum levels of PAPP-A and the later risk of complications. PAPP-A is a protease for insulin-like growth factor-binding proteins (IGFBPs) 4 and 5. Higher levels of IGFBPs would be expected to lead to lower levels of insulin-like growth factors, which are key foetal growth factors. Consistent with this, low PAPP-A is associated with a higher risk for SGA and stillbirth and knocking out PAPP-A in mice results in FGR [5,21]. Other placenta-derived proteins

are predictive of later complications. First trimester levels of placental growth factor (PlGF) and soluble fms-like tyrosine kinase (sFlt-1) [22] are negatively associated with the subsequent risk of delivering an SGA infant. The latter association is in contrast to the pattern during later stages of pregnancy, where higher sFlt1-1 is associated with an increased risk of adverse outcome. Low PAPP-A and PlGF are also associated with an increased risk of preterm birth, underlining the relationship between early placental dysfunction, FGR and spontaneous preterm birth [5,22].

Second trimester biochemical predictors

Elevated maternal serum levels of AFP are associated with an increased risk of placenta-related complications including stillbirth due to FGR [9]. AFP is a foetal oncoprotein and high maternal levels are thought to reflect abnormal placental permeability. A combination of low PAPP-A in the first trimester and high AFP in the second is particularly strongly predictive for complications [23]. However, AFP is less commonly measured now as a screening tool for Down's syndrome and this screening has moved to the first trimester and has also been somewhat overtaken by non-invasive testing for Down's syndrome using maternal plasma cell free DNA. However, some clinicians perform the quad test as a means of assessing placental function and have incorporated this into the assessment of high-risk women [24].

Third trimester biochemical predictors

Performance of biochemical tests of placental function used to be widespread before the availability of high-quality ultrasound. Biochemical testing was both introduced and withdrawn without strong evidence for its utility or lack of utility. Combining ultrasonic assessment of foetal size with biochemical assessment of placental function would be an attractive screening approach, and there is recent evidence that maternal serum levels of placenta-derived proteins can differentiate between healthy SGA and FGR infants [25]. Prospective studies have been conducted and analyses are in progress; however, there is currently no strong evidence to support the combination as a screening test for FGR. Data do exist in the context of women presenting with suspected pre-eclampsia that measurement of PlGF performs well as a predictor of FGR [26].

The clinical approach to screening

Current clinical practice

Currently, all women are screened for FGR; however, the method for screening differs in worldwide. In the USA and UK, women are not routinely scanned after the 20-week anomaly scan. Rather, women are assessed for risk factors at booking (e.g. previous FGR or relevant medical comorbidities) and for complications arising during the pregnancy (e.g. reduced foetal movements or antepartum haemorrhage) and by clinical examination (e.g. using the symphyseal fundal height), and they are selected for ultrasound scanning on the basis of clinical indications. The justification for this view is the Cochrane review of 13 randomised controlled trials (RCTs) including ~35,000 women, which demonstrated no benefit from universal screening [27]. There are a number of issues, however, on the quality of the meta-analysis. First, the 13 RCTs did not have a standard definition of a positive screening result. Hence, it is not clear whether combining their results is informative. Second, none of the 13 RCTs coupled screening to a known effective intervention. Clearly, screening will only result in better outcomes when coupled to a disease-modifying intervention. Third, sample size calculations indicate that, even with very optimistic estimates of the predictive strength of ultrasound and with very optimistic effects of interventions on the outcome, a sample size of 35,000 would be inadequate to detect a reduction in the risk of perinatal death [28].

The weaknesses in the evidence base have led some countries to implement universal screening in the absence of evidence for safety or effectiveness. A study of outcomes in France, where a routine scan was implemented between 30 and 34 weeks, failed to demonstrate improved outcomes among SGA infants

correctly identified by the programme (true positives) compared with SGA infants who were missed (false negatives) [29]. Indeed, a number of adverse outcomes were more common in the true positives. Moreover, the analysis clearly demonstrated that, among low-risk women who had a false-positive diagnosis of SGA, there was an increased risk of preterm birth, need for neonatal resuscitation and admission to neonatal intensive care. These findings suggest that the *ad hoc* implementation of universal screening for FGR should not be considered, as it has the potential to be more harmful than beneficial.

Why would screening cause harm?

The sections below document the ways in which a positive screening test might change clinical management. However, in essence, there is a single highly effective disease-modifying therapy: medically indicated delivery removes the subsequent risk of antepartum stillbirth associated with FGR. However, medically indicated delivery, by definition, entails delivering the baby at an earlier gestational age than would have occurred had the pregnancy been allowed to continue, and prematurity is one of the major determinants of neonatal morbidity. Hence, the timing of the screening test and the delivery are clearly key. These may be the main factors which resulted in the increased risk for complications following implementation of universal screening using ultrasound during late pregnancy in France. The gestational age window for screening was 30–34 weeks, which means that all foetuses identified as being at high risk and thought to require immediate delivery would then be exposed to the risk of prematurity. If the test had a very high positive predictive value for stillbirth, the net effect could still be beneficial. However, the absolute risk of stillbirth for each week of gestational age in this interval is about 1 loss per 2000 women per week [30]. Hence, even if a test had a positive likelihood ratio of 100, the positive predictive value would still only be ~5%. There are, potentially, some combinations of ultrasound measurements which would achieve this level of predictive value, e.g. EFW <3rd combined with absent or reversed end diastolic flow (EDF) in the umbilical artery. However, the net harm which appears to be caused by screening likely reflects early delivery for less severe ultrasonic abnormalities.

Clinical methods used for screening FGR

Currently, the main clinical approach to screening for FGR is to measure the symphyseal-fundal height (SFH) with a tape measure. It is well recognised that this has low sensitivity to detect SGA [18]. A refinement of the approach is to use a method that accounts for the mother's characteristics, discussed above and termed customisation. This method has two elements, the Gestation-Related Optimal Weight (GROW) method for customising SFH and EFW measurements and a related package of care which has been widely implemented in the UK called the Growth Assessment Protocol (GAP). Interestingly, the 2008 NICE Guideline considered use of customised assessment of SFH measurement [31]. Despite their conclusion that the method was unproven, it has been extensively implemented. The key piece of evidence supporting this approach is a RCT of the method. This demonstrated an increase in the sensitivity for the detection of SGA from 29% to 48% [32]. However, there was no information on the specificity. It is impossible to assess two screening approaches based on sensitivity alone. Subsequent studies have shown an apparent fall in the stillbirth rate in England and Wales which coincided with the implementation of GAP [33]. Such 'before and after' comparisons should be interpreted very carefully, given the multiple possible influences on overall rates of stillbirth. Moreover, it has been estimated that the use of customised SFH might prevent around 5% of stillbirths, even assuming that this screening test performs when implemented as it did in the research studies [34]. Finally, comparison of publicly available data on stillbirth rates in Scotland, where GROW and GAP were not widely implemented, demonstrates that the fall in stillbirth rate was actually greater in Scotland than in England and Wales (Fig. 2). This does not mean that GROW and GAP are not clinically effective, rather it means that the fall in stillbirth rates in England and Wales cannot be used as evidence to suggest that they are effective. A stepped wedge cluster RCT of the method is currently on-going (The DESiGN Trial, London UK, PI Dr Dharmintra Pasupathy), and this may provide definitive evidence for the safety and effectiveness of the approach.

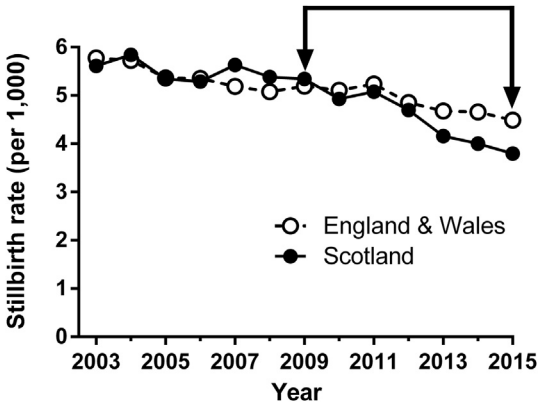

Fig. 2. Stillbirth rates from England and Wales and from Scotland, between 2003 and 2015. Data were obtained from publicly available sources: Office of National Statistics (https://www.ons.gov.uk) and National Records of Scotland (https://www.nrscotland.gov.uk). The arrows indicate the period when GROW implementation increased from <10% to >90% for births in England.

How can we evaluate new methods for screening?

There are two elements to screening. The first is to perform the screening test, and the second is to apply the intervention in the women who screen positive. A trial of screening could fail to demonstrate net benefit either because the test does not work or because the intervention is ineffective (or any good done is outweighed by harm). Evaluating a screening test is not as straightforward as it might appear on first consideration. It is well accepted that high-quality studies of new drugs require a blinded comparison of the new drug versus either a placebo or an active comparator. However, it is less widely recognised that a screening test has to be evaluated in a blinded fashion with the standard of care. The issue is that if the result of the new test is revealed, it could lead to biases. An example is considering delivery for suspected foetal distress as an outcome. A number of studies have reported associations with non-blinded ultrasonic markers and delivery for foetal distress [35,36]. However, it is self-evident that the association could be explained by the individual performing the delivery doing so wholly or in part based on the test result. This could lead to a false-positive association between the test and adverse outcome. However, it could also be that the test is useful and the interventions initiated by the test prevent the adverse outcome. In this setting, knowledge of the test result could lead to a false-negative association, e.g. the test correctly identified foetuses at increased risk of intrapartum asphyxia, but there was no increased risk of the asphyxia due to interventions that were conducted because of the test result. There is a view that it might be unethical not to reveal the result of the test being evaluated. However, this makes an assumption that revealing the test result would be more beneficial than harmful. When the evidence does not exist to support this view, it is hard to see what ethical argument could be applied. One could equally argue that a woman should not be exposed to the risks associated with an unproven screening test. The lessons learned from interventional studies, namely, that there is no alternative to conducting high-quality unbiased studies, urgently need to be learned in the context of screening for FGR.

The major beneficial effect of screening for preterm FGR is to prevent stillbirth. It is possible that identifying FGR infants might lead to interventions that prevent non-lethal complications, e.g. true positives may benefit from antenatal corticosteroids and pre-labour caesarean delivery. However, it is likely that any beneficial effect on non-lethal complications through intervention would be outweighed by harm caused by iatrogenic preterm delivery. It follows, therefore, that trials would ideally study the effect of screening on the risk of perinatal death. Otherwise, trials of screening may be underpowered to demonstrate benefit and only powered to detect harm. The problem from the perspective of research is that stillbirth is rare. Sample size calculations demonstrate that assuming a 5% screen-positive rate, a positive likelihood ratio of 10 and an intervention which reduced the risk of perinatal death by 50%, a total of ~130,000 women would be required. Even this assumes that

researchers would know the positive likelihood ratio for perinatal death for the given test. Establishing this would, in turn, require an extremely large study. One practical way forward is to develop screening tests based on non-lethal proxies for stillbirth, such as severe FGR. A programme of screening and intervention could then be designed where screen-positive women were randomised to intervention or concealment of the test result. The advantages of this approach are a lower sample size and that the validation of the screening test and assessment of the intervention are performed in a single study. The alternative approaches are illustrated in Fig. 3 and the issues are discussed in detail elsewhere [37].

Interventions in women who screen positive

As discussed above, screening is only justified if a disease-modifying intervention is available. There are on-going studies that examine novel therapeutic approaches such as the selective phosphodiesterase inhibitor sildenafil citrate ('Viagra') in severe early-onset FGR [38]. However, the primary justification for screening at present is to target a series of interventions, which would normally be implemented in a high-risk population, and to screen positive but otherwise low-risk women.

Antenatal management

Identification of women with FGR would lead to a number of changes to their antenatal care. These patients are advised to have a particularly low threshold to report any concerns regarding foetal movements. They are asked to attend frequent ultrasound scans with foetal biometry generally assessed at 2 weekly intervals. Plotting of serial measurements would reveal whether there was further slowing in growth, confirming the diagnosis of FGR. Doppler ultrasound assessment of the foetus, often including umbilical artery, MCA and ductus venosus, are generally performed more frequently, typically twice per week in women with FGR. In the USA, formal assessment of the biophysical profile (BPP) might be performed. In Europe, although the BPP is less widely employed, both the liquor volume and foetal activity are generally assessed.

An important large-scale study that compared different methods of antenatal monitoring in the context of early-onset FGR, the Truffle study [39], is discussed elsewhere. It is worth flagging one slightly artificial element of the study, namely, the reduction of decision-making to a single aspect of foetal monitoring. In practice, assessment of such cases and a diagnosis of a deteriorating state would frequently be based on multiple measurements. In essence, the main concern is that intra-uterine foetal death may occur during the period between assessments. In high-risk cases, a woman might be asked to attend for daily computerised assessment of the CTG and in very high-risk cases might be admitted to hospital to allow assessment twice or three times per day. Intra-uterine foetal death is very uncommon when a foetus is identified as having high-risk. In the Truffle study, only 2% of cases resulted in intra-uterine death despite the fact that this was an extremely high-risk group.

Timing of delivery

The primary disease-modifying intervention to prevent stillbirth is planned delivery. In the context of post-date delivery, where there is no adverse effect of the earlier gestational age of delivery on the risk of neonatal death, routine induction of labour decreased the risk of perinatal death by ~70% [40]. However, at preterm gestational ages, the main decision-making in the context of a baby known to be FGR is balancing the risks of neonatal morbidity and mortality if the baby is delivered early versus the risk of antepartum stillbirth if it is not. There is a continuum of risk associated with gestational age at delivery both for short-term and long-term adverse effects. This has been nicely illustrated for the risk of special educational needs (SEN), where the risk of the outcome was seen to decline with advancing gestational age at delivery up to the 40th week of gestational age (Fig. 4). Moreover, the strength of the association between gestational age and SEN raises the sobering consideration that the iatrogenic morbidity caused by *ad hoc* screening and early delivery of false positives could be lifelong. This underlines the importance of a strong evidence base before implementation of a screening programme.

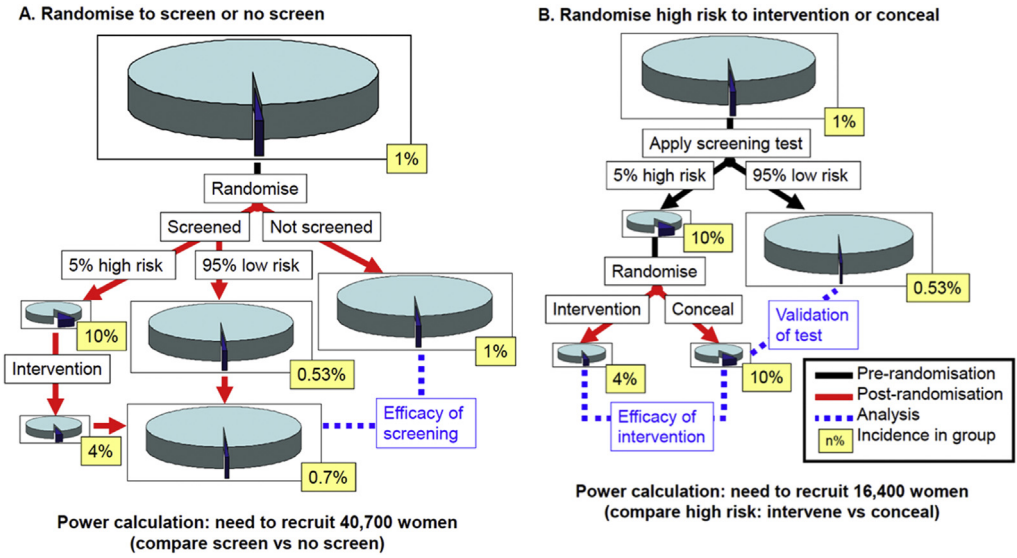

Fig. 3. Alternative study designs for randomised controlled trials of screening. Randomising women who had a positive screening test result (to intervention or conceal) reduces the sample size and allows separate assessment of the diagnostic effectiveness of the screening test and the clinical effectiveness of the intervention. The assumptions are a primary outcome affecting 1% of women, a screening test which is positive in 5% of the population and has a positive predictive value of 10%, and an intervention which reduces the risk of the outcome by 60%. Note: as the sample size calculation assumes a 1% background risk, this would be appropriate for a more common outcome than perinatal death. Graph taken from Smith GCS, PLoS Med 2012; 9: e1001274 (paper and content open access).

Mode of delivery and intrapartum management

Infants with FGR are at increased risk of intrapartum asphyxia. In severe cases, pre-labour caesarean may be considered, as this eliminates the possibility of asphyxia due to labour. However, there are multiple concerns about caesarean delivery and, wherever possible, induction of labour should be considered. Conduct of labour will be influenced by the diagnosis of FGR. The evidence base supporting

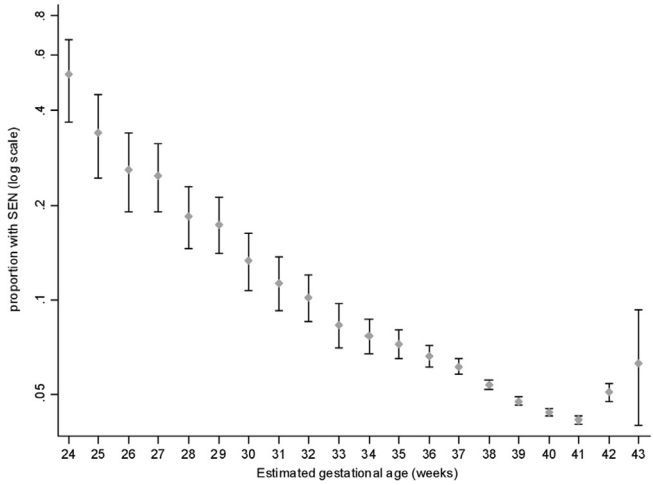

Fig. 4. Relationship between the week of gestational age at delivery and the risk of special educational needs in childhood. Graph taken from Mackay et al. PLoS Medicine 2010; 7:e1000289 (paper and content open access).

continuous electronic foetal monitoring (EFM) is less than ideal and it is not clear whether it reduces the risk of perinatal death in low-risk women [41]. However, in the high-risk situation of pregnancies complicated by FGR, continuous EFM is the norm. Moreover, knowledge that the baby is FGR and at increased risk of asphyxia may result in a lower threshold to check foetal scalp blood for pH and to perform emergency caesarean delivery for suspected foetal distress.

Summary

There is a great deal of knowledge about the ultrasonic indicators of FGR, about the placental causes of FGR and other associated tests that indicate FGR. Moreover, FGR remains a major cause of stillbirth, and our current clinical methods for screening are crude and ineffective. It seems plausible that applying high-tech methods to low-risk women would result in benefit. However, the direct evidence to support this is lacking: rather, there is evidence that suggest such methods cause net harm. There is an urgent need for better screening tests for FGR, identified through carefully designed studies. There is also an urgent need for better data about how the existing tests perform in screening low-risk women for FGR. Future studies of screening and intervention need to be appropriately powered to address the most important outcomes, need to be conducted with high-quality data on the screening properties of the test and need to couple the test with a known, effective intervention. As the major harm associated with false positives is the effect of iatrogenic prematurity, there is a rationale to focus initial research efforts on screening women for late FGR.

Conflicts of interest

GS received research support from GE (supply of two diagnostic ultrasound systems) and from Roche (supply of equipment and reagents for biomarker studies, approx. £600,000 in value). GS has been paid to attend advisory boards by Roche. GS has other commercial interests outside the scope of the current paper.

Practice points

- Clinical assessment for FGR performs poorly as a screening test.
- Screening all pregnant women using universal late pregnancy ultrasound does not appear to confer benefit and may cause harm by increasing iatrogenic prematurity.
- Foetal abdominal circumference growth velocity appears to be the best-performing single ultrasonic measure to differentiate between healthy SGA and FGR infants.
- Foetal biometry should be associated with a foetal growth-based standard: the appropriateness of EFW measurements should not be quantified in relation to a BW-based normal range, as this will result in false-negative results at preterm gestational ages.
- The evidence base supporting the use of customisation of EFW measurements is weak.

Research agenda

- Screening tests need to be evaluated using high-quality study designs. It is difficult to evaluate a novel screening test without blinding.
- Combining ultrasonic measurement of foetal size with biochemical assessment of placental function may allow more sensitive and specific detection of FGR.
- Future trials of screening and intervention need to couple the screening test with an effective disease-modifying intervention.
- Late FGR may be more amendable to screening and intervention for future trials due to the reduced risk of iatrogenic delivery.

References*

[1] Resnik R. Intrauterine growth restriction. Obstet Gynecol 2002;99(3):490–6.
*[2] Barker DJ. Adult consequences of fetal growth restriction. Clin Obstet Gynecol 2006;49(2):270–83.
[3] Gaillard R, Steegers EA, de Jongste JC, Hofman A, Jaddoe VW. Tracking of fetal growth characteristics during different trimesters and the risks of adverse birth outcomes. Int J Epidemiol 2014;43(4):1140–53.
[4] Partap U, Sovio U, Smith GC. Fetal growth and the risk of spontaneous preterm birth in a prospective cohort study of nulliparous women. Am J Epidemiol 2016;184(2):110–9.
[5] Smith GC, Stenhouse EJ, Crossley JA, Aitken DA, Cameron AD, Connor JM. Early pregnancy levels of pregnancy-associated plasma protein a and the risk of intrauterine growth restriction, premature birth, preeclampsia, and stillbirth. J Clin Endocrinol Metab 2002;87(4):1762–7.
[6] Smith GC, Shah I, White IR, Pell JP, Crossley JA, Dobbie R. Maternal and biochemical predictors of spontaneous preterm birth among nulliparous women: a systematic analysis in relation to the degree of prematurity. Int J Epidemiol 2006; 35(5):1169–77.
[7] Iliodromiti S, MacKay DF, Smith GC, Pell JP, Sattar N, Lawlor DA, et al. Customised and noncustomised birth weight centiles and prediction of stillbirth and infant mortality and morbidity: a cohort study of 979,912 term singleton pregnancies in Scotland. PLoS Med 2017;14(1). e1002228.
[8] Oliver-Williams C, Fleming M, Monteath K, Wood AM, Smith GC. Changes in association between previous therapeutic abortion and preterm birth in Scotland, 1980 to 2008: a historical cohort study. PLoS Med 2013;10(7). e1001481.
[9] Smith GC, Shah I, White IR, Pell JP, Crossley JA, Dobbie R. Maternal and biochemical predictors of antepartum stillbirth among nulliparous women in relation to gestational age of fetal death. BJOG 2007;114(6):705–14.
*[10] Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh IL, 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;384(9946):869–79.
*[11] Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup JL, et al. The World Health Organization fetal growth charts: a multinational longitudinal study of ultrasound biometric measurements and estimated fetal weight. PLoS Med 2017; 14(1). e1002220.
[12] Kalache KD, Duckelmann AM. Doppler in obstetrics: beyond the umbilical artery. Clin Obstet Gynecol 2012;55(1): 288–95.
[13] Kingdom JC, Burrell SJ, Kaufmann P. Pathology and clinical implications of abnormal umbilical artery Doppler waveforms. Ultrasound Obstet Gynecol 1997;9(4):271–86.
[14] Rodriguez A, Tuuli MG, Odibo AO. First-, second-, and third-trimester screening for preeclampsia and intrauterine growth restriction. Clin Lab Med 2016;36(2):331–51.
[15] Everett TR, Peebles DM. Antenatal tests of fetal wellbeing. Semin Fetal Neonatal Med 2015;20(3):138–43.
[16] Altman DG, Hytten FE. Intrauterine growth retardation: let's be clear about it. Br J Obstet Gynaecol 1989;96(10):1127–32.
[17] Deter RL. Individualized growth assessment: evaluation of growth using each fetus as its own control. Semin Perinatol 2004;28(1):23–32.
*[18] Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. 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;386:2089–97.
[19] Smith GC. First trimester origins of fetal growth impairment. Semin Perinatol 2004;28(1):41–50.
[20] Schreurs CA, de Boer MA, Heymans MW, Schoonmade LJ, Bossuyt PMM, Mol BWJ, et al. Prognostic accuracy of cerebroplacental ratio and middle cerebral artery Doppler for adverse perinatal outcomes: a systematic review and meta-analysis. Ultrasound Obstet Gynecol 2018;51(3):313–22.
[21] Conover CA, Bale LK, Overgaard MT, Johnstone EW, Laursen UH, Fuchtbauer EM, et al. Metalloproteinase pregnancy-associated plasma protein A is a critical growth regulatory factor during fetal development. Development 2004; 131(5):1187–94.
[22] Smith GC, Crossley JA, Aitken DA, Jenkins N, Lyall F, Cameron AD, et al. Circulating angiogenic factors in early pregnancy and the risk of preeclampsia, intrauterine growth restriction, spontaneous preterm birth, and stillbirth. Obstet Gynecol 2007;109(6):1316–24.
[23] Smith GC, Shah I, Crossley JA, Aitken DA, Pell JP, Nelson SM, et al. Pregnancy-associated plasma protein a and alpha-fetoprotein and prediction of adverse perinatal outcome. Obstet Gynecol 2006;107(1):161–6.
[24] Proctor LK, Toal M, Keating S, Chitayat D, Okun N, Windrim RC, et al. Placental size and the prediction of severe early-onset intrauterine growth restriction in women with low pregnancy-associated plasma protein-A. Ultrasound Obstet Gynecol 2009;34(3):274–82.
*[25] Cleaton MA, Dent CL, Howard M, Corish JA, Gutteridge I, Sovio U, et al. Fetus-derived DLK1 is required for maternal metabolic adaptations to pregnancy and is associated with fetal growth restriction. Nat Genet 2016;48(12):1473–80.
[26] Griffin M, Seed PT, Duckworth S, North R, Myers J, Mackillop L, et al. Prediction of delivering a small for gestational age infant and adverse perinatal outcome in women with suspected pre-eclampsia. Ultrasound Obstet Gynecol 2018;51(3):387–95.
*[27] Bricker L, Medley N, Pratt JJ. Routine ultrasound in late pregnancy (after 24 weeks' gestation). Cochrane Database Syst Rev 2015;6. CD001451.
*[28] Flenady V, Wojcieszek AM, Middleton P, Ellwood D, Erwich JJ, Coory M, et al. Stillbirths: recall to action in high-income countries. Lancet 2016;387(10019):691–702.
*[29] Monier I, Blondel B, Ego A, Kaminiski M, Goffinet F, Zeitlin J. Poor effectiveness of antenatal detection of fetal growth restriction and consequences for obstetric management and neonatal outcomes: a French national study. BJOG 2015; 122(4):518–27.

* Studies of particular importance to the current topic.

[30] Smith GC. Life-table analysis of the risk of perinatal death at term and post term in singleton pregnancies. Am J Obstet Gynecol 2001;184(3):489–96.
[31] National Collaborating Centre for Women's and Children's Health. NICE guideline: antenatal care. London: RCOG Press; 2008.
[32] Gardosi J, Francis A. Controlled trial of fundal height measurement plotted on customised antenatal growth charts. Br J Obstet Gynaecol 1999;106(4):309–17.
[33] Gardosi J, Giddings S, Clifford S, Wood L, Francis A. Association between reduced stillbirth rates in England and regional uptake of accreditation training in customised fetal growth assessment. BMJ Open 2013;3(12). e003942.
[34] Smith GCS. Prevention of stillbirth. Obstet Gynaecol 2015. https://doi.org/10.1111/tog.12197.
[35] Khalil AA, Morales-Rosello J, Morlando M, Hannan H, Bhide A, Papageorghiou A, et al. Is fetal cerebroplacental ratio an independent predictor of intrapartum fetal compromise and neonatal unit admission? Am J Obstet Gynecol 2015;213(1). 54–10.
[36] Valino N, Giunta G, Gallo DM, Akolekar R, Nicolaides KH. Biophysical and biochemical markers at 35-37 weeks' gestation in the prediction of adverse perinatal outcome. Ultrasound Obstet Gynecol 2016;47(2):203–9.
[37] Smith GC. Researching new methods of screening for adverse pregnancy outcome: lessons from pre-eclampsia. PLoS Med 2012;9(7). e1001274.
[38] Ganzevoort W, Alfirevic Z, von DP, Kenny L, Papageorghiou A, van Wassenaer-Leemhuis A, et al. STRIDER: sildenafil therapy in dismal prognosis early-onset intrauterine growth restriction–a protocol for a systematic review with individual participant data and aggregate data meta-analysis and trial sequential analysis. Syst Rev 2014;3:23.
*[39] Lees CC, Marlow N, van Wassenaer-Leemhuis A, Arabin B, Bilardo CM, Brezinka C, et al. 2 year neurodevelopmental and intermediate perinatal outcomes in infants with very preterm fetal growth restriction (TRUFFLE): a randomised trial. Lancet 2015;385(9983):2162–72.
*[40] Gulmezoglu AM, Crowther CA, Middleton P, Heatley E. Induction of labour for improving birth outcomes for women at or beyond term. Cochrane Database Syst Rev 2012;6. CD004945.
[41] Brocklehurst P, Field DJ, Juszczak E, Kenyon S, Linsell L, Newburn M, et al. The INFANT trial. Lancet 2017;390(10089):28.