UK National Screening Committee

Screening to prevent stillbirth

External review against programme appraisal criteria for the UK National Screening Committee

Version: Consultation document

Author: Costello Medical

Date: 8 February 2019

The UK National Screening Committee secretariat is hosted by Public Health England.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes.
Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK NSC, Floor 2, Zone B, Skipton House, 80 London Road, London SE1 6LH
www.gov.uk/uknsc
Twitter: @PHE_Screening Blog: phescreening.blog.gov.uk

For queries relating to this document, please contact: phe.screeninghelpdesk@nhs.net

Published Month 2018

Contents

Plain English summary

This is the fist time that the UK National Screening Committee looked at screening during pregnancy to prevent stillbirths.

A stillbirth is a baby born dead after 24 or more weeks of pregnancy. The rate of stillbirths in the UK has decreased a lot in the last 10 years, but 3,600 stillbirths still happen in the UK every year.

Some women are at higher risk of having a stillborn baby. There are many different causes of stillbirth, for example:

- illness in the mother,
- genetic problems in the baby,
- characteristics of the mother (for example obesity, smoking or stillbirth in a previous pregnancy) or problems in the placenta (the placenta a is an organ that grows in the womb during pregnancy. Its job is to provide oxygen and food to the baby and remove waste products from the baby's blood).

This review concentrates only on placenta problems because around 40% of stillbirths are due to such problems.

To prevent stillbirths it is important to know which pregnancies are at risk of it. But, there is currently no screening programme to asses the risk of stillbirth during pregnancy in the UK. This review looked if there are tests that can identify pregnancies at risk of stillbirth and what is the best way to follow up such pregnancies. It also explores what is the best thing to do to prevent stillbirth, and the costs of such actions on the baby.

This review concluded there is not enough evidence to recommend screening for stillbirth in the UK. This is because:

- there are no tests that can predict stillbirth accurately enough
- there is not enough evidence to say how best to monitor the women that are at risk of stillbirth
- there is not enough evidence to say that any treatments can reduce the risk of stillbirth when there is a problem with the placenta
- there is not enough evidence on how to plan the baby birth to avoid stillbirth, without risk for the baby.

Executive summary

Purpose of the review

This review was conducted to check whether a programme of routine screening for risk of stillbirth should be recommended.

Background

Nearly 3 million stillbirths occur annually worldwide, 98% of which are in low-income and middle-income countries.[8] In England, stillbirth rates have fallen to 4.4 per 1,000 total births in 2016; nevertheless, there are still over 3,000 stillbirths occurring in the UK every year.[9] and the UK's stillbirth rates are amongst the highest within high-income countries, with a stillbirth rate more than double that of the best-performing nation, Iceland. In 2015, the Department of Health announced a new national goal of "halving the rate of stillbirths, neonatal and maternal deaths in England by 2030, with a 20% reduction by 2020".[1-3] A proportion of stillbirths without anomalies are considered to be preventable with the provision of antenatal care that allows for accurate identification of at-risk pregnancies, careful monitoring, and timely applied interventions and delivery of the baby.[4-7] It is therefore important to identify the elements of antenatal care where evidence supports their implementation for prevention of stillbirth. The first step to stillbirth prevention is to identify pregnancies at risk; once identified, such pregnancies need to be monitored for signs of worsening to decide whether, and when, to intervene.

In support of the 2020 national goal set out by the NHS, which aims to reduce the number of stillbirths by 20%, the UK NSC have commissioned a rapid review to ascertain if there is sufficient evidence to support recommending a programme of screening low-risk pregnancies to prevent stillbirth caused by placental dysfunction. There are currently no screening programmes aimed at identifying women at risk of stillbirth due to placental insufficiency.

Focus of the review

The review focused on stillbirths caused by placental dysfunction, distinguishing between early (pre-term) and late (term) stillbirths where possible. The following uncertainties formed the basis of the review questions:

- Are there any effective tests to identify pregnancies at risk of pre-term or term stillbirths due to placental dysfunction, which are potentially preventable? (criterion 4)
- What strategies are most effective at monitoring the identified at-risk pregnancies, and should these be different depending on when in pregnancy the dysfunction is thought to have started? (criterion 7)

- Are there any interventions other than elective birth effective at preventing pre-term or term stillbirth? (criteria 9 and 10)
- How effective is elective birth at reducing stillbirth risk, including the influence of timing of delivery on risks arising from prematurity versus prevention of pre-term and term stillbirths? (criteria 9 and 10)

Recommendation under review

The UK NSC has not previously considered whether such programmes should be recommended, thus, no prior evidence reviews have been conducted and no recommendations are currently in place.

Findings and gaps in the evidence of this review

Within the scope of the review, 40 articles reporting on 39 unique studies were identified. A summary of question level results is presented below.

Criterion 4: 'There should be a simple, safe, precise and validated screening test.'

There were 27 articles relevant to this criterion. There were 26 primary publications reporting on 19 unique cohorts of women, and one SLR/meta-analysis. The majority of studies identified by this review evaluated the accuracy of screening tests for the prediction of all-cause stillbirth occurring at any gestational age (depending on the definition of stillbirth used). Based on the evidence, there are currently no tests that are appropriate for use in a screening programme aimed at predicting pregnancies at risk of pre-term or term stillbirth due to placental dysfunction in clinical practice. Therefore, criterion 4 is not met.

Criterion 7: 'There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.'

There were 3 studies that assessed monitoring regimes for pregnancies at high risk of stillbirth. The investigated monitoring strategies were electronic fetal heart rate monitoring (CTG), Doppler flow velocimetry, fetal movement counting and maternal serum AFP. Stillbirth was a considerably rare event, which increased the uncertainty around the outcome. The studies were not powered to detect a difference in stillbirth or severe SGA, and no studies differentiated between pre-term and term stillbirth. The limited evidence base for management strategies to prevent pre-term or term stillbirths does not allow for any conclusions to be drawn on the effectiveness of monitoring regimes in reducing the risk of stillbirth. Therefore, criterion 7 is not met.

Criterion 9: There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered'
Criterion 10: 'There should be agreed evidence-based policies covering which individuals should be offered interventions and the appropriate intervention to be offered'

Six studies reported in 7 articles reported on possible interventions for high-risk pregnancies.[8-13] Of the 6 relevant studies, one investigated unfractionated heparin (UFH),[10] 2 sildenafil citrate,[13, 14] 2 aspirin alone,[8, 11] and one compared aspirin and enoxaparin with aspirin alone.[9] Even among pregnancies at risk, stillbirth was a considerably rare event, which increased the uncertainty around the outcome. Based on the evidence found by this review, it is not possible to ascertain the effectiveness of interventions to prevent pre-term or term stillbirths or stillbirth overall. Without further studies, no intervention can be recommended as effective or preferable to elective birth.

This review also found 3 studies in 3 unique cohorts that reported on the risk of stillbirth upon induction of labour compared with expectant management.[15-17] Stillbirths were only reported in one of the 3 identified studies and it appears that induction of labour may be beneficial for preventing pre-term but not term stillbirths. However, the poor quality of that study precludes drawing any definite conclusions. Due to the poor quality and targeted scope of the evidence considered in this review, the effectiveness and safety of induced delivery for the prevention of pre-term or term stillbirth in screen-detected high-risk pregnancies cannot currently be ascertained. Therefore, criteria 9 and 10 are not met.

Recommendations on screening

Based on the synthesis of evidence against the UK NSC criteria, screening of pregnant women to prevent stillbirths due to placental dysfunction is not recommended.

Overall, a moderate number of good-quality studies have been identified, including some conducted in the UK, but none of the screening tests examined were sensitive and specific enough to be recommended for use in a screening programme using the thresholds that are widely used for evaluating the accuracy of diagnostic tests (LR+ >10 and LR− <0.1). Further, there was insufficient good-quality evidence on appropriate monitoring strategies and interventions that could be used in screen-identified high-risk pregnancies to prevent stillbirth; this includes very limited, evidence on effectiveness of elective birth in preventing pre-term or term stillbirths.

Limitations

There is an inherent risk of bias in the measurement of test accuracy to predicit risk of stillbirth in studies where women and treating obstetricians were not blinded to the interpretation of test results. In an effort to minimised this, this review used a combined reference standard where babies could have been either stillborn or liveborn with severe SGA, i.e. <3rd centile for gestational age. Severe SGA was also considered a relevant outcome in studies of monitoring and interventions. It is noted that this could also lead to an overestimation of test accuracy/intervention efficacy, as only a proportion of severe SGA babies will represent "prevented" stillbirths; nevertheless, no studies were found where the reference standard included both stillborn and liveborn babies.

This review did not find sufficient evidence to support any changes to the recommendation for elective birth, which is the only intervention currently used in clinical practice to prevent stillbirths in pregnancies that would be considered at-risk by this review.[18] It is however noted, that while not directly applicable to women at risk of stillbirth, there is evidence to support that elective birth is safe at term. Therefore, although this review only included women at risk of stillbirth, it is recognised that delivery has been shown to reduce the risk of both perinatal death and stillbirth in the unselected population.[19] While these observations do not alter the primary conclusion that screening and intervention for stillbirth is not currently justified, they do suggest that future research might focus on screening and intervention to prevent stillbirths at term, where there is an intervention that is effective in other contexts and would not be expected to lead to harm through iatrogenic prematurity.

Methodological limitations included limiting the searches to records published since 2000, and only including peer-reviewed, English-language journal articles. The titles, abstracts and full texts were screened by one reviewer, with a second reviewer verifying all included, 10% of excluded decisions and any articles where there was uncertainty about their inclusion.

Evidence uncertainties

Given that growth monitoring is considered to be more specific for identifying pregnancies at risk of pre-term stillbirth, there is a need for high-quality, prospective studies investigating growth potential as a screening test, but these studies would need to report results separately for pre-term and term stillbirths. Data on whether the interventions helped to avoid elective birth were also lacking as were studies of elective birth in screen-detected high-risk pregnancies. Finally, there is uncertainty around the accuracy of the screening test in studies without test result blinding, and as such, studies utilising different approaches to measuring test performance may be more informative than traditional measures of predictive values or likelihood ratios.

Introduction and approach

Background

In the UK, stillbirth is defined through the Stillbirth (Definition) Act 1992 as 'a baby who has been delivered with no signs of life and who is known to have died after the 24th week of pregnancy'.[20] This definition is used by the Confidential Enquiry into Maternal and Child Health (CEMACH), the Royal College of Obstetricians and Gynaecologists (RCOG) and the Office for National Statistics (ONS). However, there is considerable heterogeneity in the definitions used between and within countries, which can range from a fetal death occurring at or after 20 weeks of gestation[21-23] to a fetal death occurring at or after 28 weeks of gestation (definition used by the World Health Organization [WHO]).[24] This makes a consistent evidence synthesis difficult, especially for early gestation (<28 weeks) stillbirths.[25, 26]

Importance of stillbirth

Nearly 3 million stillbirths occur annually worldwide, 98% of which are in low-income and middle-income countries.[4] However, even between and within high-income countries, stillbirth rates vary widely.[25] In England, stillbirth rates have fallen from 5.7 per 1,000 total births in 2004 to 4.4 per 1,000 total births in 2016; nevertheless, there are still over 3,000 stillbirths occurring in the UK every year.[27] In fact, the UK's stillbirth rates are amongst the highest within high-income countries, with a stillbirth rate more than double that of the best-performing nation, Iceland. There is also a 25% variation in the stillbirth rates across different English regions.[1, 27]

In 2015, the Department of Health announced a new national goal of "halving the rate of stillbirths, neonatal and maternal deaths in England by 2030, with a 20% reduction by 2020".[1-3] The increased rate compared with other countries and variation among English regions indicates that a lower stillbirth rate is achievable. A proportion of stillbirths without anomalies are considered to be preventable with the provision of antenatal care that allows for accurate identification of at-risk pregnancies, careful monitoring, and timely applied interventions and delivery of the baby.[4-7] It is therefore important to identify the elements of antenatal care where evidence supports their implementation for prevention of stillbirth.

Causes of stillbirth

Stillbirth is associated with a range of causes, such as maternal conditions, infections, autoimmune disorders, chromosomal or structural abnormalities, problems with placentation or placental function.[26] There are many different systems for classifying stillbirths by cause, further

complicating the synthesis of evidence from different countries.[4, 28] The current review will focus on stillbirths caused by placental dysfunction, which are estimated to account for ~40% of stillbirths in high-income countries.[25] Poor placental function, which disrupts the delivery of nutrients and oxygen to the fetus, can lead to growth restriction, placental abruption, pre-eclampsia or hypoxaemia, all of which are known to put the baby at risk of stillbirth.[29-31]

Before term, the nutritional demands of the fetus increase exponentially to support growth and development.[32, 33] Early onset of placental dysfunction can therefore lead to fetal growth restriction (FGR), caused by the limited supply of nutrients to the developing baby.[30, 32, 34] By contrast, at and beyond term, the respiratory demands of the fetus exponentially increase (while nutritional demands plateau). Therefore, a late onset of placental dysfunction can restrict the delivery of oxygen to the fetus, increasing the risk of stillbirth due to hypoxaemia, but is unlikely to affect fetal growth.[32, 33] FGR is therefore considered to be most characteristic of early-onset placental dysfunction, and poses a serious risk if further action to deliver the fetus is not taken.[1] Conversely, stillbirths at and beyond term are less likely to be associated with FGR.

Due to the different outcomes based on the timing of placental dysfunction onset, it has been suggested that stillbirth could be subdivided into pre-term (before 37 completed weeks of gestation) and term (past 37 completed weeks of gestation), which may have implications for incidence, diagnostic accuracy, and effectiveness and timing of interventions. The choice of 37 weeks relies on the delivery before 37 completed weeks of gestation (pre-term) being associated with a high risk of infant mortality; over 70% of neonatal deaths occur in babies born before 37 weeks.[35] Babies born pre-term also have an increased risk of major and minor morbidities and chronic diseases in later life.[35] These risks decrease sharply when the baby is born after 37 weeks of gestation; one study demonstrated that rates of neonatal intensive care unit admission, length of stay and neonatal morbidities were significantly lower in births at term, compared with late pre-term births (34 to 36 weeks of gestation).[36] These are important implications to consider when evaluating screening programmes to prevent stillbirth arising from placental dysfunction; for example, pregnancies at risk of pre-term stillbirth may necessitate pre-term delivery but those at risk of term stillbirth can be delayed to avoid the risks that are associated with prematurity.

SGA and FGR: definitions, relation to each other and power to predict stillbirth

It is claimed that FGR is the single largest risk factor for stillbirth in normally formed fetuses and that many stillbirths in normally formed fetuses are due to FGR.[37] These stillbirths could potentially be avoidable if FGR is antenatally detected, and if the baby is mature enough to be delivered.[5] Thus, measures of restricted fetal growth are considered an indication of placental dysfunction, though diagnostic issues exist in determining whether a fetus is growth restricted or constitutionally small. One indication of FGR could be small for gestational age (SGA), defined most often as a baby with a birthweight in the lowest 10% of the norm, the latter often based on population-

adjusted growth curves or charts customised for maternal characteristics.[38] However, FGR is not synonymous with SGA. A fetus can be constitutionally small, and it is estimated that 50–70% of SGA fetuses have appropriate growth according to maternal size and ethnicity.[18] Conversely, a baby can be growth restricted while also being appropriate for gestational age.[38] By definition, 10% of babies will be SGA in an average pregnancy cohort, and although SGA babies are at an increased risk of perinatal mortality and morbidity, most adverse outcomes occur in the FGR group.[38]

It appears that identifying SGA has a low predictive power for stillbirth,[39-41] though the likelihood of FGR is higher in severe SGA (<3rd centile estimated fetal weight [EFW] or abdominal circumference [AC]) infants.[38] A prospective cohort study reported that universal screening for SGA in the third trimester had a much higher sensitivity in detecting severe SGA (likely to be FGR) and predicting neonatal morbidity than selective screening (by indication),[42] but it remains unclear how informative detection of FGR through SGA or even severe SGA would be for prediction of stillbirth.[38, 42]

One critique to approaches of detecting FGR based on SGA is that a baby could be within normal limits but not realising its full growth potential; instead, serial growth measurements assessed against a predicted customised slope of a growth curve have been proposed, in order to identify when a baby is not growing at a predicted rate.[43] Thus, growth monitoring strategies could identify the pregnancies with FGR that are at a high risk of stillbirth and potentially preventable with appropriately timed delivery or other interventions.[5]

It is also important to consider that FGR is most often an indication of early-onset placental dysfunction; pregnancies with late-onset placental dysfunction may not be SGA or growth-restricted,[31] making it increasingly difficult to identify those at risk of term stillbirths with the use of growth curves alone. As such, alternative means of identifying placental dysfunction that do not solely rely on detection of growth restriction should also be considered. Tests based on biochemical markers or various parameters measured via Doppler ultrasound may be useful for detecting other signs of placental dysfunction and more appropriate for identifying those at risk of term stillbirth.[44] A systematic literature review (SLR) published in 2015 found that in the first or second trimester there were no accurate tests for predicting stillbirth, but two tests (uterine artery pulsatility index [UtA-PI] and maternal serum pregnancy associated plasma protein [PAPP-A] levels) appeared to be good predictors of stillbirths related to placental dysfunction.[44]

Risk factors for stillbirths and SGA

In addition to tests, numerous risk factors for stillbirth, including maternal characteristics, health conditions and previous pregnancy complications, may be helpful to identify pregnant women who may be at risk of stillbirth. Pre-existing diabetes, hypertension, history of stillbirth, advanced

maternal age, primiparity, smoking and increased body mass index (BMI)/obesity have each been found to be associated with an increased risk of stillbirth, either through individual studies or meta-analyses.[37, 45-50] The strength of association is moderate (odds ratio [OR] 2 to 5) for diabetes, hypertension and prior stillbirth, but weak (OR 1 to 2) for the other individual factors, and the uncertainty around the associations is small (narrow confidence intervals).[37, 45-50] However, it is unclear how combinations of these minor risk factors, which have only a weak association with stillbirth, influence the risk of stillbirth in an individual pregnancy. This has implications for screening and defining at-risk populations, as it is unlikely that any single risk factor would have sufficient predictive power; instead, risk may need to be estimated with algorithms that consider weighted contributions from all the different factors. It is also unclear whether and how these factors influence the timing of placental dysfunction onset and thus differ in their predictive power for pre-term and term stillbirths. Maternal conditions, characteristics and previous pregnancy history that put the pregnancy at risk of stillbirth are also (among others) risk factors for an SGA baby,[38] possibly because of the association between FGR and SGA. However, as with predicting stillbirth, it is not clear how combinations of these factors influence the risk of FGR or SGA in an individual pregnancy.[38]

Current clinical practice

There is currently no single guidance document covering all aspects of placental insufficiency, SGA and stillbirth in UK clinical practice; instead, recommendations are covered by a number of NICE guidelines (CG62, CG70, CG190)[51-53] and RCOG guidelines (GreenTop55, GreenTop27, GreenTop31),[18, 54, 55] as well as the Saving Babies' Lives care bundle developed by NHS England.[56] Guidance on antenatal care for uncomplicated pregnancies is presented in NICE CG62 and recognises that certain pregnancies (including in women who smoke, with pre-existing diabetes or hypertension, history of stillbirth, a BMI ≥30 kg/m^2 or those who are ≥40 years of age) may need additional care. The RCOG GreenTop31 guideline also identifies a number of risk factors for having an SGA baby: pregnancy conceived through in vitro fertilisation (IVF), nulliparity, BMI 25 to 30 kg/m^2, pregnancy interval <6 or >60 months, paternal or maternal SGA, or doing daily vigorous exercise.[18] In such at-risk pregnancies, the RCOG also recommend referral for assessment at 20 weeks to determine fetal growth.[18]

Recommendations to reduce stillbirth

A number of specific recommendations are in place in the UK to reduce stillbirth across various guidelines and programmes. For the NHS, reducing stillbirth is a priority; the specific recommendations are published in the Saving Babies' Lives care bundle and can be summarised as:[56]

- Reducing smoking in pregnancy

- Risk assessment and surveillance for FGR (using the RCOG algorithm or the algorithm published in Saving Babies' Lives to assess initial risk, both highly similar)
- Raising awareness of reduced fetal movement
- Effective fetal monitoring during labour

In addition to the Saving Babies' Lives care bundle designed to tackle stillbirth, the Growth Assessment Programme (GAP), introduced by the Perinatal Institute in 2008, is aimed at increasing the antenatal detection of problems with fetal growth.[57] The programme consists of GROW accreditation workshops, which include training in the assessment of risk factors, standardised fundal height measurement, plotting on customised charts, and evidence-based referral pathways and protocols based on the RCOG guidelines.[57] Stillbirth rates decreased in the 3 regions with high uptake of GROW, with the decrease being specifically among stillbirths due to FGR.[57] Based on these reports, an enhanced GAP was introduced in 2013, with focus on training, identifying training needs through audits, evidence-based protocols, monitoring of IUGR and detection rates and cooperation between the Perinatal Institute and specific Trusts.[57] Evaluation of the GAP programme in 15 UK NHS trusts concluded that while the antenatal detection of SGA babies progressively increased over its implementation from 2016 to 2018, there is insufficient direct evidence to infer that the GAP programme directly increased SGA prediction.[58] Nevertheless, the authors note it is plausible that improved monitoring of fetal growth using serial fetal growth measurements (as used in the GAP programme) could improve detection of SGA babies before birth.[58]

Identification and monitoring recommendations

The first step to stillbirth prevention is to identify pregnancies at risk; once identified, such pregnancies need to be monitored for signs of worsening to decide whether, and when, to intervene. There are currently no screening programmes specifically aimed at identifying women at risk of stillbirth due to placental insufficiency. Nevertheless, existing recommendations based around identification and monitoring of SGA pregnancies are similar across all relevant UK guidelines. Antenatal detection of FGR has been shown to be associated with reduced stillbirth rates in a large cohort of singleton pregnancies.[37] According to both NICE and RCOG, the current recommendation for identification of SGA fetuses in low-risk pregnancies is a standardised serial measurement of symphysis-fundal height (SFH) plotted on customised height charts from 24 weeks onwards.[18, 51] Effectiveness of SFH for preventing stillbirth is unclear; a Cochrane review of SFH found only one study, and concluded that there was insufficient evidence to evaluate the use of SFH measurement in antenatal care for reducing stillbirth.[59] For pregnancies identified with SFH <10th centile, monitoring growth by ultrasound is recommended by GreenTop31 at each antenatal appointment from 24 weeks onwards and fetal growth should be assessed with serial ultrasound throughout the third trimester if the increased risk persists.[18] In terms of other tests, the CG62 guideline does not recommend routine Doppler ultrasound or ultrasound scanning after 24 weeks

in low-risk pregnancies.[51] Similarly, formal fetal movement-counting and cardiotocography should not be offered in low-risk pregnancies.[51] However, umbilical artery Doppler is recommended as the primary surveillance tool in the SGA fetus.[18]

Intervention recommendations

It is currently unclear which, if any, interventions are effective at reducing the risk of stillbirth in general, or late and early stillbirth specifically. Other than steroids for planned earlier/caesarean (CS) delivery and early aspirin for women at risk of pre-eclampsia, no pharmaceutical interventions are currently recommended,[18] likely because no interventions effective at decreasing the risk of stillbirth due to placental insufficiency are known.[29, 60] However, antenatal detection of SGA babies and FGR is strongly recommended as a means to reduce stillbirth risk as this gives the option to consider earlier delivery.[18, 37, 57]

Currently, the only intervention in place to reduce the risk of stillbirth associated with FGR is the induction of labour.[18, 56] The RCOG guideline for delivery of FGR babies recommends planned delivery by 37 weeks, the exact timing depending on the perceived threat from FGR and risks arising from prematurity based on ultrasound measurements.[18] Up to 29 weeks, intact survival of the baby mostly depends on gestational age. Delivery may be indicated depending on prematurity risk evaluation or surveillance tests for fetal distress. For example, in fetuses with umbilical artery absent/reversed end diastolic velocities (AREDV), who are predicted to be viable, delivery should be scheduled once DV Doppler is abnormal or UV pulsations appear, but it is also recommended to deliver before 32 weeks, even if DV Doppler is normal.[18] If SGA is detected after 32 weeks, delivery should occur no later than 37 weeks if umbilical artery Doppler is abnormal or should be offered at 37 weeks, and a senior obstetrician should determine the time and mode of birth.[18]

Planned CS is not recommended for SGA, unless umbilical artery AREDV is present,[18] and neither the NICE CG190 (intrapartum care) nor the CG132 (caesarean section) guidelines include recommendations for how to deliver FGR babies.[53, 61] Nonetheless, it should be noted that the frequency of caesarean sections has risen in high-income countries, partly due to increasing concern over stillbirth, and it is unclear if recommendations for elective birth would be likely to cause these to increase further.[4]

Questions and uncertainties

There are various guidelines and recommendations in place covering different aspects of placental insufficiency, SGA and stillbirth in UK clinical practice. In support of the 2020 national goal set out by the NHS which aims to reduce the number of stillbirths by 20%, the UK NSC have commissioned a rapid review to ascertain if there is sufficient evidence to support recommending a programme of screening low-risk pregnancies to prevent stillbirth caused by placental dysfunction.

The following uncertainties will form the basis of the review questions:

- Are there any effective tests to identify pregnancies at risk of pre-term or term stillbirths due to placental dysfunction, which are potentially preventable?
- What strategies are most effective at monitoring the identified at-risk pregnancies, and should these be different depending on when in pregnancy the dysfunction is thought to have started?
- Are there any interventions other than elective birth effective at preventing pre-term or term stillbirth?
- How effective is elective birth at reducing stillbirth risk, including the influence of timing of delivery on risks arising from prematurity versus prevention of pre-term and term stillbirths?

The review will focus on stillbirths caused by placental dysfunction, distinguishing between early (pre-term) and late (term) stillbirths where possible.

Current policy context and previous reviews

There are currently no programmes screening for risk of stillbirth in pregnancy. The UK NSC has not previously considered whether such programmes should be recommended, thus, no prior evidence reviews have been conducted.

Objectives

This review aims to assess whether there is sufficient evidence to consider introducing a screening programme for stillbirth. The review will appraise evidence on the questions in Table 1, which each relate to the criteria set out by the UK NSC for assessing the suitability of a screening programme.

Table 1. Key questions for the evidence summary, and relationship to UK NSC screening criteria

	Criterion	Key questions	Studies included
	THE TEST		
4 and 5	• There should be a simple, safe, precise and validated screening test. • The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	Is there an effective test to predict the risk of stillbirth?[a] a. Before 37 completed weeks of gestation b. From 37 completed weeks of gestation	27
	THE SCREENING PROGRAMME		

	Criterion	Key questions	Studies included
11	There should be evidence from high quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an "informed choice" (eg. Down's syndrome, cystic fibrosis carrier screening), there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.	What is the appropriate monitoring regime for pregnancies that have been identified by screening to be at risk of stillbirth? a. Before 37 completed weeks of gestation b. From 37 completed weeks of gestation	3
	THE INTERVENTION		
9 and 10	• There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered. • There should be agreed evidence-based policies covering which individuals should be offered interventions and the appropriate intervention to be offered.	Are there any effective interventions to prevent stillbirth in women identified as high risk through screening that is not elective birth? a. Before 37 completed weeks of gestation b. From 37 completed weeks of gestation	6 (7 articles)
		How effective is elective caesarean section (CS) or induction of labour to prevent stillbirth in pregnancies at risk? a. Before 37 completed weeks of gestation b. From 37 completed weeks of gestation	3

[a] A combined reference standard where babies could have been stillborn or born with SGA/FGR <3rd centile was used, to account for stillbirths that may have been avoided due to knowledge of the test result.

Methods

The current review was conducted by Costello Medical in collaboration with the UK NSC, in keeping with the UK NSC evidence review process. The search strategy is presented in Appendix 1, and methods of study selection (including full eligibility criteria and quality assessment checklists used) are detailed below.

Eligibility for inclusion in the review

The following review process was followed:

1. Each abstract was reviewed against the inclusion/exclusion criteria by one reviewer. Where the applicability of the inclusion criteria was unclear, the article was included at this stage in order to ensure that all potentially relevant studies were captured. A second independent reviewer provided input in cases of uncertainty and validated 10% of the first reviewer's excluded decisions and all of the included decisions. Any disagreements were resolved by discussion until a consensus was met.
2. Full-text articles required for the full-text review stage were acquired if freely available at the Cambridge University Library. For any paywalled articles unavailable at the Cambridge University Library, the abstracts were re-reviewed to determine the probability that the full text may be highly relevant to the review questions. Authors of any highly relevant articles were contacted to provide the full texts and any articles that were not obtained from the authors within 2 weeks were purchased.
3. Each full-text article was reviewed against the inclusion/exclusion criteria by one reviewer, who determined whether the article was relevant to one or more of the review questions. A second independent reviewer provided input in cases of uncertainty and validated 10% of the first reviewer's excluded decisions and all of the included decisions. Any disagreements were resolved by discussion until a consensus was met.

Eligibility criteria for each question are presented in Table 2 below. For all questions, systematic literature reviews (SLRs) and meta-analyses (MAs) were considered for inclusion. If the scope of a SLR or MA was very closely aligned to one of the topics of this review, it was included in its own right. However, where the scope was not closely aligned to one of the topics of this review but some of the included articles may have been of interest, the reference list of the SLR or MA was hand-searched. Any relevant primary research articles identified that were relevant were included, but the SLR itself was excluded.

Table 2. Inclusion and exclusion criteria for question 1 — screening for stillbirth

Domain	Target condition	Population	Intervention	Outcome	Study type	Setting	Other considerations
Inclusion criteria	Stillbirth	Unselected or low-risk pregnant women[a]	Test to predict stillbirth before or after 37 completed weeks of gestation. Index test: • Combined risk calculated from at least two maternal risk factors • Biochemical markers • Ultrasound and Doppler markers of placental dysfunction • Fetal size or growth • Combinations of these Reference standard: • Stillbirth, or livebirth with SGA<3rd centile[b]	Measures of screening accuracy: • Sensitivity • Specificity • Positive predictive value • Negative predictive value • Accuracy • Likelihood ratio	Cross-sectional studies, cohort studies, case-control studies, systematic reviews RCTs and interventional studies will be included if the intervention is the screening test and appropriate measures of screening accuracy are reported	Studies conducted in the UK Studies conducted in high-income countries where the screening methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)	Peer-reviewed studies in the English language
Exclusion criteria		Studies only including women with high-risk pregnancies who would be outside the expected screening population, such as women who have: • Previously had a stillbirth • Had recurrent miscarriages • Pre-existing diabetes • Pre-existing	Studies where the reference standard is SGA under the 5th or 10th centile	Any other outcomes (including area under the receiver-operator curve or measures of association between risk factors/test values and risk of stillbirth)	Case reports, case series, narrative reviews, editorials, commentaries, letters, conference abstracts or other publication types that have not been peer-reviewed	Studies in ineligible countries, or international studies that consider eligible and ineligible countries, but outcomes for eligible countries are not presented separately to outcomes from ineligible	Studies with full text not in English language

		hypertension, eclampsia or pre-eclampsia • Multiple pregnancies • Pregnancies with known chromosomal or structural abnormalities • Autoimmune disorders • Rhesus isoimmunisation or other significant blood group antibodies • HIV or hepatitis B virus (HBV) infection • Infections such as toxoplasmosis, rubella, cytomegalovirus, herpes simplex, syphilis, chorioamnionitis • Women >13 days beyond their due date				countries	

[a] Women outside the CG 62 NICE guidance are considered at high-risk;1 [b] A combined reference standard where babies could have been either stillborn or liveborn with SGA<3rd centile counters the intervention bias of scheduled birth, which could preclude stillbirth in a high-risk individual

Abbreviations: EEA: European Economic Area; OECD: Organisation for Economic Co-operation and Development; RCTs: randomised controlled trials; SGA: small for gestational age.

Table 3. Inclusion and exclusion criteria for question 2 — monitoring pregnancies at risk of stillbirth

Domain	Target condition	Population	Intervention/ comparators	Outcome	Study type	Setting	Other considerations
Inclusion criteria	Stillbirth	Women at risk of stillbirth detected through screening or	**Intervention** • Monitoring regime	• Stillbirth (>24 weeks) (early or late)	RCTs, interventional studies, cohort studies, case-	• Studies conducted in the UK • Studies conducted in high-income countries	Peer-reviewed studies in the English language

	otherwise[a]	• Fetal movements • Fetal growth • Doppler flow velocimetry (utero-placental and fetal) • CTG (computerised and analogue) • Biophysical profile **Comparator** • Normal care	• Livebirth with SGA<3rd centile	control studies, systematic reviews	where the screening methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)	
Exclusion criteria	Women not at risk of stillbirth Women at risk of stillbirth as part of a larger cohort where results for the women of interest are not reported separately	Interventions not aimed at monitoring pregnancies at risk of stillbirth	Studies not reporting any outcomes of interest. Specifically, outcome of SGA under 5th or 10th centile will not be appropriate	Cross-sectional studies, case reports, case series, narrative reviews, editorials, commentaries, letters, conference abstracts or other publication types that have not been peer-reviewed	Studies in non-eligible countries or international studies that consider eligible and non-eligible countries, but outcomes for eligible countries are not presented separately to outcomes from non-eligible countries	Studies with full text not in the English language

[a] Studies where the risk of stillbirth is determined other than through screening (i.e. through any of the tests specified in Table 2, but not as part of a screening programme) were only be considered as relevant if insufficient studies in screen-predicted stillbirth are found

Abbreviations: CTG: computer tomography; EEA: European Economic Area; OECD: Organisation for Economic Co-operation and Development; RCTs: randomised controlled trials; SGA: small for gestational age.

Table 4. Inclusion and exclusion criteria for questions 3 and 4 — interventions to prevent stillbirth

Domain	Target condition	Population	Interventions/ comparators	Outcome	Study type	Setting	Other considerations
Inclusion criteria	Stillbirth	Women at risk of stillbirth	Interventions directed to prevent stillbirth or SGA <3rd centile (that are placental in origin), including but not limited	Risk of stillbirth or livebirth <3rd centile, reported as: • Absolute risk • Risk ratio	RCTs, interventional studies, cohort studies, case-control studies, systematic reviews	Studies conducted in the UK Studies conducted in high-income countries where the screening	Studies in the English language

		to: • Anti-platelet agents • Aspirin • Dalteparin • Anti-coagulant agents (such as Heparin (low-molecular weight, unfractionated), Dipyridamole, Tinzaparin, Enoxaparin and Nadroparin) • Anti-hypertensive agents (such as Beta-blockers, Nitric oxide, Labetalol, Hydralazine and Calcium supplementation) • Planned delivery Comparators: • Any or none	• (RR) • Odds ratio (OR) Any other relevant measure of risk		methods and technology are expected to be similar to that of the UK (OECD and EEA countries excluding South Korea and Mexico)	
Exclusion criteria	Women not at risk of stillbirth Women at risk of stillbirth as part of a larger cohort where results for the women of interest are not reported separately	Interventions aimed at preventing stillbirth where the cause of stillbirth is not placental in origin, such as: • Anti-retroviral therapies and strategies to prevent stillbirth related to HIV/AIDS • Antibiotics • Anti-malarial	Studies not reporting any outcomes of interest	Cross-sectional studies, case reports, case series, narrative reviews, editorials, commentaries, letters, conference abstracts or other publication types that have not been peer-reviewed	Studies in non-eligible countries or international studies that consider eligible and non-eligible countries, but outcomes for eligible countries are not presented separately to outcomes from non-eligible countries	Studies with full text not in the English language

		medication and strategies to prevent malaria • Anti-helminthic medication • Preventive strategies aimed at reducing life-style risk factors for stillbirth (smoking, obesity) • Strategies to prevent pre-term labour e.g. flunarizine	

Abbreviations: EEA: European Economic Area; OECD: Organisation for Economic Co-operation and Development; OR: odds ratio; RCTs: randomised controlled trials; RR: risk ratio; SGA: small for gestational age.

Risk of bias considerations

In studies of stillbirth, knowledge of the test result indicating that the pregnancy is at risk would likely trigger efforts aimed at preventing negative outcomes, including stillbirth. Pregnancies where stillbirth is avoided may appear as false-positives instead of true-positives, decreasing sensitivity and specificity of the test. Therefore, any unblinded screening studies (where the index test result is known prior to delivery) are at a high risk of bias. Equally, in studies of monitoring, knowledge that the pregnancy is at risk may prompt the healthcare provider to intervene, regardless of the monitoring arm, again resulting in ascertainment bias. In an effort to minimise this, a combined reference standard where babies could have been stillborn or liveborn with severe SGA, i.e. <3rd centile was used. This should capture pregnancies with severe FGR, where knowledge of the test result or the pregnancy being at risk may have caused an earlier delivery, thereby preventing stillbirth.

Appraisal for quality/risk of bias tool

The following tools were used to assess the quality and risk of bias of each study included in the review

- Diagnostic accuracy studies: Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[62]
- Observational and interventional studies: adapted Downs and Black Checklist[63]
- SLRs: AMSTAR checklist[64]

The full guidance used for the quality assessments is available in Table 28 to Table 30; Appendix 4.

Databases/sources searched

The following databases were searched:

- MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print
- Embase
- The Cochrane Library, including the following:
    - Cochrane Database of Systematic Reviews (CDSR)
    - Cochrane Central Register of Controlled Trials (CENTRAL)
    - Database of Abstracts of Reviews of Effects (DARE)

Searches were conducted in June 2018. Full details of the searches, including the search strategy for each database, are presented in Appendix 1.

Overall results

Database searches yielded 6,069 results, of which 38 articles were judged to be relevant to one or more questions. An additional 2 references were identified through hand-searching reference lists, so 40 articles were ultimately included.

Appendix 2 contains a full PRISMA flow diagram (Figure 5), along with a table of the included publications and details of which questions these publications were identified as being relevant for (Table 19).

Question level synthesis

Criterion 4 – Screening for stillbirth

4: 'There should be a simple, safe, precise and validated screening test.'

Question 1 – Is there an effective test to predict the risk of stillbirth?
a) Before 37 completed weeks of gestation
b) From 37 completed weeks of gestation
No prior evidence reviews have been conducted to support the decision of whether a programme of screening to prevent stillbirth should be recommended by the UK NSC.

As recommended by the NICE CG62 antenatal care guideline, low-risk pregnant women in the UK currently undergo growth measurement with methods such as abdominal palpitation, SFH measurements, ultrasound scanning, fetal biometry and customised growth charts to identify SGA or large for gestational age (LGA) babies, as well as various screening tests to detect genetic diseases and other abnormalities.[35] The aim of this question was to identify and synthesise evidence published since 2000 on accuracy parameters of tests that predict pre-term (before 37 completed weeks of gestation) or term stillbirth (from 37 completed weeks) due to placental dysfunction in UK women, or women similar to a low-risk or unselected UK population.

Eligibility for inclusion in the review

This review searched for cohort, cross-sectional and case-control studies, as well as SLRs or MAs of the above, for screening tests that can be used to predict stillbirth. Randomised controlled trials (RCTs) and interventional studies were included if the intervention was the screening test and appropriate measures of screening accuracy were reported. Studies had to assess the performance of a test that aimed to predict the risk of stillbirth (before or from 37 weeks of gestation). In order to avoid ascertainment bias in unblinded studies, where the knowledge of the test result would prompt action preventing the outcome, a combined reference standard where babies could have been stillborn or liveborn with SGA <3rd centile was used. Only studies directly reporting test accuracy parameters were included; no calculations were performed in this review to obtain measures of test accuracy.

The eligible population was unselected or low-risk pregnant women, as specified by the NICE GC62 guideline for antenatal care.[51] Studies that only included women with high-risk pregnancies such as those who had previous complications or maternal comorbidities were excluded. These women would already be identified as being at risk through existing antenatal care pathways and

receive antenatal care under different guidelines than the rest of the pregnant population, and would therefore fall outside of the expected screening population.

Full details of the eligibility criteria are presented in Table 2.

Description of the evidence

There were 27 articles relevant to this question. There were 26 primary publications reporting on 19 unique cohorts of women, and one SLR/meta-analysis.

The majority of the identified studies were of a prospective cohort design (N=14),[50, 65-77] with 10 retrospective cohort or database studies,[40, 78-86] two case-control studies,[69, 87] and one case-cohort study.[88] Chaiworapongsa 2013 reported data for a prospective cohort study and a case-control study in two separate cohorts, which are considered as two separate studies in this review.[69] Most studies were conducted in the UK (N=15)[40, 50, 65-68, 73, 74, 76, 77, 80-83, 86] and the USA (N=9),[70-72, 79, 84, 85, 87, 88] with the remaining studies conducted in Canada (N=1),[75] Chile (N=1),[88] and Finland (N=1).[78] In particular, 8 of the included articles report data from a large prospective screening programme in the UK, which recruited women with singleton pregnancies attending routine antenatal care at the King's College Hospital and Medway Maritime Hospital (with a small number of studies additionally recruiting women from University College London Hospital).[50, 65-68, 73, 76, 77] Whilst these 8 publications reported screening with different biochemical or ultrasound markers at different gestational ages, they are likely to have been conducted in the same cohort of pregnant women, or in cohorts with substantial cross-over. Similarly, three publications report on a cohort of women enrolled in in the FaSTER Trial (First- and Second-Trimester Evaluation of Risk), a multicentre, prospective study on combined ultrasound and biochemical screening for Down syndrome.[70, 71, 89]

The majority of identified studies evaluated the accuracy of screening tests for the prediction of all-cause stillbirth at any gestational age, depending on the definition of stillbirth used. Only 5 studies further distinguished between pre-term and term stillbirths, or by gestational age. [40, 74, 75, 82, 86] Four articles reporting on one cohort additionally reported measures of test accuracy using a reference standard of 'stillbirths arising from impaired placentation'.[65-67, 73]

Biochemical and ultrasound markers for the prediction of stillbirth

The articles included in this review reported measures of test accuracy for the prediction of stillbirth using 9 distinct biochemical markers or 6 ultrasound-based markers, used in isolation or as combinations of markers, maternal factors or fetal biometry:

- Biochemical markers
    - alpha fetal protein (AFP)

  - free beta human chorionic gonadotropin (β HCG)
  - inhibin A
  - pregnancy associated plasma protein A (PAPP-A)
  - placental growth factor (PlGF)
  - soluble endoglin (sEng)
  - soluble fms-like tyrosine kinase 1 (sFlt-1)
  - soluble vascular endothelial growth factor receptor-1 (sVEGFR-1)
  - unconjugated oestriol
- Ultrasound-based markers
  - uterine artery pulsatility index (UT-PI/UtA-PI)
  - uterine artery resistance index (UtA-RI)
  - UtA notching
  - ductus venosus Doppler (reversed A wave)
  - ductus venosus pulsatility index for veins (DV-PIV)
  - nuchal translucency (NT)

Eleven of the identified publications examined combinations of maternal characteristics, biochemical and ultrasound markers for the identification of pregnancies at risk of stillbirth,[65-67, 72, 73, 76, 77, 90] or specific ratios of biomarkers such as the cerebroplacental ratio (middle cerebral artery pulsatility index [MCA-P]/UT-PI)[68] and angiogenic index-1 (PlGF/sVEGFR-1 ratio).[69, 88] Six studies combined individual maternal characteristics only.[50, 65, 66, 83, 85, 86]

Most of the screening tests for biomarkers were conducted during the first and second trimester of pregnancy; however, 4 studies examined individual biomarkers or a combination of these in the third trimester of pregnancy (30 to 34 weeks where specified).[68, 69, 76, 77]

In addition, relevant results from the Conde-Agudelo 2015 SLR and MA on accuracy of biochemical tests or fetal growth assessment for the prediction of stillbirth, were also included in this review.[44] To avoid double-counting these results, no studies included in Conde-Agudelo 2015 were included in this review in their own right.

Fetal growth assessment for the prediction of stillbirth

Overall, the review identified 5 primary studies[40, 79, 84, 86, 87] and one SLR[44] that reported measures of test accuracy for the prediction of stillbirth through the assessment of fetal size or growth. Four studies assessed the use of fetal growth standards to screen for SGA as a predictor of stillbirth: 2 evaluated population-based growthgrowth standards, customised growth standards and ultrasound-based growth standards,[79, 87] one study assessed the Scottish fetal growth standard (adjusts for optimal weights *in utero*)[40] and one study compared the use of non-sex specific and sex-specific growth standards.[84] Familiari 2016 evaluated fetal femur length,[86] and the Conde-

Agudelo 2015 SLR reported measures of test accuracy for fetal femur length, as well as for suboptimal fetal growth for the prediction of stillbirth.[44]

Discussion of findings

Quality assessment

The quality and applicability of each study was appraised with a modified Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) assessment checklist.[62] The full quality assessments for the included studies, as well as the quality assessment of the Conde-Agudelo 2015 SLR, are presented in Appendix 3.

Table 5. Summary of QUADAS-2 assessments for stillbirth screening studies (biochemical and ultrasound tests, part 1)

Question	Akolekar 2016a[66]	Akolekar 2016b,[65] Aupont 2016[67]	Bakalis 2016[68]	Chaiworapongsa 2017[88]	Chaiworapongsa 2013a[69]	Chaiworapongsa 2013b[69]	Dugoff 2004[70]	Dugoff 2005[72]	Dugoff 2008[71]	Marttala 2010[78]
PARTICIPANT SELECTION										
Risk of bias	Low	Low	Low	High	Unclear	High	Unclear	Low	Low	High
Concern about applicability	Low	Low	Low	Low	High	High	Low	Low	Low	Low
INDEX TESTS										
Risk of bias	High	High	High	Low	High	Unclear	Low	Low	High	Low
Concern about applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
REFERENCE STANDARD										
Risk of bias	Unclear	Unclear	High	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Concern about applicability	Low	Low	Low	Low	High	High	Low	Low	Low	High
PARTICIPANT FLOW										
Risk of bias	Low	Low	Low	High	Low	High	Unclear	Low	High	Low

Table 6. Summary of QUADAS-2 assessments for stillbirth screening studies (biochemical and ultrasound tests, part 2)

Question	Matrodima 2016[73]	Poon 2013[80]	Singh 2012[81]	Smith 2007a[82]	Smith 2007b[74]Smith 2007b[74]	Sutan 2010[83]	Tancrede 2015[75]	Trudell 2017[85]	Valino 2016a[77]	Valino 2016b[76]	Yerlikaya 2016[50]
PARTICIPANT SELECTION											
Risk of bias	Low	High	High	Low	Low	High	Low	High	Low	Low	Low
Concern about applicability	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low
INDEX TESTS											
Risk of bias	High	Low	Low	Low	Low	High	Low	High	High	High	High

Question	Matrodima 2016[73]	Poon 2013[80]	Singh 2012[81]	Smith 2007a[82]	Smith 2007b[74]Smith 2007b[74]	Sutan 2010[83]	Tancrede 2015[75]	Trudell 2017[85]	Valino 2016a[77]	Valino 2016b[76]	Yerlikaya 2016[50]
Concern about applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
REFERENCE STANDARD											
Risk of bias	Unclear	Low	Low	Unclear	Unclear	Unclear	Low	Unclear	High	Unclear	Unclear
Concern about applicability	Low	Low	High	Low	Low	High	High	High	Unclear	Low	Low
PARTICIPANT FLOW											
Risk of bias	Low	High	High	Low	Unclear	Unclear	Low	High	High	High	Low

Table 7. Summary of QUADAS-2 assessments for stillbirth screening studies (fetal growth measurements)

Question	Familiari 2016[86]	Hemming 2011[40]	Odibo 2012[79]	Smith 2014[87]	Trudell 2015[84]
PARTICIPANT SELECTION					
Risk of bias	High	High	High	High	High
Concern about applicability	Low	Low	Low	Low	Low
INDEX TESTS					
Risk of bias	High	Low	Low	Low	Low
Concern about applicability	Low	Low	Low	High	Low
REFERENCE STANDARD					
Risk of bias	Low	Unclear	High	Unclear	Low
Concern about applicability	Low	Low	High	Low	High
PARTICIPANT FLOW					
Risk of bias	Low		High	High	Low

Participant selection

Studies were considered to be at high risk of bias with regard to participant selection if they were retrospective or made inappropriate exclusions. Overall, the risk of bias in participant selection was judged high in 10 out of 17 cohorts of pregnant women. In the Chaiworapongsa 2017 study,[88] women were randomly sampled from a cohort enrolled in a previous study by the same group (where participation required women to have at least three blood samples available for analysis). This may have introduced bias if women with complete blood samples had higher-risk pregnancies than those with incomplete samples. Also, additional cases of stillbirth from the same previous study were then selected and added to the cohort, which could be a further source of selection bias. In the Dugoff 2004 study, only pregnant women with complete index test and outcome data from the FASTER trial were initially included in the study, although it is not clear how many women were not considered for inclusion because of missing data.[70] Eleven studies were considered to be at high risk of bias because of their retrospective,[40, 78-86] or case-control designs.[69, 87] The

remaining 11 studies were all at low risk of bias for this domain, as they included pregnant women from unselected or low-risk populations.

While some studies which were conducted in unselected populations included women from high-risk groups, such as those with pre-existing diabetes, pre-eclampsia or antiphospholipid syndrome/systemic lupus erythematosus, these women constituted a very small proportion (<7%) of the overall sample in these cohorts and were thus unlikely to have biased the outcomes.[65-67, 73]

The concern about applicability was low in the majority (22) of studies, and high in only 3 studies (reported in 2 publications).[69, 81] Pregnancies with impaired fetal growth at recruitment were excluded from the Chaiworapongsa 2013 prospective cohort study,[69] therefore excluding pregnancies that may have been at risk of stillbirth due to placental dysfunction, and the sample of pregnancies included in the case-control study is unlikely to be representative of the low-risk screening population due to the substantial risk of selection bias. Singh 2012 included only singleton pregnancies who were nulliparous or had a previous pregnancy with a history of placental syndromes, and so there is high concern that the study population is not representative of an unselected or low-risk screening population.[81]

Index tests

Studies were at high risk of bias if the index test results were interpreted with knowledge of stillbirth outcomes, and if the threshold for which an index test result was considered positive or negative was not pre-specified. If the cut-off values were not pre-specified, the results of the index test may have influenced the chosen thresholds, allowing for the potential overestimation of test accuracy.

In general, there was a low risk of bias in the conduct of the index test. However, 5 articles reported measures of test accuracy for screening models that used data from a large prospective screening study in the UK.[65-68, 73] The screening models presented in those articles were based and tested on the same datasets of maternal characteristics and medical history. In addition, threshold values for the index test were not pre-specified or reported. These results are therefore at high risk of bias. Cut-off values for the index test were also not pre-specified or adequately reported in eight other studies.[50, 71, 76, 77, 83, 85, 86]

By contrast, studies where the index test results were likely interpreted with knowledge of the reference standard, but used pre-specified thresholds, were not considered to be at a high risk of bias.[40, 70, 72, 74, 78-82, 84, 87, 91] As the index test results were recorded before stillbirth was confirmed and with pre-defined thresholds, the interpretation of the index test results are unlikely to be affected by knowledge of stillbirth having occurred. There was no concern that the index test may have differed from the review question in any of the included studies.

Reference standard

The conduct of the reference standard was poorly described or defined in the majority (17) of studies, and so it was unclear how stillbirth was ascertained and whether it was likely to have been correctly classified. As all of the included studies were conducted in high-income countries, where legal requirements and standardised protocols are in place for recording stillbirths, this information may have been deemed unnecessary to include by the authors.

The lack of information provided regarding the conduct of the reference standard made it difficult to ascertain if stillbirth was confirmed with knowledge of the index test results; the risk of bias for this domain was therefore unclear in 19 publications (Table 5 to Table 7). If the results of the index test were known when a stillbirth was confirmed, this could have led to bias in the recording of a diagnosis of perinatal death outcomes including miscarriage, stillbirth, or neonatal death. Furthermore, in three articles reporting data from the same prospective screening study, it was stated that the results of the 30 to 34 weeks scan were made available to the obstetricians, who may have subsequently intervened in the pregnancy, thereby potentially avoiding complications such as stillbirth.[68, 76, 77] This could have led to an underestimation of test accuracy. It was pre-specified that studies reporting "stillbirth or SGA <3rd centile" as a combined reference standard would be eligible for inclusion in the review to mitigate this potential underestimation; however, no studies reporting this were identified.

In 11 studies, the definition of stillbirth differed from the UK definition, thereby limiting the applicability of these studies' findings to this evidence review. Nine studies that were conducted in the USA,[75, 79, 83-85, 88], Canada[75] and Chile[69] defined stillbirth as a fetal death >20 weeks of gestation. Stillbirth was defined as a fetal death occurring at or after 22 weeks in the Marttala 2010 study, conducted in Finland.[78] The Conde-Agudelo 2015 SLR noted that a wide variety of definitions for stillbirth exist, but for the purpose of the SLR, stillbirth was defined as a structurally and chromosomally normal fetus whose death occurred at or after 20 weeks of gestation.[44] Smith 2007b defined stillbirth as: a) delivery of an infant which showed no signs of life and b) all intrauterine fetal deaths subsequent to the measurement of uterine artery Doppler (which occurred at 22 to 24 weeks, at a median gestational age of 23 weeks).[74] Though this does not match the UK definition exactly, the similarity is high and so the concern about applicability was judged to be low. Seven publications reporting on the same cohort from a large prospective screening programme either reported the UK definition of stillbirth, or it was probable that this was used, as it was reported that only the pregnancies that delivered a phenotypically normal live birth or stillbirth at ≥24 weeks' gestation were included.[50, 65-67, 73, 76, 80] Two additional publications on this cohort did not clearly define stillbirth or report their inclusion criteria, but as they report data from the same prospective screening programme conducted in the UK, there is little concern about the applicability of these findings.[68, 76] Stillbirth was defined as a fetal death >20 weeks in

Chaiworapongsa 2017, but outcome data was reported for fetal death ≥24 weeks of gestation in the results section. Concern about the applicability of this study's findings is therefore low.

The reference standard in the included studies measured stillbirths resulting from all causes, whereas the index test may have been specifically targeted to predict stillbirths arising due to placental insufficiency as many of the biochemical and ultrasound markers are associated with placental function. As such, the reference standard would be expected to detect a larger number of cases (including stillbirths due to placental insufficiency as well as other causes) than the index test, which could result in lower measures of test accuracy for the index test.

Participant flow

The reference test was poorly described in the majority of studies; it was not clear how stillbirth was confirmed at birth. This raises the possibility of bias, as there may have been differences in determining whether the baby was stillborn or died shortly after birth. Studies were not considered to be at high risk of bias based on this criterion alone, as this is an unlikely scenario. Studies were however considered to be at a high risk of bias if pregnant women were excluded from the analyses; in 8 studies a considerable number of eligible pregnant women were not included in the analyses due to missing index test or outcome data.[69, 76, 77, 81, 84, 85, 87, 88] This could have introduced selection bias, potentially leading to the underestimation or overestimation of test accuracy. In the Poon 2013 study, it was reported that antepartum stillbirths were diagnosed by ultrasonography in women presenting with reduced or absent fetal movements, while intrapartum stillbirths were diagnosed at birth, and stillbirths due to placental abruption were diagnosed retrospectively.[80] As it is unclear whether all women with antepartum stillbirths received a reference standard and stillbirth was diagnosed differently within the cohort, the risk of bias for this domain is considered to be high. It is however noted that antepartum stillbirths were also very likely to have been confirmed at birth.

Conde-Agudelo 2015 SLR quality

The Conde-Agudelo 2015 SLR was of good quality as assessed by the AMSTAR checklist,[64] with concerns around only 4 of the 11 items. The limitations of the review were a) not listing conflicts of interest for each study included in the review, b) not listing studies excluded at full text screening stage, c) not providing a quality assessment for each study and d) the quality of the studies was not used to formulate conclusions. Overall, the applicability of the SLR was good, with the majority of the included studies being conducted in high-income countries, where the setting and population were similar to that of the UK.

All studies included in the SLR were assessed with the QUADAS-2 checklist, using 6 of the 14 items (Adequate study design, Description of selection criteria, Appropriate reference standard,

Adequate description of the test, Blinding, Adequate reporting of results). Only 10% were judged by the SLR authors to be of high quality (that is, fulfilling 5 or 6 criteria), whereas 79% had ≥3 methodological flaws. The most common sources of bias were lack of information in the reporting of the reference standard, mainly on blinding of clinicians to index test results and the ascertainment of fetal death before birth. Most studies also failed to report on predictive accuracy of the index test for specific stillbirth categories, which was judged as high risk of bias by the authors.

Results

Only 5 studies reported measures of test accuracy for screening tests separately for pre-term and term stillbirths, or at different gestational ages.[40, 74, 75, 82, 86] These results are presented in Table 8, and the performance of these tests has been visualised with graphs of sensitivity versus false positive rate (1−specificity) (Figure 1 to Figure 4; see Appendix 6 for further explanation of likelihood ratios and interpretation of these graphs). A further 5 publications on a single cohort assessed the accuracy of a screening test compared to a reference standard of 'stillbirths due impaired placentation'. These studies are considered separately, below.

The majority of studies identified by this review evaluated the accuracy of screening tests for the prediction of all-cause stillbirth occurring at any gestational age (depending on the definition of stillbirth used). It should be noted that many of the evaluated markers are likely to be indicative of placental dysfunction; therefore, the reported test parameters are expected to be highly relevant to predicting the risk of stillbirth due to placental dysfunction. Measures of test accuracy for the detection of all-cause stillbirth as reported by these studies are presented and discussed in Appendix 8.

A 'perfect' diagnostic test is one that is able to discriminate between test subjects who truly have and truly do not have the test condition (that is, 100% sensitivity and 100% specificity [FPR rate of 0%]), however this is rarely achievable clinically. The general consensus is that tests with a positive likelihood ratio (LR) greater than 10 and a negative LR of less than 0.1 are considered to have an acceptable accuracy, and could be considered for use in screening for a condition in clinical practice.[92, 93] These thresholds have been used a guide for the interpretation of test accuracy in this evidence review.

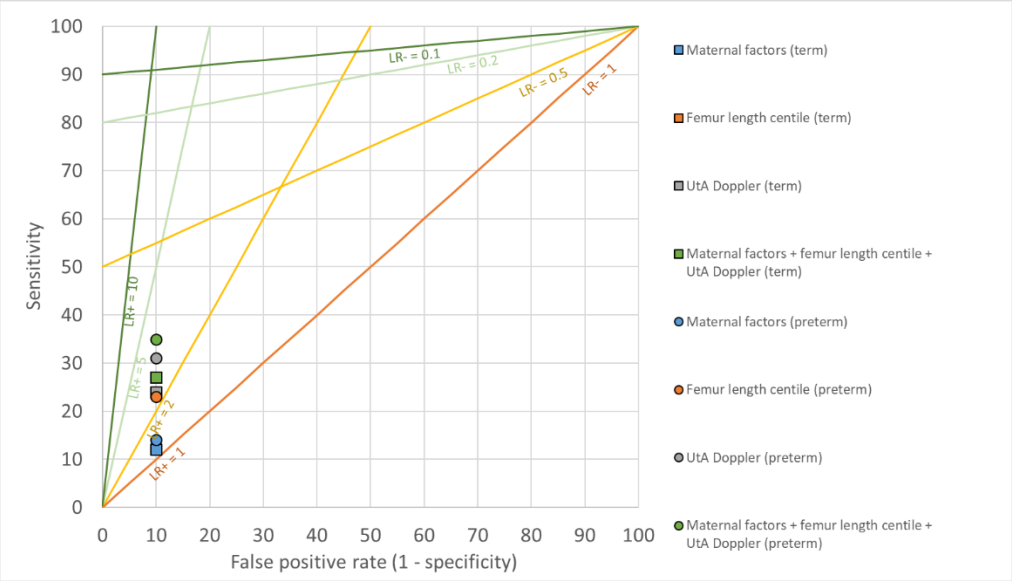
Table 8. Measures of test accuracy for screening tests for all-cause stillbirth by pre-term/term birth or gestational age at birth

Study	Test	Gestational age at birth	Women	Sens (%)	Spec (%)	PPV	NPV	LR+	LR−
Familiari 2016	Maternal factors	Term	23,894	12	90	NR	NR	NR	NR
		Pre-term	23,894	14	90	NR	NR	NR	NR
	Femur length centile	Term	23,894	27	90	NR	NR	NR	NR
		Pre-term	23,894	23	90	NR	NR	NR	NR
	Uterine artery Doppler	Term	23,894	24	90	NR	NR	NR	NR
		Pre-term	23,894	31	90	NR	NR	NR	NR
	Maternal factors + femur length centile + uterine artery Doppler	Term	23,894	27	90	NR	NR	NR	NR
		Pre-term	23,894	35	90	NR	NR	NR	NR
Hemming 2011	Fetal growth standard	Term	540,849	30	91	NR	NR	NR	NR
		34 to 36 weeks	540,849	34	90	NR	NR	NR	NR
		32 to 33 weeks	540,849	36	90	NR	NR	NR	NR
		24 to 31 weeks	540,849	43	90	NR	NR	NR	NR
Smith 2007a[a]	Maternal characteristics + AFP + hCG[a]	24 to 28 weeks (top 5%)	84,769	36.68	95.04	NR	NR	7.80	0.65
		24 to 28 weeks (top 10%)	84,769	41.51	90.04	NR	NR	4.17	0.65
		24 to 28 weeks (top 20%)	84,769	54.72	80.04	NR	NR	2.74	0.57
		29 to 32 weeks (top 5%)	84,769	18.67	95.01	NR	NR	3.74	0.86
		29 to 32 weeks (top 10%)	84,769	32.00	90.02	NR	NR	3.21	0.76
		29 to 32 weeks (top 20%)	84,769	44.00	80.02	NR	NR	2.20	0.70

Study	Test	Gestational age at birth	Women	Sens (%)	Spec (%)	PPV	NPV	LR+	LR−
		33 to 36 weeks (top 5%)	84,769	25.00	95.13	NR	NR	5.13	0.79
		33 to 36 weeks (top 10%)	84,769	34.21	90.23	NR	NR	3.50	0.73
		33 to 36 weeks (top 20%)	84,769	46.05	80.45	NR	NR	2.36	0.67
		37 to 43 weeks (top 5%)	84,769	16.78	95.02	NR	NR	3.38	0.88
		37 to 43 weeks (top 10%)	84,769	22.15	90.03	NR	NR	2.22	0.86
		37 to 43 weeks (top 20%)	84,769	35.57	80.03	NR	NR	1.78	0.81
Smith 2007b	Doppler alone (mean pulsatility index, unilateral notch and bilateral notch)	32 weeks of less	30,519	58.5	95.2	NR	NR	12.1	0.44
		33 weeks or more	30,519	6.6	95.1	NR	NR	1.3	0.91
	Doppler and maternal characteristics	32 weeks of less	30,519	53.7	95.1	NR	NR	10.9	0.49
		33 weeks or more	30,519	21.3	95.0	NR	NR	4.3	0.83
	Maternal characteristics alone	32 weeks of less	30,519	31.7	95.1	NR	NR	6.4	0.72
		33 weeks or more	30,519	18.0	95.1	NR	NR	3.7	0.91
Tancrede 2015[b]	AFP >2.0 MoM	Term	2110	-	-	0	-	NR	NR
		Pre-term	2110	0%	95.9	0	99.8	NR	NR
	hCG >2.0 MoM	Term	2125	-	-	0	-	NR	NR
		Pre-term	2125	40	89.9	0.9	99.8	NR	NR

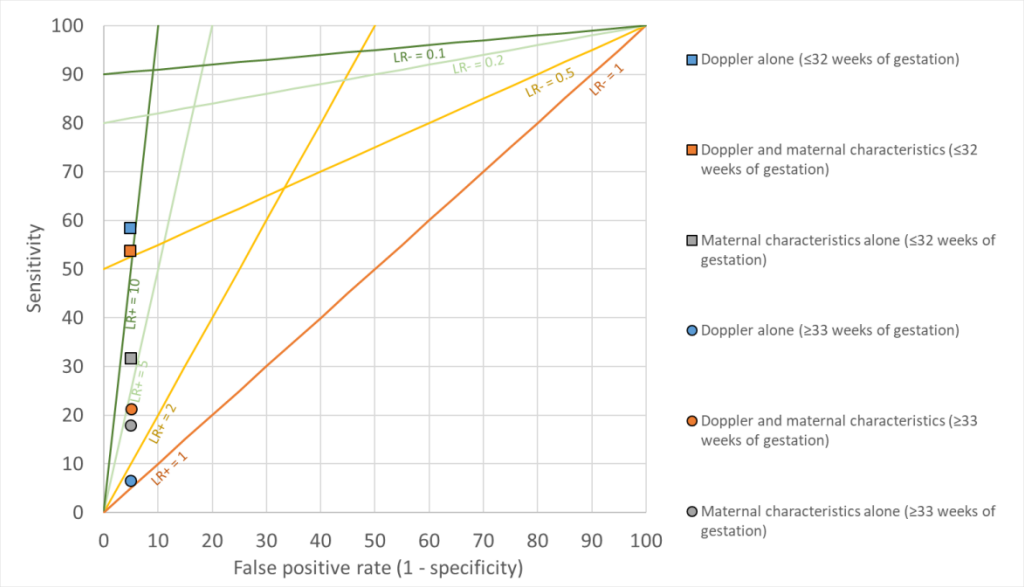
[a] Results are presented for women in the top 5%, 10% and 20% of predicted risk
[b] 3466 women were enrolled in the study but it was reported that 2110 and 2125 of these women had data available for serum AFP and serum hCG
Abbreviations: AFP: alpha fetoprotein; hCG: human chorionic gonadotropin; LR: likelihood ratio; MoM: multiple of the median; PPV: positive predictive value; NPV: negative predictive value: NR: not reported; SGA: small-for-gestational age.

Figure 1. Tests for the prediction of stillbirth by pre-term/term births (Familiari 2016)



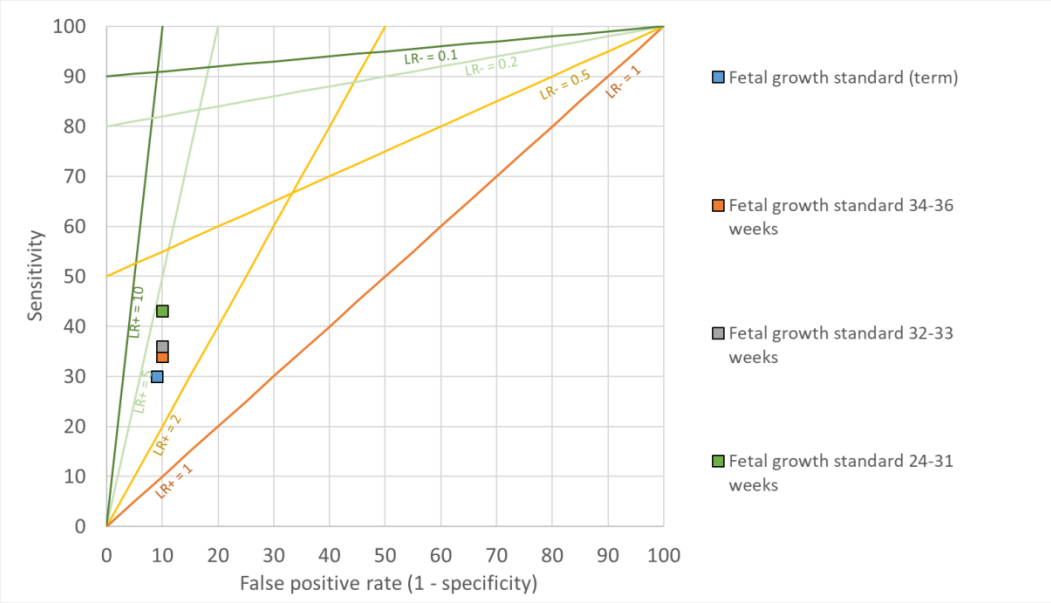
Abbreviations: LR: likelihood ratio; UtA: uterine artery.

Figure 2. Tests for the prediction of stillbirth at ≤32 weeks versus ≥33 weeks of gestation at birth (Smith 2007b)



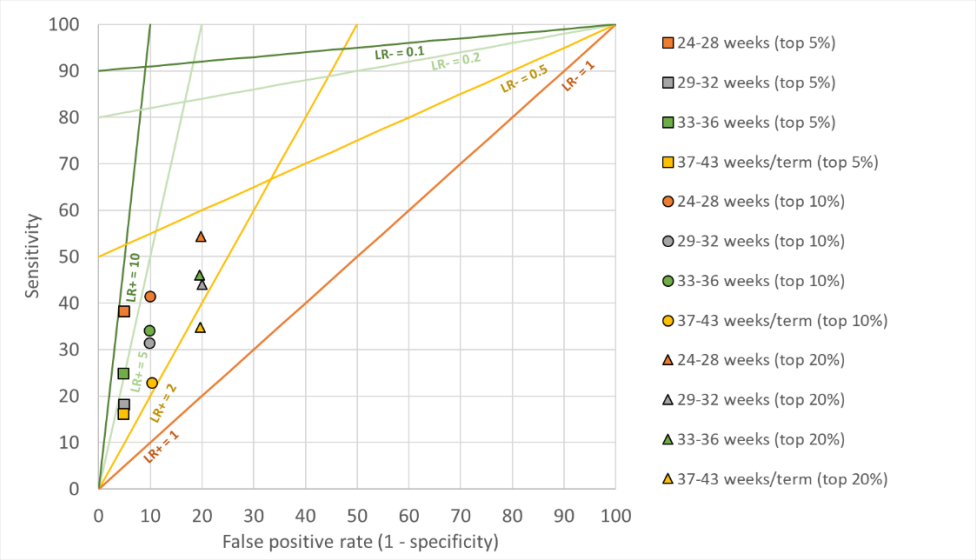
Abbreviations: LR: likelihood ratio.

Figure 3. Fetal growth standards for the prediction of stillbirth by gestational age at birth (Hemming 2011)



Abbreviations: LR: likelihood ratio.

Figure 4. Maternal characteristics, AFP and hCG for the prediction of stillbirth by gestational age at birth, for women in the top 5%, 10% and 20% of predicted risk of screening positive (Smith 2007a)



Abbreviations: AFP: alpha fetoprotein; hCG: human chorionic gonadotropin; LR: likelihood ratio.

Familiari 2016 specifically reported measures of test accuracy for 'pre-term' and 'term' stillbirths. Assessment of UtA-PI at 19 to 24 weeks of gestation had a higher sensitivity for the detection of

pre-term stillbirths (31% sensitivity, 90% specificity) than for stillbirths at term (24% sensitivity, 90% specificity). By contrast, femur length centile had a slightly higher sensitivity for term stillbirths (23% vs. 27% at 90% specificity, respectively). A test of UtA-PI and femur length centile in combination with maternal factors achieved a sensitivity of 35% for pre-term births, in comparison with 27% for stillbirths at term. However ultimately, there was not a large difference in test accuracy for the detection of pre-term and term stillbirths using maternal characteristics, femur length centile, UtA-PI, or a combination of these three tests. It is worth noting that determination of stillbirth risk in UK practice is currently often based on maternal risk factors alone, which, as highlighted by this study, is not a very reliable method of risk quantification (sensitivity 12 to 14 %).

Smith 2007a presented measures of test accuracy for stillbirths at term (37 to 43 weeks' gestation), and further stratified by gestational age for pre-term stillbirths.[82] As shown in Figure 5, test accuracies for the combination test of maternal characteristics, AFP and hCG were highest for pre-term stillbirths at 24-28 weeks' gestation, for women in the top 5%, 10% and 20% of predicted risk (orange markers).[82]

For the prediction of pre-term stillbirths at ≤32 weeks' gestation, the use of UtA Doppler alone achieved a very high LR+ of 12.1, and UtA Doppler in combination with maternal characteristics achieved a LR+ of 10.9. However, the LR- for these tests were 0.44 and 0.49, respectively, which is higher than desirable to rule out stillbirth on the basis of a negative test result.[74]

Overall, it appears that there was an improvement in test accuracy for the detection of stillbirths occurring at earlier gestational ages using: fetal growth standards;[40] ultrasound Doppler (mean pulsatility index, unilateral notch and bilateral notch,[74] and UtA-PI[90]); and combination tests of maternal characteristics with biochemical markers (AFP and hCG)[74] or ultrasound Doppler.[74]

The Tancrede 2015 study did not present sufficient data for comparison of test accuracy for pre-term versus term stillbirths.[75] The data presented in Table 8 and Figure 1 to Figure 4 suggests that the accuracy of screening tests may be higher for pre-term stillbirths and at earlier gestational ages than for term stillbirths (and at later gestational ages). Nevertheless, no tests for the detection of stillbirth at any gestational age adequately quantified the risk of stillbirth for use in pregnant women in clinical practice.

Placental dysfunction

Many of the biochemical markers evaluated in the studies included in this review are associated with impaired placentation; however, only four articles, all reporting data from one large cohort, specifically explored the accuracy of these tests in identifying stillbirths caused by placental insufficiency. In these studies, stillbirths were considered to be the result of impaired placentation when they were associated with pre-eclampsia, placental abruption or a birthweight <10th

percentile for gestational age. It is important to consider that the use of birthweight <10th centile as a marker of placental dysfunction is likely to have biased the results toward the detection of pre-term stillbirths, which are more commonly associated with growth restriction. Any stillbirths arising from placental dysfunction with a birthweight >10th centile will have been excluded from this analysis.

All tests were conducted in the first or second trimester, at 11 to 13 weeks,[66, 73] or 19 to 24 weeks of gestation.[65, 67] The measures of test accuracy for the combination tests evaluated for the prediction of stillbirths due to impaired placentation are presented in Table 9. Graphs of sensitivity versus false positive rate (1−specificity) are presented in Appendix 7.

Table 9. Measures of test accuracy for screening tests for the prediction of stillbirths arising from impaired placentation

Test	Study	Women	Sens (%)	Spec (%)
Combination tests				
Maternal factors	Akolekar 2016b (FPR 5%)	70,003	22.6	95
	Akolekar 2016b (FPR 10%)	70,003	34.0	90
	Mastrodima 2016 (FPR 5%)	76,897	23.6	95
	Mastrodima 2016 (FPR 10%)	76,897	36.3	90
Maternal factors + PlGF	Akolekar 2016a (FPR 5%)	45,452	40.5	95
	Akolekar 2016a (FPR 10%)	45,452	51.1	90
Maternal factors + DV-PIV	Akolekar 2016a (FPR 5%)	45,452	28.2	95
	Akolekar 2016a (FPR 10%)	45,452	37.4	90
Maternal factors + UT-PI	Akolekar 2016a (FPR 5%)	45,452	35.1	95
	Akolekar 2016a (FPR 10%)	45,452	45.8	90
Maternal factors + PlGF + DV-PIV	Akolekar 2016a (FPR 5%)	45,452	42.0	95
	Akolekar 2016a (FPR 10%)	45,452	51.4	90
Maternal factors + UT-PI + DV-PIV	Akolekar 2016a (FPR 5%)	45,452	36.6	95
	Akolekar 2016a (FPR 10%)	45,452	48.1	90
Maternal factors + PlGF + UT-PI	Akolekar 2016a (FPR 5%)	45,452	47.9	95
	Akolekar 2016a (FPR 10%)	45,452	60.8	90
Maternal factors + PlGF + DV-PIV + UT-PI	Akolekar 2016a (FPR 5%)	45,452	48.1	95
	Akolekar 2016a (FPR 10%)	45,452	61.1	90
Maternal factors + fetal biometry	Akolekar 2016b (FPR 5%)	70,003	52.8	95
	Akolekar 2016b (FPR 10%)	70,003	34.0	90
Maternal factors + UtA-PI	Akolekar 2016b (FPR 5%)	70,003	62.3	95
	Akolekar 2016b (FPR 10%)	70,003	73.6	90

Maternal factors + biometry + UtA-PI	Aupont 2016 (FPR 5%)	70,003	69.8	95
	Aupont 2016 (FPR 10%)	70,003	74.8	90
Maternal factors + biometry + UtA-PI + PlGF	Aupont 2016 (FPR 5%)	70,003	76.1	95
	Aupont 2016 (FPR 10%)	70,003	83.6	90
Maternal factors + PAPP-A + UT-PI + DV-PIV	Mastrodima 2016 (FPR 5%)	76,897	47.8	95
	Mastrodima 2016 (FPR 10%)	76,897	54.8	90

Predictive values and likelihood ratios were not reported.

Abbreviations: DV-PIV: ductus venous pulsatility index; FPR: false positive rate; PAPP-A: pregnancy-associated plasma protein; PlGF: placental growth factor; UtA/UT-PI: uterine artery pulsatility index;

The sensitivity of combination tests for the prediction of stillbirths arising from placental dysfunction was found to be notably improved (by up to 15%) compared with the sensitivity for the prediction of all-cause stillbirth.[65-67, 73] A combination test of maternal factors, fetal biometry, UtA-PI and PlGF at 19 to 24 weeks gestation was able to identify 83.6% of stillbirths arising from impaired placentation, in comparison with 57.6% all-cause stillbirths at a specificity of 90%. [67] Whilst the sensitivity of these tests is moderate, none of these tests could be considered sufficiently accurate for use in clinical practice using the threshold of a LR+>10 and LR− <0.1 (Appendix 7). The evidence base for screening specifically for stillbirths related to placental dysfunction is currently based on results from a single cohort; the potential for use of any of these tests for screening cannot be substantiated without further research. As the diagnosis of stillbirths in most countries does not currently distinguish between stillbirths arising from placental dysfunction and those arising from other causes, this is a limiting factor in the current evidence base and in the available literature.

Measures of test accuracy for the prediction of all-cause stillbirth

The best performing index test, as reported by the Conde-Agudelo 2015 SLR, achieved a LR+ >10 and a LR− <0.1, thereby falling into the 'acceptable accurate'* category. This test, using UtA bilateral notch measurements, reportedly achieved a LR+ of 13.3 and a LR− of 0 (100% sensitivity and 93% specificity). However, as this study had a very small sample size with only 10 cases of stillbirth, it is likely that the test accuracy has been overestimated. A meta-regression conducted in the Conde-Agudelo 2015 SLR found that studies with fewer than 25 stillbirths in the study sample often overestimated the accuracy of the test, additionally concluding that using UtA bilateral notch is unlikely to be a sufficiently accurate test for stillbirth due to placental dysfunction.[44] Thus, this finding cannot be confirmed without further high-quality prospective screening studies.

No other screening tests identified by this review could be considered to be capable of accurate quantification of the risk of stillbirth. For ultrasound markers, three studies evaluated UtA-RI as a screening test for stillbirth, but measures of test accuracy were only adequately reported for one test included in the Conde-Agudelo 2015 SLR.[44] There is therefore insufficient evidence to draw conclusions on the usefulness of this test for predicting stillbirth. Due to uncertainty around the study quality, there is also insufficient evidence to draw conclusions on the usefulness of fetal nuchal translucency or ductus venosus reversed A wave in predicting all-cause stillbirth.

This review additionally identified studies examining screening for biomarkers in the third trimester of pregnancy.[68, 69, 76] This is not reflective of general UK clinical practice for low-risk pregnancies; tests for biochemical abnormalities are typically performed in the first or second trimester. High

* A test is considered to have an an acceptable accuracy if +LR greater than 10 and a - LR of less than 0.1

measures of test accuracy for the angiogenic-1 index and the PlGF/sEng ratio performed at 30 to 34 weeks of gestation were reported by the Chaiworapongsa 2013 cohort and case-control studies.[69] As discussed above, these studies were at high risk of bias, so this review did not find sufficient evidence to suggest that screening for biochemical markers in isolation or as a combination in the third trimester could have superior accuracy than screening in the first or second trimester.

Seven studies evaluating methods of fetal growth monitoring for the prediction of all-cause stillbirth were included in this review (one of which was reported by the Conde-Agudelo 2015 SLR).[44] In general, measures of test accuracy for population-based birth weight standards for stillbirth were poorer than for customised (and sex-specific) growth standards, although clinically meaningful comparison between the fetal growth assessment tests as reported by different studies is limited by the use of different fetal growth standards.

Conclusions

This evidence summary included 5 studies that evaluated the use of screening tests for the prediction of pre-term and term stillbirths separately, or by gestational age at birth. Four out of 5 studies were conducted in large cohorts, ranging from 23,894 to 540,849 pregnant women.[40, 90] While the data reported by these studies suggests that there is a possibility that the accuracy of screening tests is higher for pre-term stillbirths and at earlier gestational ages than for term stillbirths (and at later gestational ages), this finding was not consistent. It is important to consider the high risk of intervention bias in the included screening studies; pregnant women and health providers were not known to be blinded to test results in any study, and therefore could have intervened with pregnancies suspected to be at high risk of negative outcomes, including stillbirth. By preventing stillbirth in some pregnancies, test accuracy may have been underestimated.

Based on the current evidence, there are currently no tests that are appropriate for use in a screening programme aimed at predicting pregnancies at risk of stillbirth due to placental dysfunction in clinical practice. No tests for the detection of stillbirth at any gestational age have been demonstrated to be suitable to quantify the risk of stillbirth in early or late pregnancies. There was also insufficient evidence to support any adequately accurate screening tests for the prediction of all-cause stillbirth or stillbirth specifically arising from placental insufficiency. Nevertheless, it should be noted that the risk assessment strategies currently used in clinical practice may perform worse than some tests identified in this review and while these tests have not been demonstrated to quantify risk sufficiently well to be used in a screening programme, they may represent an improvement to the status quo.

Summary of findings relevant to criterion 4: criterion not met

Quantity: A moderate number of studies was identified and included in this evidence review to address this question, mostly undertaken in moderate to large-sized cohorts. However, only 6 studies separately presented measures of test accuracy for term and pre-term stillbirths, or at different gestational ages, and screening for stillbirths using a reference standard of stillbirths arising from placental dysfunction was assessed in only one large cohort.

Quality: The quality of the included studies was variable; the review identified large, prospective studies for biochemical markers, UtA-PI and fetal growth charts. Fetal nuchal translucency, UtA-RI and ductus venosus reversed A wave were only reported through retrospective, lower-quality studies. The reference standard was poorly described in the majority of studies.

Applicability: All of the studies included in this review were conducted in high-income countries that are considered to be reflective of the UK setting, and the majority of the biomarkers investigated in the included studies are already used in UK clinical practice in guideline-directed monitoring of high-risk pregnancies. However, in most cases, test accuracies are reported for all-cause stillbirth despite these biochemical markers being biologically associated with placental dysfunction. Seventeen publications reporting on six unique cohorts used or were likely to have used the UK definition of stillbirth (a fetal death occurring at ≥24 weeks of gestation),[40, 50, 65-68, 70, 71, 73, 74, 76, 80, 82, 86, 89] thereby limiting the applicability of the findings reported in the remaining 10 publications reporting on 13 cohorts.[69, 74, 75, 78, 79, 81, 83-85]

Consistency: A wide range of tests were investigated, with only a small number of individual tests investigated by multiple studies. Where the same test was investigated in more than one study, different thresholds were to classify the index test result, thereby restricting the direct comparison between the results. As a result, limited evidence-based conclusions can be drawn regarding the diagnostic accuracy for any particular diagnostic test for pre-term or term stillbirths, stillbirths specifically due to placental dysfunction, or all-cause stillbirths.

Conclusions:
Only 5 studies separately presented measures of test accuracy for term and pre-term stillbirths, or at different gestational ages, and only 4 records reporting on a single cohort specifically reported measures of test accuracy for screening for stillbirth using a reference standard of ‘stillbirth arising from impaired placentation’. Overall, no tests using biochemical markers, ultrasound tests, fetal growth charts, or a combination of these methods were found to be to have an acceptable accuracy* by high-quality studies, suggesting that there are currently no known screening tests that are potentially appropriate for clinical use in the UK. Until further high-quality evidence of a sensitive and specific screening test for pre-term or term stillbirths associated with placental dysfunction becomes available, a screening programme for stillbirth is unlikely to be recommended.

Criterion 7 – Monitoring for preventing stillbirth

'There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.'

Question 2 – What is the appropriate monitoring regime for pregnancies that have been identified by screening to be at risk of stillbirth?

a) Before 37 completed weeks of gestation
b) From 37 completed weeks of gestation

No prior evidence reviews have been conducted to support the decision of whether a programme of screening to prevent stillbirth should be recommended by the UK NSC.

There are currently no specific recommendations on how best to monitor pregnancies at risk of stillbirth due to placental dysfunction. In pregnancies at risk of other adverse perinatal outcomes, monitoring may include collection of various Doppler ultrasound parameters, cardiotocography (CTG), or decreased fetal movement (DFM) counting. This review searched for relevant data published since 2000 that reported on the effect of different monitoring regimes on the risk of pre-term or term stillbirth in at-risk pregnancies.

Eligibility for inclusion in the review

This review searched for RCTs, interventional, cohort and case-control studies, as well as SLRs and MAs of the above, for monitoring regimes that could be used to refine and manage the risk ofstillbirth. Studies had to assess different ways that pregnancies at risk of stillbirth could be monitored. Examples of monitoring regimes included fetal movement counting, heart rate monitoring or evaluating the fetal biophysical profile. All included studies required a normal care comparator; that is, a cohort of women being monitored as per standard care guidelines (which may include no monitoring).

Publications were only included if they reported risk of severe SGA (birthweight <3rd centile) or stillbirth, stratified by pre-term (before 37 completed weeks of gestation) or term (from 37 completed weeks) occurrence where possible. Included studies were conducted in the UK or other high-income countries, where the screening methods and technology are expected to be similar to that of the UK.

The eligible population were women at risk of stillbirth due to placental dysfunction detected through screening, testing or due to having specific maternal risk factors.

Full details of eligibility criteria are presented in Table 3.

Description of the evidence

There were 3 studies that assessed monitoring regimes for pregnancies at high risk of stillbirth. Of these, one was a RCT, one a prospective study and one a retrospective study.[94-96] The RCT was conducted in multiple EEA countries — Germany, Italy, the Netherlands, Austria and the UK.[94] The cohort studies were conducted in France and Norway.[95, 96]

Monitoring strategies investigated by the 3 studies included electronic fetal heart rate monitoring (CTG), Doppler flow velocimetry and maternal serum AFP. Lees 2015 compared fetal and neonatal mortality in 3 antenatal monitoring strategies: reduced cardiotocograph fetal heart rate STV (CTG STV), early DV changes (PI >95th percentile; DV p95), or late DV changes (A wave [the deflection within the venous waveform signifying atrial contraction] at or below baseline; DV no A). The study also compared neurologic impairment at 2 years in the 3 groups.[94] Spaggiari 2013 aimed to assess maternal-fetal outcomes in pregnancies associated with persistently elevated second-trimester maternal serum AFP.[95] Finally, Tveit 2009 examined two cohorts of women with DFM before and during two consensus-based interventions aiming to improve care through: 1) written information to women about fetal activity and DFM, 2) guidelines for management of DFM for health-care professionals.[96]

Women were monitored at different gestational ages; the earliest gestational age that monitoring commenced was in the second trimester in Spaggiari 2013 (14 to 18 weeks), where second trimester maternal serum markers used in screening for Down syndrome in singleton pregnancies, including hCGb and AFP, were assayed.[95] Median gestational age at sampling was 16^{+1} weeks (IQR: 15^{+3} to 17^{+5}).[95] Although not specified as such, stillbirths in this study were pre-term, reported between 15 and 30 weeks (median: 17 weeks).[95] Pregnancies with a second screen were also pre-term, with stillbirths reported between 21 and 31 weeks (median: 26 weeks).[95] Conversely, in Lees 2015 and Tveit 2009 the monitoring was only started in the third trimester (26 to 32 weeks, and ≥28 weeks respectively).[94, 96] In Lees 2015, median gestation at delivery was 30.7 weeks (IQR 29.1 to 32.1).[94] Tveit 2009 included women with singleton pregnancies of at least 28 weeks gestation or more who reported a concern for DFM (either by spontaneous reporting or upon questioning).[96] It is unclear whether women were pre-term or at term when reporting DFM or when the reported stillbirths occurred.

In addition to stillbirth, Spaggiari 2013 also reported the outcome of severe SGA (defined here as birthweight <3rd centile); a condition that may indicate that the fetus could have been stillborn had the birth been delayed, and as such, is of interest in studies where birth may be premature or differ between arms.

Discussion of findings

Quality assessment

The quality of the 3 identified primary studies was assessed using a modified Downs and Black checklist, presented in Appendix 4, Table 29. A summary of the risk of bias and applicability to the UK setting is presented in Table 10, with the full appraisal presented in Appendix 3, Table 25.

Table 10. Summary of Downs and Black assessments for monitoring to prevent stillbirth studies

Domain	Lees 2015[94]	Spaggiari 2013[95]Spaggiari 2013[95]	Tveit 2009[96]
EXTERNAL VALIDITY	Unclear	Low	High
BIAS	Unclear	Unclear	High
POWER	Unclear	Unclear	Unclear

External validity

Only one included study, Spaggiari 2013, included a population closely aligned to the review question.[95] In this study, women with singleton pregnancies from an otherwise unselected population were included if they had a maternal serum alpha fetoprotein (AFP) level of ≥ 2.5 MoM; as AFP has been suggested as a potential screening test for stillbirth,[75, 82] this population is highly applicable to the review question.[95]

In Lees 2015, women were eligible for inclusion if they were admitted to hospital with singleton pregnancies, FGR diagnosis and abnormal UA Doppler with a PI >95th percentile with or without reversed or absent end-diastolic flow.[94] However, details relating to the reasons for admission were not reported in the study and randomisation took place in specialist units. As such there is uncertainty about the applicability of the study participants to a population requiring monitoring as a consequence of screening to predict stillbirth caused by placental dysfunction. In addition, the study population presented between 26 and 32 weeks gestation. The applicability of the study results outside this gestational age range, and therefore its relevance to monitoring at different gestational ages, is also uncertain.

In Tveit 2009, eligibility criteria included women with singleton third trimester pregnancies at ≥28 weeks gestation who reported a concern for DFM, either by spontaneous reporting or upon questioning; it is unclear how many women were questioned about DFM, thus women with awareness of fetal movement counting may have been more likely to present to the unit than those without awareness who were not questioned.[96] This also leads to concerns over applicability of the evidence, as women were not assessed to be at risk of stillbirth by screening and so the population may not be aligned with a screen-detected population.

Bias

Lees 2015 was an RCT with appropriate randomisation and allocation concealment, and intervention arms recruited from the same population and balanced in baseline characteristics.[94] However, the definition of stillbirth is unclear, which could affect data collection and risk of bias. The study was unblinded for both study subjects and those measuring the main outcome (due to the nature of the trial), which may impact confounding. As such, it was at a high risk of ascertainment bias, as investigators or medical staff were likely to intervene with pregnancies known to be at high risk regardless of study group. The additional care could contribute to fewer stillbirths, and thus has the potential to lead to an underestimation of comparative effectiveness in the prevention of stillbirth. Losses to follow-up were taken into account in Lees 2015 with appropriate intention-to-treat (ITT) analysis carried out.

Spaggiari 2013 and Tveit 2009 were observational studies with a greater risk of confounding. Neither study was blinded and there was no randomisation to monitoring groups. Spaggiari 2013 also did not report baseline characteristics for the different groups and no adjustment for confounding was reported.[95] Furthermore, although stillbirth was clearly defined, it is unclear whether this was measured reliably in line with the definition reported, resulting in an unclear risk of ascertainment bias.[95] Tveit 2009 gave no definition for stillbirth and therefore there is uncertainty on how well stillbirths were recorded given the observational nature of the study, resulting in a high risk of bias.[96]

While Spaggiari 2013 provided a clear definition of stillbirth, this was not aligned with the UK definition.[95] Definitions of stillbirth could be inferred from the eligibility criteria of the remaining studies; however, both of these definitions were still different to that in the UK.[94, 96] Furthermore, none of the studies specifically reported stillbirths due to placental dysfunction.

Power

The cohorts ranged from 92 to 3,038 women enrolled per study arm.[95, 96] However, few or no stillbirths occurred in most arms. Power calculations were not performed in any of the 3 included studies, and as such, their power to detect stillbirth or severe SGA (birthweight <3rd centile) was unclear.

Results

A summary of stillbirth and severe SGA outcomes for the identified monitoring regimes is presented in Table 11. Full study results and study details are provided in Appendix 3, Table 22.

Table 11. Measures of monitoring to prevent stillbirth/severe SGA

Study	Intervention	N per arm	SGA, n (%)	P value	Stillbirth, n (%)	Effect on risk of stillbirth	P value
CTG/Doppler flow velocimetry							
Lees 2015[94]	CTG STV[a]	166	NR	NR	0 (0)[d] [2 (1)][e]	NR	NR
	DV p95[b]	167	NR	NR	3 (2)[d] [4 (2)][e]		
	DV no A[c]	170	NR	NR	4 (2)[d] [6 (4)][e]		
Maternal serum							
Spaggiari 2013[95]	AFP >2.5 MoM – No 2nd assay	273	18 (6.6)	NR	46 (16.8)	NR	NR
	2nd Assay – AFP remained high	92	10 (10.9)	0.01	9 (9.8)	Pregnancies with AFP remaining ≥2.5 MoM had significantly higher rates of IUFD	0.005
	2nd Assay – AFP returned to normal	226	8 (3.5)		5 (2.1)		
Consensus-based information							
Tveit 2009[96]	Written information to women about fetal activity and DFM / guidelines for management	3,038	NR	NR	73 (2.4)	Adjusted OR: 0.51 95% CI: 0.32 to 0.81	0.004
	No written information	1,215	NR	NR	50 (4.2)		

Abbreviations: AFP: alpha-fetoprotein; CI: confidence interval; CTG: cardiotocography; CTG STV: cardiotocograph short-term variation; DFM: decreased fetal movement; DV: ductus venosus Doppler waveform; IUFD: intrauterine fetal demise; MSAFP: maternal serum alpha-fetoprotein; NR: not reported; OR: odds ratio.
[a] reduced cardiotocograph fetal heart rate STV
[b] early DV changes (pulsatility index >95th percentile)
[c] late DV changes (A wave [the deflection within the venous waveform signifying atrial contraction] at or below baseline)
[d] reported as unexpected death between scheduled follow-up appointments
[e] reported as sum of unexpected fetal death (see d above) and fetal death where parents declined delivery despite being indicated according to study criteria

Data on pre-term and term stillbirths was not reported separately in any of the included studies. However, based on the information given, Lees 2015 and Spaggiari 2013 appear to limit the data collected to pre-term stillbirths.

The Lees 2015 RCT included women with singleton fetuses at 26 to 32 weeks gestation.[94] The median age at delivery was 30.7 weeks. The study allowed for comparison between the 3 monitoring arms. In relation to stillbirth events, the results of the ITT and per protocol analyses were in the same direction. Both analyses reported that the lowest rate of stillbirth was in the CTG group compared to the two Doppler groups (Table 11). In terms of the proportion of babies surviving without age-adjusted neurological impairment at 2 years of age, the study reported no difference between the 3 arms. The highest rate of overall survival was in the CTG group.

However, amongst the babies who survived, highest rate of neurological impairment was found in the CTG group and the lowest rate was found in the DV no A Doppler group.

Spaggiari 2013 found that a high maternal serum AFP level is associated with a high rate of pregnancy complications, including stillbirth and severe SGA.[97] The cohort of women having a second AFP assay had a lower number of stillbirths than the cohort of women only having had one AFP measurement, although no statistical analysis was conducted so the statistical significance of this result is unconfirmed. Further, it appears that there was no difference in severe SGA (birthweight <3rd centile) between the 2 cohorts of women. As expected, among those who did have a second assay, the risk of stillbirth was significantly decreased for women where AFP returned to normal compared with those where AFP remained ≥2.5 MoM on a second assay ($p<0.005$). Monitoring commenced in the second trimester in Spaggiari 2013 (14 to 18 weeks). Late stillbirth was not reported so it is unclear whether late stillbirth would have been prevented in this study.

Tveit 2009 was the only study where the monitoring strategy investigated was shown to be significantly better at reducing the risk of stillbirth.[96] The study concluded that improved information on management of DFM and providing uniform information to women is associated with fewer stillbirths (aOR: 0.51; 95% CI: 0.32 to 0.81; $p=0.004$).[96] This study therefore offers some evidence that monitoring for DFM results in a reduction of stillbirth compared with no intervention. However, unlike the TRUFFLE study, there was no analysis of the longer-term outcomes which may accompany the higher rate of survival. This limits the conclusions that can be drawn. In addition, precise monitoring strategy details were not provided and the distinction between pre-term and term stillbirths is unclear.

Conclusions

Overall, there was a lack of data on which monitoring strategies are appropriate in pregnancies at risk of pre-term or term stillbirths; paricurarlly in relation to term pregnancy. These studies were not powered to detect a difference in stillbirth or severe SGA, and no studies differentiated between pre-term and term stillbirth. The 3 studies show that monitoring strategies can impact on stillbirth. However only one study included a no monitoring group so the degree to which they do reduce stillbirth is uncertain.

The evidence base relating to monitoring strategies was limited in volume. Only 3 studies were identified and each explored different monitoring modalities in different populations. For example, the 3 studies varied in the gesational age range in which the modality was applied.

Importantly, in 2 studies there was concern about the applicability of the study participants to a screen-detected population. And only 1 study explored the long-term outcomes associated with survival following monitoring.

Summary of findings relevant to criterion 11: Criterion not met

Quantity: The evidence base for question 2 was small. Only 3 studies with a small number of participants were identified as relevant for this question.

Quality: Studies were of mixed quality and external validity. Bias was generally unclear or high, mostly due to unclear definitions of stillbirth/severe SGA and the observational nature of 2 studies. Power calculations were not performed in any studies, and as such, their power to detect stillbirth or severe SGA was unclear.

Applicability: In all 3 included studies there was a high or an uncertain concern about applicability to the review question. This is mostly due to the fact that results were not reported separately for pre-term and term stillbirth. Further, no studies used the UK definition of stillbirth, limiting their potential applicability to the UK setting. In 2 out of 3 studies there was also a concern that the included cohort of women was not representative of women with screen-detected risk of stillbirth.

Consistency: None of the examined monitoring strategies were investigated in more than one study, thus it is not possible to assess consistency between the results.

Conclusions: The evidence base for management strategies to prevent pre-term or term stillbirths is limited. It does not allow for any conclusions to be drawn on the optimal monitoring strategy for pregnancies identified as high risk in a screening programme focusing on risk of stillbirth related to placental dysfunction. While the studies may suggest candidates for monitoring in this context, this criterion cannot be not met without further high-quality studies comparing monitoring regimes in screen-detected high-risk pregnancies.

Criteria 9 and 10 – Interventions for preventing stillbirth

Criteria 9 and 10 of the UK NSC Screening Criteria state that:

9: 'There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered'

10: 'There should be agreed evidence-based policies covering which individuals should be offered interventions and the appropriate intervention to be offered'

No prior evidence reviews have been conducted to determine whether a programme of screening to prevent stillbirth should be recommended by the UK NSC.

Currently, the only intervention used in clinical practice to prevent stillbirth due to placental dysfunction is elective birth. It is unclear if any pharmacological interventions may be effective at preventing stillbirth and by that facilitating the avoidance of elective birth. If interventions are available, timing of administration needs to be determined. This review searched for relevant data published since 2000, and assessed questions relating to the effectiveness of interventions for preventing pre-term or term stillbirth. Evidence relating to these criteria was assessed through two questions, considering evidence on interventions other than elective birth, or planned delivery and expectant management, respectively.

Question 3 – Are there any effective interventions other than elective birth to prevent stillbirth in women identified as high risk through screening?

a) Before 37 completed weeks of gestation

b) From 37 completed weeks of gestation

Eligibility for inclusion in the review

This review searched for randomised controlled trials, interventional studies, cohort and case-control studies, as well as SLRs and MAs of the above. Studies were included if the population comprised pregnant women at high risk of stillbirth due to placental dysfunction, determined through screening, testing or due to having a combination of maternal risk factors. Once identified, women could receive interventions directed at preventing stillbirth caused by placental insufficiency; aspirin, anti-coagulants, anti-platelets, or anti-hypertensives were particularly relevant, though other similar interventions were also eligible. Publications were only included if

they reported risk of stillbirth; ideally this would be stratified by pre-term or term (where possible) or severe SGA (<3rd centile). Studies had to be conducted in the UK or other high-income countries.

Full details of eligibility criteria are presented in Table 4.

Description of the evidence

Six studies reported in 7 articles were included in the review for question 3.[8-13] Of the 6 relevant studies, one investigated unfractionated heparin (UFH),[10] two sildenafil citrate,[13, 14] 2 aspirin alone,[8, 11] and one compared aspirin and enoxaparin with aspirin alone.[9] Comparators were aspirin alone, placebo, standard care or no treatment, and in one study only a single arm was relevant for inclusion.[11]One study was conducted in 19 fetal medicine units in the UK,[14] with 2 studies conducted in Canada,[10, 13] 2 in France,[9, 11] and one in Spain.[8] None of the studies reported the gestational age at which stillbirths or livebirths with severe SGA occurred or stratified the results by pre-term and term, but where possible information on gestational ages at enrolment and delivery was used to infer whether the results were more applicable to pre-term or term stillbirth. In addition to stillbirth, 3 studies also reported the outcome of severe SGA (defined here as birthweight <3rd centile); a condition that may indicate that the fetus could have been stillborn had delivery been delayed, and as such, is of interest in studies where delivery may be premature or differ between treatment arms.

Discussion of findings

Quality assessment

The quality of the 6 included studies was appraised using an adapted Downs and Black checklist, Appendix 4. A summary of the risk of bias and applicability to the UK setting is presented in Table 12, and the full appraisal is presented in Table 26.

Table 12. Summary of Downs and Black assessments for interventions to prevent stillbirth studies

Question	Ayala 2012[8]	Haddad 2016[9]	Kingdom 2011[10]	STRIDER[14]	Subtil 2003a[11] Subtil 2003b[12]	von Dadelszen 2011[13]
EXTERNAL VALIDITY	High	High	High	High	Low	High
BIAS	Low	Low	Low	Unclear	Unclear	High
POWER	Unclear	High	High	High	NA	High

External validity

Only one of the 6 studies included a population directly applicable to the review question.[11] Women in Subtil 2003 were selected from a low-risk population, and randomly assigned to receive or not receive uteroplacental artery Doppler ultrasound. Of those randomised to ultrasound, only those with abnormal results were then prescribed the intervention (aspirin).[11] In Ayala 2012, Kingdom 2011, von Dadelszen 2011 and STRIDER, women were at high risk but it is unclear how they were selected.[8-10, 14] Thus, the evidence is of limited applicability for this review, as the risk for these conditions can only approximately indicate that a pregnancy is also at an increased risk for stillbirth and thus would not be aligned with the screen-detected UK population at high risk of stillbirth.

One of the 5 studies was conducted in the UK, in which results were reported for fetal death <26 weeks' gestation (which would therefore include cases of miscarriage) and at ≥26 weeks' gestation,[14] which allows for the measurement of stillbirth but unclear if in line with the UK definition. Only 2 studies provided clear definitions of stillbirth, but neither were aligned with the UK definition.[9, 11] Furthermore, none of the studies specifically reported stillbirths due to placental dysfunction or stratified by gestational age. Though certain exclusions were made by most studies, none of these were stringent enough to narrow the reported outcome to stillbirths resulting from placental dysfunction, or pre-term or term stillbirths.

Bias

Three studies, Ayala 2012, Haddad 2016 and Kingdom 2011, were at a low risk of confounding.[8-10] They were RCTs, with appropriate randomisation and allocation concealment, and treatment arms recruited from the same population and balanced in baseline characteristics. STRIDER was also considered to be at low risk of confounding, although the method and extent of allocation concealment was unclear.[14] Some data dredging was suspected in Ayala 2012, but it was judged that this would not affect the conclusion for the risk of stillbirth in that study.[88] Blinding of women and care providers was used in Ayala 2012[8] and STRIDER,[14] and losses to follow-up in these three studies were considered where possible (a modified ITT analysis was used in Haddad 2016,[9] excluding women who have withdrawn consent for use of their data).

Subtil 2013 was at an unclear risk of bias largely because only a single arm from an RCT was relevant to the review question, and so many areas of the quality assessment checklist could not be assessed based on that arm alone.[11] Von Dadelszen was at a high risk of bias as it was a retrospective case-control study, with cases and controls selected from different populations.[13]

Power

In general, studies either did not report power calculations for stillbirth (Ayala 2012)[88] or were underpowered to detect a change in stillbirth risk (Haddad 2016, Kingdom 2011, STRIDER and

von Dadelszen 2011).[9, 10, 13, 14] Only one arm was included from Subtil 2003, so the power calculation performed for that study was not applicable and it was unclear if the single arm was powered to detect any effect of the intervention on stillbirth.[11, 12]

Results

The risk of stillbirth or livebirth with birthweight <3rd centile in the included studies is detailed in Table 13. Full study results and study details are provided in the extraction tables in Appendix 3, Table 23.

Table 13. Summary of the effectiveness of interventions to prevent stillbirth

	Ayala 2012[8]	Haddad 2016[9]	Kingdom 2011[10]	Sharp 2018 (STRIDER)[14]	Subtil 2003a[11] Subtil 2003b[12]	von Dadelszen 2011[13]
Intervention	ASA 100 mg daily (N=176)	Enoxaparin + ASA 100 mg daily (N=124)	UFH, 7500UI twice daily (N=16)	Sildenafil citrate 25 mg thrice daily (N=70)	ASA 100 mg daily (N=239)	Sildenafil citrate 25 mg thrice daily (N=10)
Comparator	Placebo (N=174)	ASA 100 mg daily (N=125)	Standard care (N=16)	Placebo (N=65)	NA – single arm relevant only	No sildenafil citrate (N=17)
Effect on risk of stillbirth	Stillbirth rate (95% CI): 1.1 (−0.4 to 2.7) in ASA 2.9 (0.4 to 5.4) in placebo No difference between arms (p=0.246)	Enoxaparin + ASA: 1/122 ASA: 3/122 No difference between arms (p=0.62)	UFH: 0/16 Standard care: 1/16 No statistical analysis provided	Fetal death ≥26 weeks' gestation Sildenafil citrate: Fetal 3/70 (4%) Placebo: 2/56 3%) No difference between arms (RR 1.50, 95% CI 0.27 to 8.34, p=0.64)	No stillbirths occurred	Sildenafil citrate: 3/10 No sildenafil: 6/17† No statistical analysis provided.
Effect on risk of severe SGA	NR	Enoxaparin + ASA: 15/122 ASA: 21/122 No difference between arms p=0.35	UFH: 4/16 Standard care: 5/16 No difference between arms p=1.000	NR	4.6% women had a baby with SGA <3rd centile	NR

†5 stillbirths resulting from late terminations have not been included here
Abbreviations: ASA: aspirin; CI: confidence interval; FGR: fetal growth restriction; NA: not applicable; NR: not reported; RR: relative risk; SGA: small for gestational age; UFH: unfractionated heparin.

In Von Dadelszen 2011 women were specifically identified as having early-onset IUGR and the median age at delivery was between 25^{+6} and 27^{+1} days, the results are thus likely be more relevant to pre-term than term stillbirths.[13] Nevertheless, no statistical analysis was provided, and the study was at a high risk of bias, precluding drawing any conclusions as to the effectiveness of sildenafil citrate for prevention of pre-term stillbirths.[13] Similarly to Von Dadelszen 2011, the STRIDER study was conducted in women with severe early-onset IUGR at between 22^{+0} and 29^{+6} weeks' gestation, thus results for the effectiveness of sildenafil citrate in preventing stillbirth is

likely to be most relevant for pre-term stillbirths, although no significant difference was detected between the study groups (for fetal death ≥26 weeks' gestation).[14]

In Kingdom 2011, women were identified in their second trimester (18^{+0} to 23^{+6} weeks) based on having placental dysfunction, which may make the results more relevant to pre-term stillbirths.[10] However, in 31.5% women the delivery occurred before 32 weeks, potentially preventing both pre-term and term stillbirths, given that only one stillbirth was recorded but severe SGA was more common. No significant difference between severe SGA was found, but the study was underpowered, so no conclusions can be drawn.[10]

Women in Subtil 2003 were identified as at-risk during their Doppler scan at 22 to 24 weeks, which may be more likely to detect early-onset dysfunction.[11] However, no stillbirths were reported in that arm of the study, which was unlikely to be due to earlier delivery as only 8.3% births occurred <37 weeks. Interestingly, women who had normal Doppler results and who had not received ASA were still at a lower risk of induced delivery<37 weeks, indicating that ASA was not sufficient to avoid elective birth.[11]

In Ayala 2012 and Haddad 2016, women were recruited earlier in pregnancy (≤16 and <14 weeks respectively) but it is unclear if that would have prioritised the detection of early-onset placental dysfunction, as some selection was based on medical history, rather than measured risk.[8, 9] Equally, it is unclear whether results are applicable to pre-term or term stillbirths; the majority of deliveries were at term and without stratification by gestational age, risk of pre-term and term stillbirths cannot be ascertained.[8, 9] Also unclear was how many pregnancies were delivered by elective birth and or whether the interventions have any effect on avoiding the use of elective birth to prevent stillbirth later on in the pregnancy.[8, 9]

Conclusions

Even among pregnancies at risk, stillbirth was a considerably rare event, and few events occurred among the small cohorts enrolled into the studies; this increased the uncertainty around the outcome. Statistical analysis was only reported in 2 of the 5 studies and it was concluded that there was no difference to the risk of stillbirth between the intervention and control arms.[8, 9] It is also worth noting that these studies were not powered to detect a difference in stillbirth or severe SGA (thus, lack of an effect does not imply that no effect exists; only that the study was unable to demonstrate it).

Based on the evidence found by this review, it is not possible to ascertain the effectiveness of interventions to prevent pre-term or term stillbirths or stillbirth overall. Without further studies, no intervention can be recommended as effective or preferable to elective birth.

Question 4 – How effective is elective caesarean section or induction of labour to prevent stillbirth in pregnancies at risk?

a) Before 37 completed weeks of gestation

b) From 37 completed weeks of gestation

Eligibility for inclusion in the review

This review searched for randomised controlled trials, interventional studies, cohort and case-control studies as well as SLRs and meta-analyses of the above. Studies were included if the population comprised pregnant women at a high risk of pre-term or term stillbirth due to placental dysfunction, determined through screening, testing or due to having a specific maternal risk factors. Once identified, women could be managed either with planned delivery (by induction of labour or planned CS) or managed expectantly to prevent stillbirth caused by placental insufficiency. Publications were only included if they reported risk of pre-term or term stillbirth or severe SGA (<3rd centile). Studies had to be conducted in the UK or other high-income countries.

Description of the evidence

This review found 3 studies in 3 unique cohorts that reported on the risk of stillbirth upon induction of labour compared with expectant management.[15-17] Two of the 3 included studies were RCTs – the DIGITAT trial and Walker 2016,[15, 17] whereas Rabinovich 2018 was a retrospective cohort study.[16] Only Walker 2016 was conducted in the UK;[17] Rabinovich 2018 took place in Israel,[16] and DIGITAT in the Netherlands.[15] The cohorts included in the two RCTs were of similar size – each randomised just over 650 participating women.[15, 17] The sample size in the retrospective study was larger, totalling 2,232 pregnancies.[16] All studies compared induction of labour with expectant management (EM), and though induction was split into induction by elective CS and induction "by other methods" in Rabinovich 2018, the statistical comparison appears to have been conducted using a combined "induction of labour" group.[16] Rabinovich 2018 was the only study where results were reported separately for late pre-term and early term.[16] Severe SGA was only reported in DIGITAT.[15]

Discussion of findings

Quality assessment

The quality of studies investigating the effectiveness of planned delivery or expectant management in preventing stillbirth was assessed using an adapted Downs and Black checklist. A summary is presented in Table 14, whereas the full assessment can be found in Appendix 3, Table 26.

Table 14. Risk of bias in studies on planned delivery versus expectant management

Question	DIGITAT[15]	Rabinovich 2018[16]	Walker 2016[17]
EXTERNAL VALIDITY	Unclear	High	High
CONFOUNDING	Unclear	High	Low
POWER	Unclear	Unclear	High

External validity

Women were selected from low-risk cohorts in all three studies. Both DIGITAT and Rabinovich 2018 included women at an increased risk of stillbirth due to having a suspected growth restriction.[15, 16] External validity was unclear in DIGITAT; it was conducted in the Netherlands and included a population that could be expected to be similar to that of the UK, however, stillbirth definition was not reported so it was unclear whether the definition used was the same as in the UK and severe SGA was based on Dutch growth charts, which are likely to be different to those used in the UK.[15] In Rabinovich 2018, there was a high concern about the applicability of the results, as that study was conducted in Israel and it appears that there may be some ethnic differences between the included cohort and the UK population.[16] Further, though stillbirth definition was not reported, the study only included women between $34^{0/7}$ and $38^{6/7}$ gestational weeks, thus only stillbirth from 34 weeks of gestational would have been included in the outcome, which is different from the UK definition.[16] Concern about external validity was also high for Walker 2016 as the recruited women were considered high-risk due to being nulliparous and of advanced maternal age (≥35 years).[17] While both are risk factors for stillbirth, it is unlikely that any screening programme would rely solely on the combination of the two to select women at risk for stillbirth. It is thus likely that the cohort included in this trial included some women who would be considered low-risk and some who would be considered high-risk by a potential screening programme.[17]

Confounding

Risk of confounding was low only in Walker 2016.[17] Though the trial was not blinded, the randomisation was performed appropriately with allocation sequence concealment. Baseline characteristics were balanced between arms and no data dredging was evident. Use of a modified intention to treat analysis was unlikely to have affected the results. Given that stillbirth was a pre-specified outcome, it is also highly likely that the staff were appropriately trained to assess the outcome reliably.

DIGITAT was judged to be at an uncertain risk of bias as it was unclear how stillbirth was defined and whether it was ascertained reliably, given that it was not a pre-specified outcome.[15] In addition, there was some indication of data dredging, and a small imbalance between the baseline characteristics in the two arms. Randomisation was performed appropriately with allocation

concealment and while the ITT analysis was not appropriate for the main hypothesis (as this was a non-inferiority study), it was unlikely to bias the analysis for risk of stillbirth or severe SGA.

Rabinovich 2018 was judged to be at high risk of confounding.[16] This was a retrospective study with significant imbalances in characteristics of women receiving different interventions. Moreover, given the retrospective design, there was a risk that stillbirth was not diagnosed or recorded reliably in the hospital records.

Power

In the DIGITAT study, the authors reported there was sufficient power to reject the null hypothesis, but it is unclear (and unlikely) that this was also the case for stillbirth and severe SGA, given that these were not pre-specified outcomes.[15] In Rabinovich 2018 the authors reported that the study was sufficiently powered; however, no calculations were provided to support that claim.[16] In Walker 2016 it was reported that the study was underpowered to detect differences in stillbirth (severe SGA was not reported).[17]

Results

A summary of the effectiveness of planned delivery compared with EM for the three studies relevant to question 4 is presented in Table 15. Fully extracted details for each study are presented in Appendix 3, Table 23.

Table 15. Effectiveness of planned delivery vs expectant management in preventing stillbirths in high-risk women

	DIGITAT[15]	Rabinovich 2018[16]			Walker 2016[17]
		Overall	Early term (37 0/7 to 38 6/7)	Late pre-term (34 0/7 to 36 6/7)	
Planned delivery	Labour induction (N=321)	Labour induction (N=1,428) (348 by elective CS, 1,080 by other induction methods)	Induction (N=951)	Induction (N=477)	Labour induction (N=304)
Expectant management (EM)	EM (N=329)	EM (N=804)	EM (N=512)	EM (N=292)	EM (N=314)
Effect on risk of stillbirth	No stillbirths in either arm	Elective CS: 0.3% stillbirths Other induction: 0.6% stillbirths EM: 1.5% stillbirths p=0.042[a]	Induction: 0.3% stillbirths EM: 0.6% stillbirths p=0.428	Induction: 0.8% stillbirths EM: 3.1% stillbirths p=0.001	No stillbirths in either arm
Effect on risk of severe FGR	18.1% fewer births <3rd centile in induction of labour compared with EM (p<0.001)	NR	NR	NR	NR
Additional outcomes					

	DIGITAT[15]	Rabinovich 2018[16]			Walker 2016[17]
		Overall	**Early term (37 0/7 to 38 6/7)**	**Late pre-term (34 0/7 to 36 6/7)**	
Neonatal death, %	No neonatal deaths in either arm	Elective CS: 0.9 Other induction: 0.2 EM: 2 p<0.001	Induction: 0.1 EM: 0.6 p=0.126	Induction: 0.8 EM: 4.5 p=0.039	NR
Apgar score <5 at 1 minute	NR	**Apgar <5, %** Elective CS: 2.6 Other induction: 1.6 EM: 5.8 p<0.001	**Apgar <5, %[b]** Induction: 1.4 EM: 3.8 P=0.003	**Apgar <5, %[b]** Induction: 2.8 EM: 9.4 p<0.001	NR
Apgar score at 5 minutes	**Apgar <7, n (%)** Induction of labour: 7 (2.2) EM: 2 (0.6) Difference in %: 1.6 (–0.2 to 3.4)	**Apgar <7, %** Elective CS: 0.9 Other induction: 0.3 EM: 1.9 p=0.002	**Apgar <7, %[c]** Induction: 0.3 EM: 1 p=0.174	**Apgar <7, %[c]** Induction: 0.6 EM: 3.6 p=0.006	**Apgar <4, n** Induction: 0 EM: 1 RR or p-value NR **Apgar 4 to 7, n** Induction: 11 EM: 11 RR 1.04 (95% CI: 0.40 to 2.69), p=0.94
Admission to NICU	**Admission to intensive care, n (%)** Induction of labour: 9 (2.8) EM: 13 (4.0) Difference in %: −1.2 (−4.0 to 1.6)	NR	NR	NR	**Admission to NICU for >4 days, n** Induction: 6 EM: 7 RR 0.88 (95% CI: 0.26 to 3.06), p=0.85
Length of stay in NICU (days)	Induction of labour: 9 (6 to 14) EM: 13 (6 to 22) p=0.2	**Neonatal hospitalisation length** Elective CS: 5 Other induction: 3 EM: 3 p<0.001	**Neonatal hospitalisation length** Induction: 3 EM: 2 p<0.001	**Neonatal hospitalisation length** Induction: 6 EM: 6 p=0.532	NR
Prematurity complications	NR	Elective CS: 3% Other induction: 1% EM: 2% p=0.032	Induction: 0.7% EM: 0.2% p=0.434	Induction: 3% EM: 5.3% p=0.136	NR
Composite adverse neonatal outcome	**Composite adverse neonatal outcome, n (%)** Induction of labour: 17 (5.3) EM: 20 (6.1) Difference in %: −0.8 (−4.3 to 2.8)	NR	NR	NR	**Hypoxia, n** Induction: 2 EM: 2 RR 1.03 (95% CI: 0.14 to 7.50), p=0.98 **Hypotonia ≥2 h, n** Induction: 1

	DIGITAT[15]	Rabinovich 2018[16]			Walker 2016[17]
		Overall	Early term (37 0/7 to 38 6/7)	Late pre-term (34 0/7 to 36 6/7)	
					EM: 0 RR or p value NR
Neonatal hypoglycaemia	NR	Elective CS: 4.5% Other induction: 5.5% EM: 7% p=0.212	Induction: 4.5% EM: 6.7% p=0.088	Induction: 6.7% EM: 7.8% p=0.617	NR
Neonatal sepsis, %	NR	Elective CS: 0.3 Other induction: 0.3 EM: 0.6 p=NS	Induction: 0.3 EM: 0.2 p=1.00	Induction: 0.2 EM: 1.2 p=0.123	NR

Abbreviations: CI: confidence interval; CS: Caesarean section; EM: expectant management; FGR: fetal growth restriction; NICU: neonatal intensive care unit; NR: not reported; NS: not significant; RR: risk ratio.

[a]It appears that the significant p value (p=0.042) is for comparison between EM and induction of labour as one group; however, this is not clearly reported, and the comparison may have been made between EM and either the induction of labour by elective CS or induction by other means, or between each of the three groups. [b]Outcome stated in publication as 'Initial APGAR score <5 (%)' so it is unclear whether this refers to Apgar score at 1 minute. [c]Outcome stated in publication as 'Subsequent APGAR score <7 (%)' so it is unclear whether this refers to Apgar score at 5 minutes.

No conclusions on the effectiveness of planned delivery versus EM can be drawn from the two included RCTs, as no stillbirths occurred in either of these studies and both were definitely or highly likely underpowered to detect a difference in the risk of stillbirth.[15, 17] Furthermore, it appears that few to no pre-term deliveries were recorded in these RCTs, indicating lack of stillbirth is unlikely to be a result of earlier delivery. It is interesting to note that a significant difference was noted in the proportion of severe SGA in the DIGITAT study, with fewer babies being born with a birthweight <3rd centile for their gestational age when labour was induced compared to when the pregnancy was managed with EM.[15] Furthermore, no neonatal deaths occurred in the study, and no significant differences were detected between planned delivery and EM for Apgar score, NICU admission and composite adverse neonatal outcome. Walker 2016 also reported no significant differences in Apgar score, admission to NICU or composite adverse outcomes of hypoxia or hypotonia. However, it is difficult to draw any conclusions from these results, as the DIGITAT and Walker 2016 studies were not sufficiently powered to detect a statistical difference.

Risk of stillbirth was reported to be reduced in the retrospective Rabinovich 2018 study, when EM was compared with labour induction by elective CS or other methods.[16] Rabinovich 2018 reported that induction of labour had a significant effect on lowering the risk of stillbirth in the late pre-term group (34 up to 36 weeks of gestation) but not in the early term group (37 up to 39 weeks of gestation). This is potentially encouraging for the prevention or pre-term stillbirths, as it also appears that neonatal outcomes such as Apgar score, fetal distress and very low birth weight are improved in the induction of labour group in the late pre-term period compared with EM. There were also no significant differences in the rates of neonatal hypoglycaemia, neonatal sepsis, or

prematurity complications between induction of labour and EM, suggesting that inducing labour between $34^{0/7}$ and $36^{6/7}$ weeks does not carry additional risk for these outcomes compared with EM. However, the study lacks a comparison on how the outcomes of the premature infants compare with those born at term, i.e. it is unclear what the added burden due to prematurity is in these babies. The result should also be considered with caution as stillbirths prior to the 34^{th} gestational week were not included in the study, thus a period where the consequences of prematurity may outweigh the benefit of preventing stillbirth through delivery, resulting in induction appearing more beneficial. It is also not applicable to the entire cohort of pregnant women, who may be at risk of stillbirth prior to 34 weeks. Moreover, the retrospective design of the study, and the imbalance in baseline characteristics between women receiving different interventions, adds further uncertainty to this result, which cannot be resolved without further high-quality research in the UK setting.[16]

Conclusions

Stillbirths were only reported in one of the three identified studies and it appears that induction of labour may be beneficial for preventing pre-term but not term stillbirths. However, the poor quality of that study precludes drawing any definite conclusions. Thus, the effectiveness of elective birth in preventing pre-term or term stillbirths hashas not been demonstrated. However, it needs to be noted that elective birth prevents stillbirths by definition (as long as no intrapartum stillbirths occur). To identify babies that may have resulted in stillbirth had they been delivered later, severe SGA was investigated as an outcome. In the DIGITAT study significantly fewer babies were born with a birthweight $<3^{rd}$ centile for their gestational age when labour was induced near term (36 to 41 weeks) than with EM.[15] If stillbirth/severe SGA are prevented, the baby may still be at risk of other adverse perinatal outcomes, particularly if the baby is delivered pre-term. In late pre-term pregnancies, those who were induced had better outcomes in terms of neonatal death, fetal distress and Apgar score than those who received EM, providing plausibility that induced delivery may be an acceptable intervention for the prevention of stillbirth after 34 weeks' gestation. Nevertheless, the effectiveness of elective delivery to prevent pre-term stillbirth before 34 weeks' gestation, with consideration of the risk of other adverse outcomes associated with pre-term delivery, remains unclear. Overall, due to the poor quality and targeted scope of the evidence considered in this review, the effectiveness and safety of induced delivery for the prevention of pre-term or term stillbirth cannot currently be ascertained.

Summary of Findings Relevant to Criteria 9 and 10: Criteria not met

Quantity: The evidence base to assess criteria 9 and 10 was small, consisting of 9 studies; 3 evaluated the risk of stillbirth in women undergoing planned delivery compared with expectant management and 6 studies reported on the effectiveness of pharmacological interventions aiming to reduce stillbirth risk.

Quality: Studies were of mixed quality and mostly at either uncertain or high risk of bias. Specifically, there was a high risk of ascertainment bias as the reliability of diagnosing stillbirth could not be judged with little or no detail reported on the method of diagnosis or how stillbirth was defined. Otherwise, all studies were either reported to be underpowered or likely to be underpowered to assess the effect of the intervention on the risk of stillbirth, thus preventing any conclusions of equivalence being drawn.

Applicability: In the majority of studies there was some concern about applicability to the review question. This was mostly due to the lack of stratification in the reported stillbirth risk by pre-term and term, and to the use of a definition of stillbirth that was not applicable to the UK setting. In addition, in 7 out of 9 studies there was a concern that the included cohort of women was not representative of women at risk of stillbirth selected through screening due to the method of participant selection as well as the baseline characteristics. Conversely, there was no concern about the interventions tested, as each is available as part of antenatal care in the UK for other certain pregnancy conditions.

Consistency: Three studies compared planned delivery with expectant management. There were no stillbirths in two of the three studies, and as these were underpowered for this outcome, it is not possible to make a comparison between the results. Three other studies evaluated low-dose (100 mg/day) aspirin; one compared aspirin with placebo, one compared aspirin with aspirin plus enoxaparin and one had no comparator arm. There were no differences between the intervention arms in either of the 2 RCTs and no stillbirths occurred in the single relevant arm of the third study. Given that these studies were also underpowered for stillbirth, it is not possible to assess consistency between the results.

Conclusions: None of the pharmacological intervention were demonstrated to be effective at reducing risk of stillbirth, either overall or pre-term or term, compared with standard care or no treatment. There is some evidence that planned delivery is effective at reducing the risk of pre-term stillbirths with improved perinatal outcomes, though this is based on a single retrospective study at high risk of bias and with concerns about applicability to the UK setting and 2 RCTs with no stillbirth events. Thus, there is currently insufficient evidence that any intervention could reduce a screen-detected risk of stillbirth.

Review summary

Conclusions and implications for policy

Based on the synthesis of evidence against the UK NSC criteria, screening of pregnant women to prevent stillbirths due to placental dysfunction is not recommended.

The main reason for this is the lack of an accurate screening test that can specifically identify women at risk of either pre-term or term stillbirth due to placental dysfunction. Very few screening studies reported measures of test accuracy specifically for pre-term and term stillbirths. When the distinction was made, tests were generally better at predicting pre-term stillbirth. This may be because pre-term stillbirths are more likely to result from early-onset placental dysfunction, and as the majority of the tests are conducted in the first and second trimester, they miss pregnancies where late-onset placental dysfunction has not yet developed. Alternatively, it may be that the tests are better at detecting pregnancies with severe placental dysfunction, which could be more likely to result in stillbirth before term is reached. Overall, a moderate number of good-quality studies have been identified, including some conducted in the UK, but none of the screening tests examined were sensitive and specific enough to be recommended for use in a screening programme, using the thresholds that are widely used for evaluating the accuracy of diagnostic tests (LR+ >10 and LR− <0.1). It may be that surpassing these thresholds is not achievable when screening for a condition as complex as stillbirth, and therefore may not be appropriate for identifying tests that could be used clinically for this purpose. In clinical practice, women are routinely induced for risk factors that are also unlikely to surpass these thresholds. For example, NICE recommend that women with pre-existing diabetes mellitus undergo elective delivery before 39 weeks' gestation.[98] It is also important to consider the high risk of ascertainment bias in the screening studies included in this review; investigators or health providers were not known to be blinded to the index test results in any study, and therefore could have intervened with pregnancies suspected to be at high risk of stillbirth. By preventing some cases of stillbirth, accuracy of the screening test may have been underestimated. This is considered to be a limitation of the evidence base for screening tests for stillbirth; blinding is often incomplete in screening studies, as both the patient and health provider need to know the test result.[99]

A significant gap in evidence is the absence of studies focusing on growth monitoring, rather than size measurement, as a screening tool. The review did identify studies screening for risk of stillbirth using growth charts to estimate fetal weight. However, the observed test accuracies were low, with sensitivity not reaching above 60%, at a specificity of only 60%.[39] The authors of one study acknowledge that customised growth standards are used to facilitate the identification of infants growing in a suboptimal way, but are not diagnostic tests.[39] Furthermore, there appears to

be some disagreement over the choice of fetal growth charts to screen for FGR to predict adverse outcomes such as stillbirth, which seems to stem from a substantial variation in these growth charts due to the heterogeneous study populations that they are derived from.[31] Given that growth monitoring is considered to be more specific for identifying pregnancies at risk of pre-term stillbirth, there is a need for high-quality, prospective studies investigating growth potential as a screening test, but these studies would need to report results separately for pre-term and term stillbirths.

Further reasons for not recommending a screening programme for stillbirth were insufficient good-quality evidence on appropriate monitoring strategies and interventions that could be used in screen-identified high-risk pregnancies to prevent stillbirth; this includes very limited, low-quality evidence on effectiveness of elective birth in preventing pre-term or term stillbirths.While it is likely that there is evidence to support that elective birth is safe at term, this evidence is not directly applicable to women at risk of stillbirth.

Between the monitoring strategies currently in place for pregnancies at risk of adverse perinatal outcomes (fetal movement counting, Doppler ultrasound, cardiotocography), it is possible that all of these strategies reduce the risk of stillbirth, although it is unclear which monitoring regime is most effective, and if there is a difference between pre-term and term stillbirth. A limitation of studies that prospectively evaluate monitoring strategies is lack of a "no monitoring" control arm, as it would be unethical not to monitor high-risk pregnant women to guide the decision on when to deliver the baby. Findings in one study did suggest that fetal movement counting awareness was associated with a decreased risk of stillbirths;[96] however, the high risk of bias and concerns about applicability to the UK prevent conclusions from being drawn. As previously discussed, unblinded studies of monitoring strategies are at risk of ascertainment bias. The question cannot be addressed without further research comparing suggested monitoring strategies with current clinical practice in the UK, and differentiating outcomes for pre-term and term stillbirth risk reduction.

This review did not find sufficient evidence to support any changes to the recommendation for elective birth, which is the only intervention currently used in clinical practice to prevent stillbirths in pregnancies that would be considered at-risk by this review.[18] Aspirin was the most commonly tested intervention, but its use was not shown to be beneficial for the prevention of stillbirths. Data on whether the interventions helped to avoid elective birth were lacking.

In terms of the effectiveness of elective birth, the benefit of avoiding stillbirth needs to be considered against the risks associated with prematurity in the decision-making for planned delivery. The only study that examined elective birth during the pre-term period was of poor quality and did not allow for conclusions on the burden of pre-term delivery to be drawn. For term pregnancies, although this review only included women at risk of stillbirth, it is important to recognise that delivery has been shown to reduce the risk of both perinatal death and stillbirth in the unselected population. For example, the ARRIVE study, conducted in low-risk nulliparous

pregnancies, found no significant difference in frequency of a composite adverse perinatal outcome between induction of labour at 39 weeks and the expectant-management group.[19] Similarly, a SLR reported that a policy of labour induction was associatedwith fewer all-cause perinatal deaths than a policy of expectant management.[100]

Furthermore, elective delivery is recommended by guidelines for other high-risk pregnancy conditions not studied in this review. For example, NICE guideline NG3 advises pregnant women with type 1 or type 2 diabetes and no other complications to have an elective birth by induction of labour, or by elective caesarean section if indicated, between 37^{+0} weeks and 38^{+6} weeks of pregnancy.[101] Clinical guideline CG107 recommends elective early birth to women with refractory severe gestational hypertension and women who have pre-eclampsia with mild or moderate hypertension at 34^{+0} to 36^{+6} weeks, depending on maternal and fetal condition, risk factors and availability of neonatal intensive care.[102]

The evidence base in this wider context is thus much more extensive than the evidence base considered for the present review. The context of the current review was to study the effect of screening and intervention on the risk of stillbirth. Hence, when the evidence around interventions was considered, it was confined to women who were considered at high risk for stillbirth following screeing.[19] While these observations do not alter the primary conclusion that screening and intervention for stillbirth is not currently justified, they do suggest that future research might focus on screening and intervention to prevent stillbirths at term, where there is an intervention that is effective in other contexts and would not be expected to lead to harm through iatrogenic prematurity.

Another limitation of the evidence is an inherent risk of bias in the measurement of stillbirth test accuracy or intervention effectiveness in studies where women and treating obstetricians were not blinded to the test results. Knowledge of the pregnancy being at high risk could prompt interventions or more careful pregnancy monitoring; a good test could paradoxically lead to an underestimation of accuracy. In an effort to minimise this, this review used a combined reference standard where babies could have been either stillborn or liveborn with severe SGA, i.e. $<3^{rd}$ centile for gestational age. Severe SGA was also considered a relevant outcome in studies of monitoring and interventions. However, it is noted that this approach is also not free of bias, in that it could also lead to an overestimation of test accuracy/intervention efficacy, as only a proportion of severe SGA babies will represent "prevented" stillbirths, with the remaining part of that group representing babies that would have always been liveborn. Nevertheless, no studies were found where the reference standard included both stillborn and liveborn babies. Conversely, this rapid review did not include studies that evaluated index tests that aimed to predict SGA $<3^{rd}$ centile. While some of these tests may also be applicable to predicting stillbirth, and inclusion of such studies may have expanded the evidence base, looking to prevent stillbirth by predicting severe SGA would be a separate screening strategy that is not within the scope of the current review.

Limitations

This section considers limitations of the review methodology. Limitations of the evidence and evidence gaps are discussed in the section above.

This rapid review was conducted in line with the UK NSC requirements for evidence summaries, as described at https://www.gov.uk/government/publications/uk-nsc-evidence-review-process/appendix-f-requirements-for-uk-nsc-evidence-summaries. All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 31 in Appendix 5.

Searches of multiple databases were conducted (see Appendix 1). Database search terms were restricted by study design and interventions, and limited to studies published since 2000. However, it is unlikely that major important studies were missed.[103, 104] A published and well validated filter was used to limit by study design,[103, 104] searches were supplemented with SLR reference list searches, and expert clinical opinion was sought on the completeness of the list of relevant records identified.

Included publication types

This review only included peer-reviewed journal publications, and excluded publications that were not peer-reviewed and grey literature. This may have led to the exclusion of relevant evidence. However, this is an accepted methodological adjustment for a rapid review and is unlikely to miss any pivotal studies.

For question 1, publications were excluded if they only presented data allowing the calculation of test accuracy parameters. This was taken as a pragmatic approach and was unlikely to result in key screening studies being missed.

Language

Only studies published in English were included. Given that this review was focusing on evidence relevant to the UK setting, this limitation should not have led to the exclusion of any pivotal studies.

Review methodology

Articles were reviewed by a single reviewer in the first instance. A second reviewer examined all included articles, 10% of excluded articles, and any articles where there was uncertainty about inclusion. This pragmatic strategy should have minimised the risk of errors.

Articles not freely available

Searches for full-text articles were carried out at Cambridge University Library. Articles that were not freely available at this library were re-evaluated and those judged highly likely to report pivotal data were requested from authors or purchased. Any remaining unavailable articles were excluded, as it was judged that they would not contain any additional pivotal data from relevant populations that would affect the conclusions of this review.

References

1. NHS England. Saving Babies' Lives: A Care Bundle for Reducing Stillbirth, 2016.
2. Department of Health. The Government's mandate to NHS England for 2016-17, 2015.
3. National Maternity Review. Better Births: Improving outcomes of maternity services in England. 2015.
4. Goldenberg RL, McClure EM, Bhutta ZA, et al. Stillbirths: the vision for 2020. Lancet 2011;377:1798-805.
5. Clifford S GS, Southam M, Williams M, Gardosi J,. The Growth Assessment Protocol: a national programme to improve patient safety in maternity care. 2013.
6. Goldenberg RL, McClure EM, Belizan JM. Commentary: reducing the world's stillbirths. BMC Pregnancy Childbirth 2009;9 Suppl 1:S1.
7. The Lancet. Ending Preventable Stillbirths: An Executive Summary for The Lancet's Series. 2016.
8. Ayala DE, Ucieda R, Hermida RC. Chronotherapy with low-dose aspirin for prevention of complications in pregnancy. Chronobiology international 2013;30:260-279.
9. Haddad B, Winer N, Chitrit Y, et al. Enoxaparin and Aspirin Compared With Aspirin Alone to Prevent Placenta-Mediated Pregnancy Complications: A Randomized Controlled Trial. Obstetrics & Gynecology 2016;128:1053-1063.
10. Kingdom JCP, Walker M, Proctor LK, et al. Unfractionated heparin for second trimester placental insufficiency: A pilot randomized trial. Journal of Thrombosis and Haemostasis 2011;9:1483-1492.
11. Subtil D, Goeusse P, Houfflin-Debarge V, et al. Randomised comparison of uterine artery Doppler and aspirin (100 mg) with placebo in nulliparous women: The Essai Regional Aspirine Mere-Enfant study (Part 2). BJOG: An International Journal of Obstetrics and Gynaecology 2003;110:485-491.
12. Subtil D, Goeusse P, Puech F, et al. Aspirin (100 mg) used for prevention of pre-eclampsia in nulliparous women: The Essai Regional Aspirine Mere-Enfant study (Part 1). BJOG: An International Journal of Obstetrics and Gynaecology 2003;110:475-484.
13. Von Dadelszen P, Dwinnell S, Magee LA, et al. Sildenafil citrate therapy for severe early-onset intrauterine growth restriction. BJOG: An International Journal of Obstetrics and Gynaecology 2011;118:624-628.
14. Sharp A, Cornforth C, Jackson R, et al. Maternal sildenafil for severe fetal growth restriction (STRIDER): a multicentre, randomised, placebo-controlled, double-blind trial. Lancet Child Adolesc Health 2018;2:93-102.
15. Boers K, Vijgen S, Bijlenga D, et al. Induction versus expectant monitoring for intrauterine growth restriction at term: randomised equivalence trial (DIGITAT). BMJ (clinical research ed.). Volume 341, 2010:c7087.
16. Rabinovich A, Tsemach T, Novack L, et al. Late preterm and early term: when to induce a growth restricted fetus? A population-based study. Journal of Maternal-Fetal and Neonatal Medicine 2018;31:926-932.
17. Walker K, Bugg G, Macpherson M, et al. Randomized Trial of Labor Induction in Women 35 Years of Age or Older. New England journal of medicine. Volume 374, 2016:813-822.

18. GreenTop31. Small for Gestational age Fetus: Investigation and Management. Royal College of Obstetricians Gynaecologists 2013.
19. Grobman WA, Rice MM, Reddy UM, et al. Labor induction versus expectant management in low-risk nulliparous women. New England Journal of Medicine 2018;379:513-523.
20. Office for National Statistics. User Guide to Child Mortality Statistics. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetochildmortalitystatistics, 2018.
21. Statistics Canada. Data quality, concepts and methodology: Definitions. Available at: https://www150.statcan.gc.ca/n1/pub/84f0210x/2008000/technote-notetech1-eng.htm. Accessed: 13.08.18, 2012.
22. Australian Institute of Health and Welfare. Perinatal deaths in Australia 2013-2014. Available at: https://www.aihw.gov.au/getmedia/78784f2e-2f61-47ea-9908-84b34441ae0a/aihw-per-94.pdf.aspx?inline=true, 2018.
23. New Zealand Ministry of Health. Miscarriage and Stillbirth. Available at: https://www.health.govt.nz/your-health/pregnancy-and-kids/services-and-support-during-pregnancy/miscarriage-and-stillbirth. Accessed: 13.08.18, 2017.
24. WHO. Maternal, Newborn, Child and Adolescent Health: Stillbirths, 2018.
25. Flenady V, Wojcieszek AM, Middleton P, et al. Stillbirths: recall to action in high-income countries. Lancet 2016;387:691-702.
26. Lawn JE, Yakoob MY, Haws RA, et al. 3.2 million stillbirths: epidemiology and overview of the evidence review. BMC Pregnancy Childbirth 2009;9 Suppl 1:S2.
27. Office for National Statistics. Child Mortality Statistics, 2016, 2018.
28. Flenady V, Froen JF, Pinar H, et al. An evaluation of classification systems for stillbirth. BMC Pregnancy Childbirth 2009;9:24.
29. Haws RA, Yakoob MY, Soomro T, et al. Reducing stillbirths: screening and monitoring during pregnancy and labour. BMC Pregnancy Childbirth 2009;9 Suppl 1:S5.
30. Silver RM. Examining the link between placental pathology, growth restriction, and stillbirth. Best Pract Res Clin Obstet Gynaecol 2018;49:89-102.
31. Thilaganathan B. Opinion: Ultrasound fetal weight estimation at term may do more harm than good. Ultrasound Obstetrical Gynaecology 2018;52:5-8.
32. Thilaganathan B. Ultrasound fetal weight estimation at term may do more harm than good. 2018.
33. Dunsworth HM, Warrener AG, Deacon T, et al. Metabolic hypothesis for human altriciality. Proc Natl Acad Sci U S A 2012;109:15212-6.
34. Baschat AA, Hecher K. Fetal growth restriction due to placental disease. Semin Perinatol 2004;28:67-80.
35. National Institute for Health and Clinical Excellence. Antenatal care: routine care for the healthy pregnant woman. Available at: www.nice.org.uk/guidance/cg62/evidence/evidence-tables-from-the-2003-version-196748322 [accessed 21st March 2018]. 2008.
36. Lubow JM, How HY, Habli M, et al. Indications for delivery and short-term neonatal outcomes in late preterm as compared with term births. Am J Obstet Gynecol 2009;200:e30-3.
37. Gardosi J, Madurasinghe V, Williams M, et al. Maternal and fetal risk factors for stillbirth: population based study. Bmj 2013;346:f108.
38. Royal College of Obstetricians & Gynaecologists. The Investigation and Management of the Small-for-Gestational-Age Fetus: Green-top Guideline No. 31. 2nd Edition ed, 2013.

39. Hemming K, Hutton JL, Bonellie S. A comparison of customized and population-based birth-weight standards: the influence of gestational age. Eur J Obstet Gynecol Reprod Biol 2009;146:41-5.
40. Hemming K, Bonellie S, Hutton JL. Fetal growth and birthweight standards as screening tools: methods for evaluating performance. BJOG: An International Journal of Obstetrics & Gynaecology 2011;118:1477-83.
41. Hutcheon JA, Zhang X, Cnattingius S, et al. Customised birthweight percentiles: does adjusting for maternal characteristics matter? Bjog 2008;115:1397-404.
42. Sovio U, White IR, Dacey A, et al. 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;386:2089-2097.
43. Gardosi J, Francis A, Turner S, et al. Customized growth charts: rationale, validation and clinical benefits. Am J Obstet Gynecol 2018;218:S609-s618.
44. Conde-Agudelo A, Bird S, Kennedy SH, et al. First- and second-trimester tests to predict stillbirth in unselected pregnant women: A systematic review and meta-analysis. BJOG: An International Journal of Obstetrics and Gynaecology 2015;122:41-55.
45. Jacob L, Kostev K, Kalder M. Risk of stillbirth in pregnant women with obesity in the United Kingdom. Obes Res Clin Pract 2016;10:574-579.
46. Flenady V, Koopmans L, Middleton P, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. Lancet 2011;377:1331-40.
47. Marufu TC, Ahankari A, Coleman T, et al. Maternal smoking and the risk of still birth: systematic review and meta-analysis. BMC Public Health 2015;15:239.
48. Lamont K, Scott NW, Jones GT, et al. Risk of recurrent stillbirth: systematic review and meta-analysis. Bmj 2015;350:h3080.
49. Schummers L, Hutcheon JA, Bodnar LM, et al. Risk of adverse pregnancy outcomes by prepregnancy body mass index: a population-based study to inform prepregnancy weight loss counseling. Obstet Gynecol 2015;125:133-43.
50. Yerlikaya G, Akolekar R, McPherson K, et al. Prediction of stillbirth from maternal demographic and pregnancy characteristics. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;48:607-612.
51. National Institute for Clinical Excellence. Antenatal Care: Routine Care for the Healthy Pregnant Woman. Clinical Guideline CG62., 2008.
52. National Institute for Clinical Excellence. Inducing labour. Clinical Guideline CG70. 2008.
53. National Institute for Health and Care Excellence. Intrapartum care: care of healthy women and their babies during childbirth. Clinical Guideline CG190.: National Institute for Health and Clinical Excellence, 2007.
54. GreenTop27. Placenta praevia, placenta praevia accreta and vasa praevia: diagnosis and management. London: RCOG 2011:1-26.
55. GreenTop55. Late intrauterine fetal death and stillbirth: October, 2010.
56. England N. Saving Babies' Lives. A care bundle for reducing stillbirth. NHS England 2016.
57. Clifford S GS, Southam M, Williams M, Gardosi J,. The Growth Assessment Protocol: a national programme to improve patient safety in maternity care. 2013.
58. Widdows K RS, Camacho EM, Heazell AEP,. Evaluation of the implementation of the Saving Babies' Lives Care Bundle in early adopter NHS Trusts in England. Manchester, UK: Maternal and Fetal Health Research Centre, University of Manchester, 2018.

59. Neilson JP. Symphysis-fundal height measurement in pregnancy. The Cochrane database of systematic reviews 2000:CD000944-CD000944.
60. Smith G. Prevention of stillbirth. The Obstetrician & Gynaecologist 2015;17:183-187.
61. National Institute for Health and Care Excellence. Caesarean section. Clinical guideline CG132. 2011.
62. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
63. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998;52:377-84.
64. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
65. Akolekar R, Tokunaka M, Ortega N, et al. Prediction of stillbirth from maternal factors, fetal biometry and uterine artery Doppler at 19-24 weeks. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;48:624-630.
66. Akolekar R, Machuca M, Mendes M, et al. Prediction of stillbirth from placental growth factor at 11-13 weeks. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;48:618-623.
67. Aupont JE, Akolekar R, Illian A, et al. Prediction of stillbirth from placental growth factor at 19-24 weeks. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;48:631-635.
68. Bakalis S, Akolekar R, Gallo DM, et al. Umbilical and fetal middle cerebral artery Doppler at 30-34 weeks' gestation in the prediction of adverse perinatal outcome. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2015;45:409-420.
69. Chaiworapongsa T, Romero R, Korzeniewski SJ, et al. Maternal plasma concentrations of angiogenic/antiangiogenic factors in the third trimester of pregnancy to identify the patient at risk for stillbirth at or near term and severe late preeclampsia. American Journal of Obstetrics and Gynecology 2013;208:287.e1-287.e15.
70. Dugoff L, Hobbins JC, Malone FD, et al. First-trimester maternal serum PAPP-A and free-beta subunit human chorionic gonadotropin concentrations and nuchal translucency are associated with obstetric complications: A population-based screening study (The FASTER Trial). American Journal of Obstetrics and Gynecology 2004;191:1446-1451.
71. Dugoff L, Cuckle HS, Hobbins JC, et al. Prediction of patient-specific risk for fetal loss using maternal characteristics and first- and second-trimester maternal serum Down syndrome markers. American Journal of Obstetrics and Gynecology 2008;199:290.e1-290.e6.
72. Dugoff L, Hobbins JC, Malone FD, et al. Quad screen as a predictor of adverse pregnancy outcome. Obstetrics and Gynecology 2005;106:260-267.
73. Mastrodima S, Akolekar R, Yerlikaya G, et al. Prediction of stillbirth from biochemical and biophysical markers at 11-13 weeks. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;48:613-617.
74. Smith GC, Yu CK, Papageorghiou AT, et al. Maternal uterine artery Doppler flow velocimetry and the risk of stillbirth. Obstet Gynecol 2007;109:144-51.

75. Tancrede S, Bujold E, Giguere Y, et al. Mid-Trimester Maternal Serum AFP and hCG as Markers of Preterm and Term Adverse Pregnancy Outcomes. Journal of Obstetrics and Gynaecology Canada 2015;37:111-116.
76. Valino N, Giunta G, Gallo DM, et al. Biophysical and biochemical markers at 30-34 weeks' gestation in the prediction of adverse perinatal outcome. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;47:194-202.
77. Valino N, Giunta G, Gallo DM, et al. Uterine artery pulsatility index at 30-34 weeks' gestation in the prediction of adverse perinatal outcome. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2016;47:308-315.
78. Marttala J, Peuhkurinen S, Laitinen P, et al. Low maternal PAPP-A is associated with small-for-gestational age newborns and stillbirths. Acta Obstetricia et Gynecologica Scandinavica 2014;89:1226-1228.
79. Odibo AO, Cahill AG, Odibo L, et al. Prediction of intrauterine fetal death in small-for-gestational-age fetuses: impact of including ultrasound biometry in customized models. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2012;39:288-292.
80. Poon L, Volpe N, Muto B, et al. Second-trimester uterine artery Doppler in the prediction of stillbirths. Fetal diagnosis and therapy. Volume 33, 2013:28-35.
81. Singh T, Leslie K, Bhide A, et al. Role of second-trimester uterine artery doppler in assessing stillbirth risk. Obstetrics and Gynecology 2012;119:256-261.
82. Smith GC, Shah I, White IR, et al. Maternal and biochemical predictors of antepartum stillbirth among nulliparous women in relation to gestational age of fetal death. BJOG: An International Journal of Obstetrics & Gynaecology 2007;114:705-14.
83. Sutan R, Campbell D, Prescott GJ, et al. The risk factors for unexplained antepartum stillbirths in Scotland, 1994 to 2003. Journal of Perinatology 2010;30:311-8.
84. Trudell AS, Cahill AG, Tuuli MG, et al. Stillbirth and the small fetus: Use of a sex-specific versus a non-sex-specific growth standard. Journal of Perinatology 2015;35:566-569.
85. Trudell AS, Tuuli MG, Colditz GA, et al. A stillbirth calculator: Development & internal validation of a clinical prediction model to quantify stillbirth risk. PLoS ONE 2017;12 (3) (no pagination).
86. Familiari A, Scala C, Morlando M, et al. Mid-pregnancy fetal growth, uteroplacental Doppler indices and maternal demographic characteristics: role in prediction of stillbirth. Acta Obstetricia et Gynecologica Scandinavica 2016;95:1313-1318.
87. Smith NA, Bukowski R, Thomas AM, et al. Identification of pathologically small fetuses using customized, ultrasound and population-based growth norms. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 2014;44:595-599.
88. Chaiworapongsa T, Romero R, Erez O, et al. The prediction of fetal death with a simple maternal blood test at 20-24 weeks: a role for angiogenic index-1 (PlGF/sVEGFR-1 ratio).[Erratum appears in Am J Obstet Gynecol. 2018 Feb 1;:; PMID: 29397906]. American Journal of Obstetrics & Gynecology 2017;217:682.e1-682.e13.
89. Dugoff L, Lynch AM, Cioffi-Ragan D, et al. First trimester uterine artery Doppler abnormalities predict subsequent intrauterine growth restriction. American Journal of Obstetrics and Gynecology 2005;193:1208-1212.

90. Familiari A, Bhide A, Morlando M, et al. Mid-pregnancy fetal biometry, uterine artery Doppler indices and maternal demographic characteristics: Role in prediction of small-for-gestational-age birth. Acta Obstetricia et Gynecologica Scandinavica 2016;95:238-244.
91. Singh R, Deo S, Pradeep Y. The decision-to-delivery interval in emergency Caesarean sections and its correlation with perinatal outcome: Evidence from 204 deliveries in a developing country. Tropical Doctor 2012;42:67-69.
92. McGee S. Simplifying Likelihood Ratios. Journal of General Internal Medicine 2002;17:647-650.
93. Hayden SR, Brown MD. Likelihood ratio: A powerful tool for incorporating the results of a diagnostic test into clinical decisionmaking. Ann Emerg Med 1999;33:575-80.
94. Lees C, Marlow N, Wassenaer-Leemhuis A, et al. 2 year neurodevelopmental and intermediate perinatal outcomes in infants with very preterm fetal growth restriction (TRUFFLE): a randomised trial. Lancet (london, england). Volume 385, 2015:2162-2172.
95. Spaggiari E, Ruas M, Dreux S, et al. Management strategy in pregnancies with elevated second-trimester maternal serum alpha-fetoprotein based on a second assay. American journal of obstetrics and gynecology 2013;208:303. e1-303. e7.
96. Tveit JVH, Saastad E, Stray-Pedersen B, et al. Reduction of late stillbirth with the introduction of fetal movement information and guidelines–a clinical quality improvement. BMC pregnancy and childbirth 2009;9:32.
97. Spaggiari E, Ruas M, Dreux S, et al. Management strategy in pregnancies with elevated second-trimester maternal serum alpha-fetoprotein based on a second assay. American Journal of Obstetrics and Gynecology 2013;208:303.e1-303.e7.
98. NICE. Diabetes in pregnancy: management from preconception to the postnatal period, 2015.
99. Welch HG, Black WC. Evaluating randomized trials of screening. J Gen Intern Med 1997;12:118-24.
100. Middleton P, Shepherd E, Crowther CA. Induction of labour for improving birth outcomes for women at or beyond term. Cochrane Database Syst Rev 2018;5:Cd004945.
101. National Institute for Health and Clinical Excellence. Diabetes in Pregnancy: Management from Preconception to Postnatal Period (NG3), 2015.
102. National Institute for Clinical Excellence. Hypertension in pregnancy: diagnosis and management (CG107): NICE, 2010.
103. Scottish Intercollegiate Guidelines Network. Search Filters (Observational Studies).
104. Scottish Intercollegiate Guidelines Network. Search Filters (Randomised Controlled Trials).
105. Clinicaltrials.gov. NCT01729468 Prevention of Pre-eclampsia and SGA by Low-Dose Aspirin in Nulliparous Women With Abnormal First-trimester Uterine Artery Dopplers. Available from: https://clinicaltrials.gov/show/nct01729468. 2012.
106. Ekeus C, Cnattingius S, Essen B, et al. Stillbirth among foreign-born women in Sweden. Eur J Public Health 2011;21:788-92.
107. Schwarze A, Nelles I, Krapp M, et al. Doppler ultrasound of the uterine artery in the prediction of severe complications during low-risk pregnancies. Archives of Gynecology and Obstetrics 2005;271:46-52.

Appendix 1 – Search strategy

Electronic databases

The search strategy included searches of the databases shown in Table 16. MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase were searched simultaneously via the Ovid SP platform. The Cochrane Library databases were searched simultaneously via the Wiley Online platform.

Table 16. Summary of electronic database searches and dates

Database	Platform	Searched on date	Date range of search
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid SP	25/06/18	1946 to Present
Embase	Ovid SP	25/06/18	1974 to 2018 June 22
The Cochrane Library, including: - Cochrane Database of Systematic Reviews (CDSR) - Cochrane Central Register of Controlled Trials (CENTRAL) - Database of Abstracts of Reviews of Effects (DARE)	Wiley Online	25/06/18	CDSR: Issue 6 of 12, June 2018 CENTRAL: Issue 5 of 12, May 2018 DARE: Issue 2 of 4, April 2015

Search Terms

Search terms included combinations of free text and subject headings (Medical Subject Headings [MeSH] for MEDLINE, and Emtree terms for Embase), grouped into the following categories:

- Disease area: Stillbirth or SGA/FGR
- Study design: Interventional or observational studies
- Interventions:
    - Screening terms (general and specific for question 1)
    - Specific monitoring terms for question 2
    - Intervention terms for questions 3 and 4

Search terms for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase are shown in Table 17, and search terms for the Cochrane Library databases are shown in Table 18.

Table 17. Search strategy for MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print and Embase

Term Group	#	Search terms	Results
Stillbirth and SGA	1	(stillbirth$ or still birth$ or stillborn or still born or ((f?etal or perinatal) adj2 (mortality or	86707

terms		death$ or loss$ or demise$ or wastage$))).tw.	
	2	stillbirth/ or perinatal death/ or fetal death/ or fetus death/ or fetal mortality/ or perinatal mortality/	81110
	3	((f?etal growth adj2 (retard$ or restrict$)) or (intrauterine growth adj2 (retard$ or restrict$)) or small for date infant or small for gestational age or FGR or IUGR or SGA).tw.	65611
	4	exp intrauterine growth retardation/ or infant, small for gestational age/ or infant, very low birth weight/ or infant, extremely low birth weight/	78758
	5	or/1-4	219739
General screening terms	6	exp mass screening/ or predictive value/ or (screen$ or predict$ or biomarker$).tw.	4981108
Screening and monitoring	7	(PIGF or placenta$ growth factor or tyrosine kinase or PAPP A or pregnancy-associated plasma protein A or (f?etal adj (cfDNA or cf DNA or cell-free DNA)) or cell-free f?etal DNA or fibronectin or PP13 or placental protein 13 or (maternal serum adj (AfP or alpha f?etoprotein or A-FP or HCG)) or free hCG or unconjugated estriol or inhibin A or activin A or estradiol or oestradiol or oestriol or human placental lactogen or hPL or schwangerschaft protein 1 or sFlt-1).tw. or exp estradiol/ or exp placental lactogen/ or exp Pregnancy-Specific beta 1-Glycoproteins/ or exp Pregnancy-Associated Plasma Protein-A/	522950
	8	(((f?etal or f?etus or maternal) adj blood flow) or ultraso$ or TAV or TVS or sonogra$ or pulsatility or (uterine artery adj2 (notching or ratio$)) or mean arterial pressure or (Doppler adj2 (velocimetry or uterine artery or ductus venosus)) or ((f?etal or f?etus) adj3 (nuchal translucency or head circumference or femur length or echogenic bowel)) or biparietal diameter or abdominal circumference or grannum grading).tw.	955096
	9	exp "Ultrasonography, Doppler, Pulsed"/	3460
	10	Pregnancy, High-Risk/ or Risk Factors/ or (((advanced maternal age or smok$ or obes$ or weight or BMI or fitness) adj2 risk) or maternal risk factor$).tw.	1301145
	11	((f?etal growth adj3 monitor$) or (f?etal movement$ adj3 (count$ or detect$ or decreas$ or reduc$)) or f?etal activity monitor$ or kick chart$ or INTERGROWTH or GROW or SCOR or AFFIRM or Estimated f?etal weight or EFW or Gestation-Related Optimal Weight).tw. or fetus movement/ or reduced fetal movement/ or growth charts/	166790
	12	or/7-11	2883634
	13	6 and 12	499588
Interventions	14	(manag$ or monitor$ or audit$ or prevent$ or avoid$ or pregnancy outcome$ or perinatal outcome$).tw.	7518277
	15	exp platelet aggregation inhibitors/ or exp anticoagulants/ or exp antihypertensive agents/ or exp induced labor/ or (anti platelet$ or antiplatelet$ or aspirin or anti coagulant$ or anticoagulant$ or heparin or anti hypertensive$ or antihypertensive$ or beta-blocker$ or beta blocker$ or nitric oxide or labetalol or hydralazine or flunarazine or dipyridamole or (calcium adj2 supplement$) or ((elect$ or plan$) adj2 (CS or c?esar?an or section or delivery or labo?r))).tw. or exp Cesarean section/	2316263
	16	14 and 15	497071
Combined	17	13 or 16	986680
	18	5 and 17	24637
Study design terms	19	exp Randomized Controlled Trials as Topic/	266311
	20	exp Randomized Controlled Trial/	969649
	21	exp Random Allocation/	173359
	22	exp Randomization/	173359
	23	exp Double Blind Method/	297200
	24	exp Single Blind Method/	56914
	25	exp Single Blind Procedure/	31634
	26	exp Double Blind Procedure/	151022
	27	exp Crossover Procedure/	55872

28	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw.	371821
29	exp Clinical Trial/	2130262
30	Clinical trial, phase i.pt.	18150
31	Clinical trial, phase ii.pt.	29277
32	Clinical trial, phase iii.pt.	13843
33	Clinical trial, phase iv.pt.	1543
34	exp Phase 1 Clinical Trial/ or exp Clinical trial, phase I/	65999
35	exp Phase 2 Clinical Trial/ or exp Clinical trial, phase II/	96476
36	exp Phase 3 Clinical Trial/ or exp Clinical trial, phase III/	48407
37	exp Phase 4 Clinical Trial/ or exp Clinical trial, phase IV/	4539
38	Controlled clinical trial.pt.	92454
39	Randomized controlled trial.pt.	462605
40	Multicenter study.pt.	234847
41	Clinical trial.pt.	510793
42	exp Clinical Trials as Topic/	585435
43	trial$.ti.	582110
44	(clinical adj trial$).tw.	745888
45	exp Placebos/	360777
46	exp Placebo/	326802
47	placebo$.tw.	471634
48	randomly allocated.tw.	54469
49	(allocated adj2 random$).tw.	61101
50	random allocation.tw.	3293
51	random assignment.tw.	4704
52	randomized.ti,ab.	1078340
53	randomised.ti,ab.	219991
54	randomisation.tw.	18142
55	randomization.tw.	60316
56	randomly.ti,ab.	675232
57	RCT.tw.	45088
58	or/19-57	4210007
59	Open-label trial$.tw.	8065
60	Open-label stud$.tw.	18829
61	Non-blinded stud$.tw.	272
62	exp Cohort Studies/	2132569
63	exp Cohort Analysis/	2132569
64	cohort analy$.tw.	15959
65	(cohort adj (study or studies)).tw.	380442
66	exp Longitudinal Studies/ or exp Longitudinal study/	229982
67	Longitudinal.tw.	476328
68	exp Follow-Up Studies/	1894661
69	exp Follow-Up/	1302243
70	(follow up adj (study or studies)).tw.	103647
71	exp Prospective Studies/ or exp Prospective study/	930782
72	(Prospective adj (study or studies)).tw.	378451
73	(evaluation adj (study or studies)).tw.	10973
74	exp Retrospective Studies/ or exp Retrospective study/	1354121
75	retrospective$.ti,ab.	1582722

	76	(chart adj3 review).tw.	98017
	77	exp Observational studies/ or exp Observational study/	194209
	78	(observational adj (study or studies)).tw.	208874
	79	((single arm or single-arm) adj3 (study or studies or trial$)).tw.	11548
	80	or/59-79	5310801
	81	58 or 80	8579586
Combined and limits	82	18 and 81	12131
	83	("Conference Abstract" or "Conference Review" or comment or letter or editorial or note or case reports).pt.	8698863
	84	(case stud$ or case report$).ti.	581179
	85	Letter/ or historical article/ or case study/	4045201
	86	exp animals/ not exp humans/	9193069
	87	or/83-86	18160271
	88	82 not 87	9753
	89	limit 88 to yr=2000-current	8034
	90	limit 89 to yr=2000-2013	4855
	91	89 not 90	3179
	92	remove duplicates from 90	3545
	93	remove duplicates from 91	2286
	94	92 or 93	5831

Table 18. Search strategy for the Cochrane Library Databases (Searched via the Wiley Online platform)

Term Group	#	Search terms	Results
Stillbirth and SGA terms	1	(stillbirth* or "still birth*" or stillborn or "still born" or ((f?etal or perinatal) near/2 (mortality or death* or loss* or demise* or wastage*))):ti,ab,kw	1418
	2	[mh stillbirth] or [mh ^"perinatal death"] or [mh ^"fetal death"]	356
	3	(("f?etal growth" near/2 (retard* or restrict*)) or ("intrauterine growth" near/2 (retard* or restrict*)) or "small for date infant" or "small for gestational age" or FGR or IUGR or SGA):ti,ab,kw	1697
	4	[mh "fetal growth retardation"] or [mh "intrauterine growth retardation"] or [mh "infant, small for gestational age"] or [mh "infant, very low birth weight"] or [mh "infant, extremely low birth weight"]	1506
	5	{or #1-#4}	4047
General screening terms	6	[mh "mass screening"] or [mh ^"predictive value"] or (screen* or predict* or biomarker*):ti,ab,kw	136003
Screening and monitoring terms	7	(PIGF or "placenta* growth factor" or "tyrosine kinase" or "PAPP A" or "pregnancy-associated plasma protein A" or (f?etal next (cfDNA or "cf DNA" or "cell-free DNA")) or "cell-free f?etal DNA" or fibronectin or PP13 or "placental protein 13" or ("maternal serum" next (AfP or "alpha f?etoprotein" or A-FP or HCG)) or "free hCG" or "unconjugated estriol" or "inhibin A" or "activin A" or estradiol or oestradiol or oestriol or "human placental lactogen" or hPL or "schwangerschaft protein 1" or sFlt-1):ti,ab,kw or [mh estradiol] or [mh "placental lactogen"] or [mh "Pregnancy-Specific beta 1-Glycoproteins"] or [mh "Pregnancy-Associated Plasma Protein-A"]	12277
	8	(((f?etal or f?etus or maternal) next "blood flow") or ultraso* or TAV or TVS or sonogra* or pulsatility or ("uterine artery" near/2 (notching or ratio*)) or "mean arterial pressure" or (Doppler near/2 (velocimetry or "uterine artery" or "ductus venosus")) or ((f?etal or f?etus) near/3 ("nuchal translucency" or "head circumference" or "femur length" or "echogenic bowel")) or "biparietal diameter" or "abdominal circumference" or "grannum grading"):ti,ab,kw	36529
	9	[mh "Ultrasonography, Doppler, Pulsed"]	157

	10	[mh "Pregnancy, High-Risk"] or [mh "Risk Factors"] or (((advanced maternal age or smok* or obes* or weight or BMI or fitness) near/2 risk) or maternal risk factor*):ti,ab,kw	30566
	11	(("f?etal growth" near/3 monitor*) or ("f?etal movement*" near/3 (count* or detect* or decreas* or reduc*)) or "f?etal activity monitor*" or "kick chart*" or INTERGROWTH or GROW or SCOR or AFFIRM or "Estimated f?etal weight" or EFW or "Gestation-Related Optimal Weight"):ti,ab,kw or [mh "fetus movement"] or [mh "reduced fetal movement"] or [mh ^"growth charts"]	1172
	12	{or #7-#11}	78213
	13	#6 and #12	14740
Interventions	14	(manag* or monitor* or audit* or prevent* or avoid* or pregnancy outcome* or perinatal outcome*):ti,ab,kw	289019
	15	[mh "platelet aggregation inhibitors"] or [mh anticoagulants] or [mh "antihypertensive agents"] or [mh "induced labor"] or ("anti platelet*" or antiplatelet* or aspirin or "anti coagulant*" or anticoagulant* or heparin or "anti hypertensive*" or antihypertensive* or "beta-blocker*" or "nitric oxide" or labetalol or hydralazine or flunarazine or dipyridamole or (calcium near/2 supplement*) or ((elect* or plan*) near/2 (CS or c?esar?an or section or delivery or labo?r))):ti,ab,kw or [mh "Cesarean section"]	63189
	16	#14 and #15	21942
Combined with limits	17	#13 or #16	36097
	18	#17 and #5 Publication Year from 2000 to 2018	554

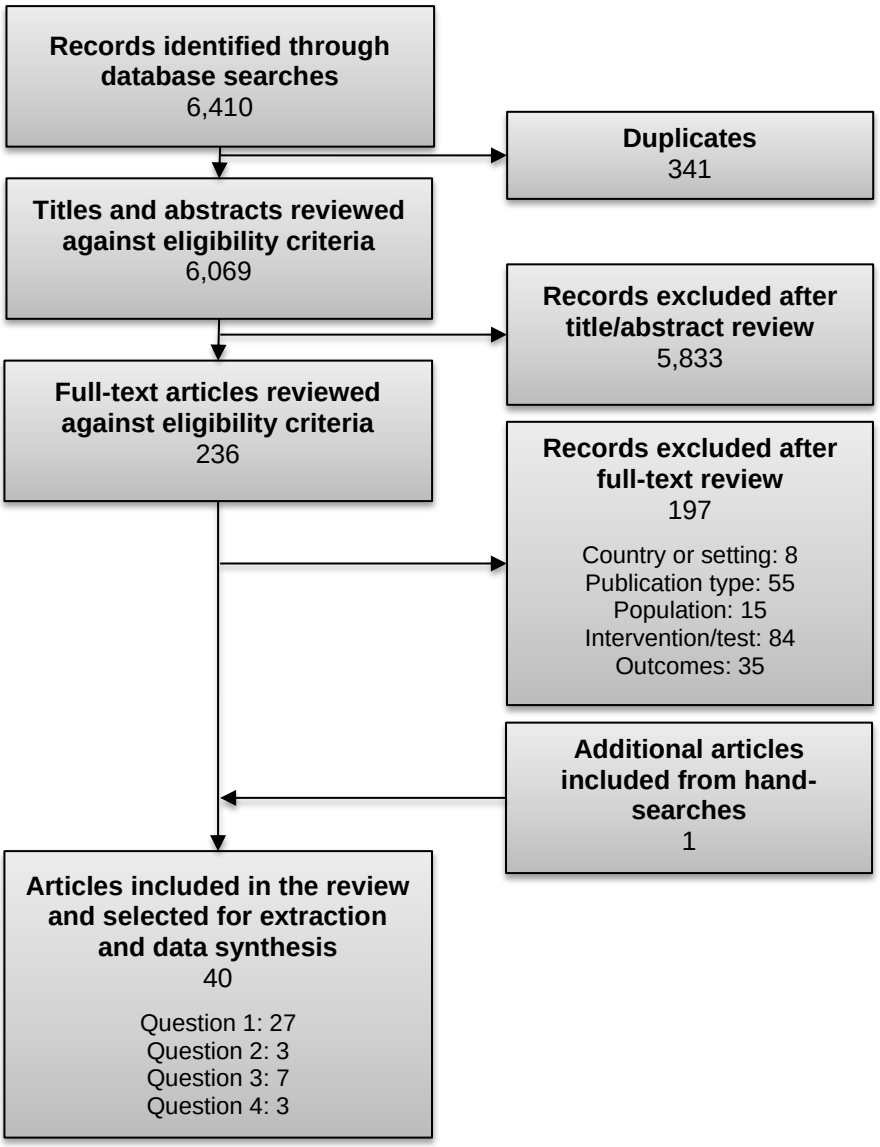
Results were imported into EndNote and de-duplicated.

Appendix 2 – Included and excluded studies

PRISMA flowchart

Figure 5 summarises the volume of publications included and excluded at each stage of the review. 40 publications were ultimately judged to be relevant to one or more review questions and were considered for extraction. Publications that were included or excluded after the review of full-text articles are detailed below.

Figure 5. Summary of publications included and excluded at each stage of the review



Publications included after review of full-text articles

The 40 publications included after review of full-texts are summarised in Table 19 below.

Table 19. Summary of publications included after review of full-text articles, and the question(s) each publication was identified as being relevant to

Study	Question	The test (Q1)/monitoring strategy (Q2)/intervention (Q3, Q4)	Comments
Akolekar 2016a[66]	Q1	Maternal risk factors, PAPP-A, PLGF, DV-PIV and UT-PI	
Akolekar 2016b[65]	Q1	Maternal risk factors, fetal biometry and UtA-PI	
Aupont 2016[67]	Q1	PlGF	Same cohort as Akolekar 2016b
Bakalis 2015[68]	Q1	Screening for low cerebroplacental ratio	Cohort may be included in Akolekar 2016a, Akolekar 2016b, Aupont 2016, Valino 2016a and b, Yerlikaya 2016
Boers 2010[15]	Q4	Induction of labour, expectant monitoring	DIGITAT trial
Chaiworapongsa 2013[69]	Q1	PlGF/sVEGFR-1 and PlGF/sEng	Study extracted consisted of a cohort and case-control parts extracted as 2013a and 2013b respectively
Chaiworapongsa 2017[88]	Q1	sVEGFR-1, sEng, PlGF, PlGF/sVEGFR-1 and PlGF/sEng	
Ayala, 2013[8, 105]	Q3	Aspirin, placebo	Study identified from a list of associated publications to the ClinicalTrials.gov reference
Conde-Agudelo 2015[44]	Q1	Fetoplacental proteins/hormone-related tests, ultrasound imaging-related tests, combinations of tests, and others	SLR
Dugoff 2004[70]	Q1	PAPP-A	
Dugoff 2005[72]	Q1	Maternal serum AFP, hCG, uE_3, and inhibin A	Study reports on the same population as Dugoff 2004
Dugoff 2008[71]	Q1	Measurement of quad screen markers (AFP, total hCG, uE_3, and inhibin A)	
Familiari 2016[86]	Q1	Maternal characteristics, femur length and UtA-PI	
Haddad 2016[9]	Q3	Enoxaparin and aspirin, aspirin	
Hemming 2011[40]	Q1	SGA by fetal or population-based growth standard	
Kingdom 2011[10]	Q3	UFH, placebo	
Lees 2015[94]	Q2	DV-PI, DV A wave with no or reversed flow, CTG-STV	
Marttala 2010[78]	Q1	PAPP-A	
Mastrodima 2016[73]	Q1	PAPP-A, UT-PI and DV-PIV	Same cohort as Akolekar 2016, Akolekar 2016b, Aupont 2016 with crossover with Bakalis 2011
Odibo 2012[79]	Q1	SGA by customised or population-based growth standard	
Poon 2013[80]	Q1	UtA-PI	
Rabinovich 2018[16]	Q4	Induction of labour, expectant monitoringmonitoring	

Singh 2012[81]	Q1	UtA-RI	
Smith 2007[82]	Q1	Maternal characteristics, AFP, hCGhCG	
Smith 2007[74]	Q1	Doppler pulsatility index, unilateral notch and bilateral notch, maternal characteristics	
Smith 2014[87]	Q1	Fetal growth curves	
Spaggiari 2013[97]	Q2	Repeat AFP scan in women with abnormal results vs not having had a second scanscan	
Sharp 2018[14]	Q3	Sildenafil citrate, placebo	STRIDER trial
Subtil 2003a[11]	Q3	UtA Doppler	
Subtil 2003b[12]	Q3	UtA Doppler	This was a supporting reference for the Subtil 2003a study
Sutan 2010[83]	Q1	Maternal characteristics	
Tancrede 2015[75]	Q1	AFP, hCG	
Trudell 2015[84]	Q1	Sex-specific and non-sex-specific growth standards	
Trudell 2017[85]	Q1	Maternal characteristics	
Tveit 2009[96]	Q2	Information and guidelines on decreased fetal movements vs no interventions or guidelines	
Valino 2016a[77]	Q1	UtA-PI	Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016 b, Yerlikaya 2016
Valino 2016b[76]	Q1	UtA-PI, UA-PI, MCA-PI, MAP, PlGF and sFlt-1	Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016 b, Yerlikaya 2016
Von Dadelszen 2011[13]	Q3	Sildenafil citrate, no sildenafil citrate	
Walker 2016[17]	Q4	Induction of labour, expectant monitoring	
Yerlikaya 2016[50]	Q1	Maternal characteristics	Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011 and Mastrodima 2016

Abbreviations: AFP, alpha-fetoprotein; CTG-STV, Cardiotocograph short-term variation; DV, ductus venosus; DV-PIV, ductus venosus pulsatility index for veins; hCG, human chorionic gonadotropin; MAP, mean arterial pressure; MCA-PI, middle cerebral artery pulsatility index; PAPP-A, pregnancy-associated plasma protein A; PI, Pulsatility index; PLGF/PlGF, placenta growth factor; RI, resistance index; sEng, soluble endoglin; sFlt-1, soluble fms-like tyrosine kinase 1; SGA, small for gestational age; sVEGFR-1, soluble vascular endothelial growth factor receptor-1; UA-PI, umbilical artery pulsatility index; uE_3, unconjugated estriol ; UFH, unfractionated heparin; UtA, uterine artery; UT-PI, uterine artery pulsatility index.

Publications excluded after review of full-text articles

Of the 236 publications included after the review of titles and abstracts, 197 were ultimately judged not to be relevant to this review. These publications, along with reasons for exclusion, are listed in Table 20.

Table 20. Publications excluded after review of full-text articles

Reference	Reason for exclusion
Abalos E, Duley L, Steyn DW. Antihypertensive drug therapy for mild to moderate hypertension during pregnancy. The Cochrane database of systematic reviews. 2014;2:CD002252.	Not a relevant pregnant population
Akolekar R, Sarno L, Wright A, Wright D, Nicolaides KH. Fetal middle cerebral artery and umbilical artery pulsatility index: effects of maternal characteristics and medical history. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;45(4):402-8.	Diagnostic test/monitoring strategy/intervention not relevant
Alfirevic Z, Devane D, Gyte GML, Cuthbert A. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. Cochrane Database of Systematic Reviews. 2017;2017 (2) (no pagination)(CD006066).	Study type not eligible
Alfirevic Z, Stampalija T, Medley N. Fetal and umbilical Doppler ultrasound in normal pregnancy. The Cochrane database of systematic reviews. 2015;4:CD001450.	Study type not eligible
Alkazaleh F, Chaddha V, Viero S, Malik A, Anastasiades C, Sroka H, et al. Second-trimester prediction of severe placental complications in women with combined elevations in alpha-fetoprotein and human chorionic gonadotrophin. American Journal of Obstetrics and Gynecology. 2006;194(3):821-7.	Diagnostic test/monitoring strategy/intervention not relevant
Alkazaleh F, Viero S, Simchen M, Walker M, Smith G, Laskin C, et al. Ultrasound diagnosis of severe thrombotic placental damage in the second trimester: An observational study. Ultrasound in Obstetrics and Gynecology. 2004;23(5):472-6.	Setting not relevant/study not in pregnant women
Allen R, Aquilina J. Prospective observational study to determine the accuracy of first-trimester serum biomarkers and uterine artery Dopplers in combination with maternal characteristics and arteriography for the prediction of women at risk of preeclampsia and other adverse pregnancy outcomes. Journal of Maternal-Fetal and Neonatal Medicine. 2017:1-18.	No relevant outcomes reported
Anderson NH, Sadler LC, McKinlay CJD, McCowan LME. INTERGROWTH-21st vs customized birthweight standards for identification of perinatal mortality and morbidity. American Journal of Obstetrics and Gynecology. 2016;214(4):509.e1-509.e7.	No relevant outcomes reported
Ayres-de-Campos D, Ugwumadu A, Banfield P, Lynch P, Amin P, Horwell D, et al. A randomised clinical trial of intrapartum fetal monitoring with computer analysis and alerts versus previously available monitoring. BMC pregnancy and childbirth [Internet]. 2010; 10:[71 p.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/364/CN-00772364/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Bais JMJ, Eskes M, Pel M, Bonsel GJ, Bleker OP. Effectiveness of detection of intrauterine growth retardation by abdominal palpation as screening test in a low risk population: An observational study. European Journal of Obstetrics Gynecology and Reproductive Biology. 2004;116(2):164-9.	Diagnostic test/monitoring strategy/intervention not relevant
Balci S. Predictive values of maternal serum PAPP-A level, uterine artery doppler velocimetry, and fetal biometric measurements for poor pregnancy and poor neonatal outcomes in pregnant women. Journal of the Turkish German Gynecology Association. 2016;17(3):143-9.	Diagnostic test/monitoring strategy/intervention not relevant
Baltajian K, Bajracharya S, Salahuddin S, Berg AH, Geahchan C, Wenger JB, et al. Sequential plasma angiogenic factors levels in women with suspected preeclampsia. American Journal of Obstetrics and Gynecology. 2016;215(1):89.e1-.e10.	Diagnostic test/monitoring strategy/intervention not relevant

Reference	Reason for exclusion
Barati M, Shahbazian N, Ahmadi L, Masihi S. Diagnostic evaluation of uterine artery Doppler sonography for the prediction of adverse pregnancy outcomes. Journal of Research in Medical Sciences. 2014;19(6):515-9.	Setting not relevant/study not in pregnant women
Barros FC, Bhutta ZA, Batra M, Hansen TN, Victora CG, Rubens CE. Global report on preterm birth and stillbirth (3 of 7): Evidence for effectiveness of interventions. BMC Pregnancy and Childbirth. 2010;10 (SUPPL. 1) (no pagination)(S3).	Study type not eligible
Bartkute K, Balsyte D, Wisser J, Kurmanavicius J. Pregnancy outcomes regarding maternal serum AFP value in second trimester screening. Journal of Perinatal Medicine. 2017;45(7):817-20.	No relevant outcomes reported
Becker R, Keller T, Kiesewetter H, Fangerau H, Bittner U. Individual risk assessment of adverse pregnancy outcome by multivariate regression analysis may serve as basis for drug intervention studies: Retrospective analysis of 426 high-risk patients including ethical aspects. Archives of Gynecology and Obstetrics. 2013;288(1):41-8.	Diagnostic test/monitoring strategy/intervention not relevant
Berkley E, Chauhan SP, Abuhamad A. Doppler assessment of the fetus with intrauterine growth restriction. American Journal of Obstetrics and Gynecology. 2012;206(4):300-8.	Study type not eligible
Bligh LN, Al Solai A, Greer RM, Kumar S. Diagnostic Performance of Cerebroplacental Ratio Thresholds at Term for Prediction of Low Birthweight and Adverse Intrapartum and Neonatal Outcomes in a Term, Low-Risk Population. Fetal Diagnosis and Therapy. 2017;27.	Diagnostic test/monitoring strategy/intervention not relevant
Bligh LN, Alsolai A, Greer RM, Kumar S. Screening for adverse perinatal outcomes: uterine artery Doppler, cerebroplacental ratio and estimated fetal weight in low-risk women at term. Journal of Maternal-Fetal and Neonatal Medicine. 2017:1-7.	Diagnostic test/monitoring strategy/intervention not relevant
Bligh LN, Greer RM, Kumar S. Screening Performance of Placental Growth Factor for the Prediction of Low Birth Weight and Adverse Intrapartum and Neonatal Outcomes in a Term Low-Risk Population. Fetal Diagnosis and Therapy. 2017;11.	Diagnostic test/monitoring strategy/intervention not relevant
Blix E, Reinar LM, Klovning A, Oian P. Prognostic value of the labour admission test and its effectiveness compared with auscultation only: A systematic review. BJOG: An International Journal of Obstetrics and Gynaecology. 2005;112(12):1595-604.	Study type not eligible
Bond DM, Gordon A, Hyett J, de Vries B, Carberry AE, Morris J. Planned early delivery versus expectant management of the term suspected compromised baby for improving outcomes. The Cochrane database of systematic reviews. 2015;11:CD009433.	Study type not eligible
Brajenovic-Milic B, Tislaric D, Zuvic-Butorac M, Bacic J, Petrovic O, Ristic S, et al. Elevated second-trimester free beta-hCG as an isolated finding and pregnancy outcomes. Fetal Diagnosis and Therapy. 2004;19(6):483-7.	No relevant outcomes reported
Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, et al. Ultrasound screening in pregnancy: A systematic review of the clinical effectiveness, cost-effectiveness and women's views. Health Technology Assessment. 2000;4(16):i-vi+1-183.	Study type not eligible
Bricker L, Medley N, Pratt JJ. Routine ultrasound in late pregnancy (after 24 weeks' gestation). The Cochrane database of systematic reviews. 2015;6:CD001451.	Study type not eligible
Cetin I, Mazzocco MI, Giardini V, Cardellicchio M, Calabrese S, Algeri P, et al. PlGF in a clinical setting of pregnancies at risk of preeclampsia and/or intrauterine growth restriction. J Matern Fetal Neonatal Med. 2017;30(2):144-9.	Diagnostic test/monitoring strategy/intervention not relevant
Chambers AE, Griffin C, Naif SA, Mills I, Mills WE, Syngelaki A, et al. Quantitative ELISAs for serum soluble LHCGR and hCG-LHCGR complex: potential diagnostics in first trimester pregnancy screening for stillbirth, Down's syndrome, preterm delivery and preeclampsia. Reprod Biol Endocrinol. 2012;10:113.	No relevant outcomes reported
Chaveeva P, Carbone IF, Syngelaki A, Akolekar R, Nicolaides KH. Contribution of method of conception on pregnancy outcome after the 11-13 weeks scan. Fetal Diagn Ther. 2011;30(1):9-22.	Diagnostic test/monitoring strategy/intervention not relevant
Chiossi G, Pedroza C, Costantine MM, Truong VTT, Gargano G, Saade GR. Customized vs population-based growth charts to identify neonates at risk of adverse outcome: systematic review and Bayesian meta-analysis of observational studies. Ultrasound in Obstetrics & Gynecology. 2017;50:156-166	Diagnostic test/monitoring strategy/intervention not relevant

Reference	Reason for exclusion
Clinicaltrials.gov. NCT01355822 Impact of the NO-donor Pentaerythrithyltetrantrate on Perinatal Outcome in High-risk Pregnancies. 2011. Available from: https://clinicaltrials.gov/show/nct01355822	Diagnostic test/monitoring strategy/intervention not relevant
Cnossen J, Morris R, ter RG, Mol B, Post J, Coomarasamy A, et al. Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis (Structured abstract). CMAJ: Canadian Medical Association Journal [Internet]. 2008; 178(6):[701-11 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/cldare/articles/DARE-12008008128/frame.html.	Study type not eligible
Coleman MA, McCowan LM, North RA. Mid-trimester uterine artery Doppler screening as a predictor of adverse pregnancy outcome in high-risk women. Ultrasound Obstet Gynecol. 2000;15(1):7-12.	Not a relevant pregnant population
Conde-Agudelo A, Villar J, Kennedy SH, Papageorghiou AT. Predictive accuracy of cerebroplacental ratio for adverse perinatal and neurodevelopmental outcomes in suspected fetal growth restriction: systematic review and meta-analysis. Ultrasound in Obstetrics & Gynecology. 2018;19.	Study type not eligible
Coomarasamy A, Honest H, Papaioannou S, Gee H, Khan KS. Aspirin for prevention of preeclampsia in women with historical risk factors: A systematic review. Obstetrics and Gynecology. 2003;101(6):1319-32.	Study type not eligible
Costa SL, Proctor L, Dodd JM, Toal M, Okun N, Johnson JA, et al. Screening for Placental Insufficiency in High-risk Pregnancies: Is Earlier Better? Placenta. 2008;29(12):1034-40.	Not a relevant pregnant population
Cui Y, Zhu B, Zheng F. Low-dose aspirin at <=16 weeks of gestation for preventing preeclampsia and its maternal and neonatal adverse outcomes: A systematic review and meta-analysis. Experimental and Therapeutic Medicine. 2018;15(5):4361-9.	Study type not eligible
Dane B, Dane C, Cetin A, Kiray M, Sivri D, Yayla M. Pregnancy outcome in fetuses with increased nuchal translucency. Journal of Perinatology. 2008;28(6):400-4.	No relevant outcomes reported
Darmstadt GL, Yakoob M, Haws RA, Menezes EV, Soomro T, Bhutta ZA. Reducing stillbirths: Interventions during labour. BMC Pregnancy and Childbirth. 2009;9 (SUPPL. 1) (no pagination)(S6).	Study type not eligible
Devane D, Lalor JG, Daly S, McGuire W, Cuthbert A, Smith V. Cardiotocography versus intermittent auscultation of fetal heart on admission to labour ward for assessment of fetal wellbeing. Cochrane Database of Systematic Reviews. 2017;2017 (1) (no pagination)(CD005122).	Study type not eligible
Di Bartolomeo A, Chauleur C, Gris JC, Chapelle C, Noblot E, Laporte S, et al. Tissue factor pathway inhibitor for prediction of placenta-mediated adverse pregnancy outcomes in high-risk women: AngioPred study.[Erratum appears in PLoS One. 2017 Jul 12;12 (7):e0181474; PMID: 28704568]. PLoS ONE. 2017;12(3):e0173596.	Diagnostic test/monitoring strategy/intervention not relevant
Dodd JM, McLeod A, Windrim RC, Kingdom J. Antithrombotic therapy for improving maternal or infant health outcomes in women considered at risk of placental dysfunction. Cochrane Database of Systematic Reviews [Internet]. 2013; (7). Available from: http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD006780.pub3/abstract.	Study type not eligible
Donnelly J, Byrne J, Murphy K, McAuliffe F. Obstetric outcome with low molecular weight heparin therapy during pregnancy. Irish Medical Journal. 2012;105(1):27-9.	Not a relevant pregnant population
Duley L, Henderson-Smart D, Knight M, King J. Antiplatelet drugs for prevention of pre-eclampsia and its consequences: Systematic review. British Medical Journal. 2001;322(7282):329-33.	Diagnostic test/monitoring strategy/intervention not relevant
Dunn L, Greer R, Flenady V, Kumar S. Sildenafil in Pregnancy: A Systematic Review of Maternal Tolerance and Obstetric and Perinatal Outcomes. Fetal Diagnosis and Therapy. 2017;41(2):81-8.	Study type not eligible
Endres LK, Krotz S, Grobman WA. Isolated low second-trimester maternal serum beta-human chorionic gonadotropin is not associated with adverse pregnancy outcome. American Journal of Obstetrics and Gynecology. 2003;189(3):755-7.	No relevant outcomes reported

Reference	Reason for exclusion
Frias AE, Jr., Luikenaar RA, Sullivan AE, Lee RM, Porter TF, Branch DW, et al. Poor obstetric outcome in subsequent pregnancies in women with prior fetal death. Obstet Gynecol. 2004;104(3):521-6.	Not a relevant pregnant population
Fujisaki M, Furuta K, Ohhashi M, Furukawa S, Kodama Y, Kawagoe Y, et al. Antithrombin improves the maternal and neonatal outcomes but not the angiogenic factors in extremely growth-restricted fetuses at <28 weeks of gestation. Journal of Perinatal Medicine. 2017;45(7):837-42.	Diagnostic test/monitoring strategy/intervention not relevant
Gaccioli F, Sovio U, Cook E, Hund M, Charnock-Jones DS, Smith GCS. Screening for fetal growth restriction using ultrasound and the sFLT1/PlGF ratio in nulliparous women: a prospective cohort study. The Lancet Child and Adolescent Health. 2018.	Diagnostic test/monitoring strategy/intervention not relevant
Garcia B, Llurba E, Valle L, Gomez-Roig MD, Juan M, Perez-Matos C, et al. Do knowledge of uterine artery resistance in the second trimester and targeted surveillance improve maternal and perinatal outcome? UTOPIA study: a randomized controlled trial. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2016;47(6):680-9.	Not a relevant pregnant population
Garcia-Tizon Larroca S, Arevalo-Serrano J, Duran Vila A, Pintado Recarte MP, Cueto Hernandez I, Solis Pierna A, et al. Human Development Index (HDI) of the maternal country of origin as a predictor of perinatal outcomes - a longitudinal study conducted in Spain. BMC Pregnancy and Childbirth. 2017;17 (1) (no pagination)(314).	Diagnostic test/monitoring strategy/intervention not relevant
Gkogkos P, Androutsopoulos G, Vassilakos P, Panayiotakis G, Kourounis G, Decavalas G. Mid-trimester maternal serum AFP levels in predicting adverse pregnancy outcome. Clinical and Experimental Obstetrics and Gynecology. 2008;35(3):208-10.	Diagnostic test/monitoring strategy/intervention not relevant
Goffinet F, Aboulker D, Paris-Llado J, Bucourt M, Uzan M, Papiernik E, et al. Screening with a uterine Doppler in low risk pregnant women followed by low dose aspirin in women with abnormal results: A multicenter randomised controlled trial. British Journal of Obstetrics and Gynaecology. 2001;108(5):510-8.	Diagnostic test/monitoring strategy/intervention not relevant
Gomes MS, Carlos-Alves M, Trocado V, Arteiro D, Pinheiro P. Prediction of adverse pregnancy outcomes by extreme values of first trimester screening markers. Obstetric Medicine. 2017;10(3):132-7.	Diagnostic test/monitoring strategy/intervention not relevant
Gomez O, Martinez JM, Figueras F, Del Rio M, Borobio V, Puerto B, et al. Uterine artery Doppler at 11-14 weeks of gestation to screen for hypertensive disorders and associated complications in an unselected population. Ultrasound in Obstetrics and Gynecology. 2005;26(5):490-4.	No relevant outcomes reported
Goyal NK, Hall ES, Greenberg JM, Kelly EA. Risk prediction for adverse pregnancy outcomes in a medicaid population. Journal of Women's Health. 2015;24(8):681-8.	Diagnostic test/monitoring strategy/intervention not relevant
Graham EM, Petersen SM, Christo DK, Fox HE. Intrapartum electronic fetal heart rate monitoring and the prevention of perinatal brain injury. Obstetrics and Gynecology. 2006;108(3 I):656-66.	Diagnostic test/monitoring strategy/intervention not relevant
Gramellini D, Piantelli G, Verrotti C, Fieni S, Delle Chiaie L, Kaihura C. Doppler velocimetry and non stress test in severe fetal growth restriction. Clinical and Experimental Obstetrics and Gynecology. 2001;28(1):33-9.	Diagnostic test/monitoring strategy/intervention not relevant
Grant A, Glazener CM. Elective caesarean section versus expectant management for delivery of the small baby. Cochrane database of systematic reviews (Online). 2001(2):CD000078.	Study type not eligible
Gris JC, Chauleur C, Faillie JL, Baer G, Mares P, Fabbro-Peray P, et al. Enoxaparin for the secondary prevention of placental vascular complications in women with abruptio placentae: The pilot randomised controlled NOH-AP trial. Thrombosis and Haemostasis. 2010;104(4):771-9.	Diagnostic test/monitoring strategy/intervention not relevant
Grivell R, Alfirevic Z, Gyte G, Devane D. Antenatal cardiotocography for fetal assessment. Cochrane database of systematic reviews (online) [Internet]. 2012; 2012(12) (no pagination). Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/776/CN-01298776/frame.html.	Study type not eligible

Reference	Reason for exclusion
Grivell RM, Alfirevic Z, Gyte GML, Devane D. Antenatal cardiotocography for fetal assessment. Cochrane Database of Systematic Reviews. 2015;2015(9):1-39.	Study type not eligible
Groom K, McCowan L, Mackay L, Lee A, Said J, Kane S, et al. Enoxaparin for the prevention of preeclampsia and intrauterine growth restriction in women with a history: a randomized trial. American journal of obstetrics and gynecology [Internet]. 2017; 216(3):[296.e1-.e14 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/017/CN-01379017/frame.html.	Study type not eligible
Groom K, McCowan L, Mackay L, Lee A, Said J, Kane S, et al. Enoxaparin for the prevention of preeclampsia and intrauterine growth restriction in women with a history: a randomized trial. American journal of obstetrics and gynecology [Internet]. 2017; 216(3):[296.e1-.e14 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/017/CN-01379017/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Groom KM, McCowan LM, Stone PR, Chamley LC, McLintock C. Enoxaparin for the prevention of preeclampsia and intrauterine growth restriction in women with a prior history - an open-label randomised trial (the EPPI trial): Study protocol. BMC Pregnancy and Childbirth. 2016;16 (1) (no pagination)(367).	Diagnostic test/monitoring strategy/intervention not relevant
Gulmezoglu AM, Crowther CA, Middleton P, Heatley E. Induction of labour for improving birth outcomes for women at or beyond term. Cochrane database of systematic reviews (Online). 2012;6:CD004945.	Study type not eligible
Harrington K, Fayyad A, Thakur V, Aquilina J. The value of uterine artery Doppler in the prediction of uteroplacental complications in multiparous women. Ultrasound Obstet Gynecol. 2004;23(1):50-5.	Diagnostic test/monitoring strategy/intervention not relevant
Harrington K, Kurdi W, Aquilina J, England P, Campbell S. A prospective management study of slow-release aspirin in the palliation of uteroplacental insufficiency predicted by uterine artery Doppler at 20 weeks. Ultrasound in Obstetrics and Gynecology. 2000;15(1):13-8.	Diagnostic test/monitoring strategy/intervention not relevant
Haws RA, Yakoob M, Soomro T, Menezes EV, Darmstadt GL, Bhutta ZA. Reducing stillbirths: Screening and monitoring during pregnancy and labour. BMC Pregnancy and Childbirth. 2009;9 (SUPPL. 1) (no pagination)(S5).	Study type not eligible
Heazell A, Bernatavicius G, Roberts S, Garrod A, Whitworth M, Johnstone E, et al. A randomised controlled trial comparing standard or intensive management of reduced fetal movements after 36 weeks gestation--a feasibility study. BMC pregnancy and childbirth [Internet]. 2013; 13:[95 p.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/300/CN-00964300/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Heazell AE, Whitworth M, Duley L, Thornton JG. Use of biochemical tests of placental function for improving pregnancy outcome. The Cochrane database of systematic reviews. 2015;11:CD011202.	No relevant outcomes reported
Hui D, Okun N, Murphy K, Kingdom J, Uleryk E, Shah PS. Combinations of maternal serum markers to predict preeclampsia, small for gestational age, and stillbirth: a systematic review. J Obstet Gynaecol Can. 2012;34(2):142-53.	Study type not eligible
Huras H, Kalinka J, Radon-Pokracka M, Kusmierska-Urban K, Kufelnicka-Babout M, Nowak M, et al. Effects of pentoxifylline and docosahexaenoic acid supplemental treatment in intrauterine growth restriction. J Matern Fetal Neonatal Med [Internet]. 2014; 27:[131 p.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/231/CN-01042231/frame.html.	Study type not eligible
Ivanovski MJ, Lazarevski S, Popovic M. Middle cerebral artery flow velocity waveforms in prediction of adverse outcome in intrauterine growth retarded fetuses. Gynaecologia et Perinatologia. 2005;14(3):133-9.	Diagnostic test/monitoring strategy/intervention not relevant
Jacquemyn Y, Martens E, Martens G. Foetal monitoring during labour: Practice versus theory in a region-wide analysis. Clinical and Experimental Obstetrics and Gynecology. 2012;39(3):307-9.	No relevant outcomes reported
Jayaballa M, Sood S, Alahakoon I, Padmanabhan S, Cheung NW, Lee V. Microalbuminuria is a predictor of adverse pregnancy outcomes including preeclampsia. Pregnancy Hypertension. 2015;5(4):303-7.	Diagnostic test/monitoring strategy/intervention not relevant
Kaijomaa M, Ulander VM, Hamalainen E, Alfthan H, Markkanen H, Heinonen S, et al. The risk of adverse pregnancy outcome among pregnancies with extremely low maternal PAPP-A. Prenatal Diagnosis. 2016;36(12):1115-20.	No relevant outcomes reported
Karim JN, Sau A. Low pregnancy associated plasma protein-A in the 1st trimester: Is it a predictor of poor perinatal outcome? Journal of Obstetrics and Gynaecology. 2013;33(4):351-4.	No relevant outcomes reported

Reference	Reason for exclusion
Khalil A, Morales-Rosello J, Townsend R, Morlando M, Papageorghiou A, Bhide A, et al. Value of third-trimester cerebroplacental ratio and uterine artery Doppler indices as predictors of stillbirth and perinatal loss. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2016;47(1):74-80.	Not a relevant pregnant population
Kozer E, Costei AM, Boskovic R, Nulman I, Nikfar S, Koren G. Effects of aspirin consumption during pregnancy on pregnancy outcomes: Meta-analysis. Birth Defects Research Part B - Developmental and Reproductive Toxicology. 2003;68(1):70-84.	Not a relevant pregnant population
Kwik M, Morris J. Association between first trimester maternal serum pregnancy associated plasma protein-A and adverse pregnancy outcome. Australian & New Zealand Journal of Obstetrics & Gynaecology. 2003;43(6):438-42	Diagnostic test/monitoring strategy/intervention not relevant
Lain SJ, Algert CS, Tasevski V, Morris JM, Roberts CL. Record linkage to obtain birth outcomes for the evaluation of screening biomarkers in pregnancy: A feasibility study. BMC Medical Research Methodology. 2009;9 (1) (no pagination)(48).	No relevant outcomes reported
Lalor JG, Fawole B, Alfirevic Z, Devane D. Biophysical profile for fetal assessment in high risk pregnancies. Cochrane Database of Systematic Reviews. 2008;(2) (no pagination)(CD000038).	Study type not eligible
Liston R, Crane J, Hamilton E, Hughes O, Kuling S, MacKinnon C, et al. Fetal health surveillance in labour. Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC. 2002;24(3):250-76; quiz 77-80.	Study type not eligible
Liston R, Sawchuck D, Young D. No. 197a-Fetal Health Surveillance: Antepartum Consensus Guideline. Journal of Obstetrics and Gynaecology Canada. 2018;40(4):e251-271.	Diagnostic test/monitoring strategy/intervention not relevant
Lithner CU, Kublickas M, Ek S. Pregnancy outcome for fetuses with increased nuchal translucency but normal karyotype. Journal of medical screening. 2016;23(1):1-6.	No relevant outcomes reported
Lobmaier SM, Huhn EA, Pildner Von Steinburg S, Muller A, Schuster T, Ortiz JU, et al. Phase-rectified signal averaging as a new method for surveillance of growth restricted fetuses. Journal of Maternal-Fetal and Neonatal Medicine. 2012;25(12):2523-8.	Diagnostic test/monitoring strategy/intervention not relevant
Lopez-Quesada E, Vilaseca MA, Vela A, Lailla JM. Perinatal outcome prediction by maternal homocysteine and uterine artery Doppler velocimetry. Eur J Obstet Gynecol Reprod Biol. 2004;113(1):61-6.	Diagnostic test/monitoring strategy/intervention not relevant
Lutomski JE, Meaney S, Greene RA, Ryan AC, Devane D. Expert systems for fetal assessment in labour. Cochrane Database of Systematic Reviews. 2015;2015 (4) (no pagination)(CD010708).	Study type not eligible
Maged AM, Saad H, Meshaal H, Salah E, Abdelaziz S, Omran E, et al. Maternal serum homocysteine and uterine artery Doppler as predictors of preeclampsia and poor placentation. Archives of Gynecology and Obstetrics. 2017;296(3):475-82.	Diagnostic test/monitoring strategy/intervention not relevant
Maiz N, Valencia C, Emmanuel EE, Staboulidou I, Nicolaides KH. Screening for adverse pregnancy outcome by Ductus venosus Doppler at 11-13+6 weeks of gestation. Obstetrics and Gynecology. 2008;112(3):598-605.	Not a relevant pregnant population
Makrydimas G, Sebire NJ, Lolis D, Vlassis N, Nicolaides KH. Fetal loss following ultrasound diagnosis of a live fetus at 6-10 weeks of gestation. Ultrasound in Obstetrics and Gynecology. 2003;22(4):368-72.	Diagnostic test/monitoring strategy/intervention not relevant
Mangesi L, Hofmeyr GJ, Smith V, Smyth RMD. Fetal movement counting for assessment of fetal wellbeing. Cochrane Database of Systematic Reviews. 2015;2015 (10) (no pagination)(CD004909).	Study type not eligible
Marchesoni D, Pezzani I, Springolo F, Ianni A, Casarsa S, Zavarise D, et al. The use of uterine artery Doppler as a screening test for pre-eclampsia. Italian Journal of Gynaecology and Obstetrics. 2003;15(1):15-20.	No relevant outcomes reported
Martis R, Emilia O, Nurdiati DS, Brown J. Intermittent auscultation (IA) of fetal heart rate in labour for fetal well-being. Cochrane Database of Systematic Reviews. 2017;2017 (2) (no pagination)(CD008680).	Study type not eligible
Mastrolia SA, Novack L, Thachil J, Rabinovich A, Pikovsky O, Klaitman V, et al. LMWH in the prevention of preeclampsia and fetal growth restriction in women without thrombophilia: A systematic review and meta-analysis. Thrombosis and Haemostasis. 2016;116(5):868-78.	Study type not eligible

Reference	Reason for exclusion
McCowan L, Harding J, Roberts A, Barker S, Ford C, Stewart A. A pilot randomized controlled trial of two regimens of fetal surveillance for small-for-gestational-age fetuses with normal results of umbilical artery doppler velocimetry. American journal of obstetrics and gynecology [Internet]. 2000; 182(1 Pt 1):[81-6 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/144/CN-00266144/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Meher S, Duley L, Hunter K, Askie L. Antiplatelet therapy before or after 16 weeks' gestation for preventing preeclampsia: an individual participant data meta-analysis. American Journal of Obstetrics and Gynecology. 2017;216(2):121-8.e2.	Not a relevant pregnant population
Menezes EV, Yakoob MY, Soomro T, Haws RA, Darmstadt GL, Bhutta ZA. Reducing stillbirths: Prevention and management of medical disorders and infections during pregnancy. BMC Pregnancy and Childbirth. 2009;9 (SUPPL. 1) (no pagination)(S4).	Study type not eligible
Messerschmidt A, Baschat A, Linduska N, Kasprian G, Brugger PC, Bauer A, et al. Magnetic resonance imaging of the placenta identifies placental vascular abnormalities independently of Doppler ultrasound. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2011;37(6):717-22.	Diagnostic test/monitoring strategy/intervention not relevant
Metcalfe A, Langlois S, Macfarlane J, Vallance H, Joseph KS. Prediction of obstetrical risk using maternal serum markers and clinical risk factors. Prenatal Diagnosis. 2014;34(2):172-9.	No relevant outcomes reported
Michikata K, Sameshima H, Urabe H, Tokunaga S, Kodama Y, Ikenoue T. The Regional Centralization of Electronic Fetal Heart Rate Monitoring and Its Impact on Neonatal Acidemia and the Cesarean Birth Rate. J Pregnancy. 2016;2016:3658527.	Not a relevant pregnant population
Miller SM. Aspirin prevents preeclampsia and complications. Journal of Family Practice. 2003;52(12):923-4.	Study type not eligible
Monaghan C, Binder J, Thilaganathan B, Morales-Rosello J, Khalil A. Perinatal Loss at Term: The Role of Uteroplacental and Fetal Doppler Assessment. Ultrasound Obstet Gynecol. 2017;24:24.	No relevant outcomes reported
Mone F, Mulcahy C, McParland P, Stanton A, Culliton M, Downey P, et al. An open-label randomized-controlled trial of low dose aspirin with an early screening test for pre-eclampsia and growth restriction (TEST): trial protocol. Contemporary clinical trials [Internet]. 2016; 49:[143-8 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/782/CN-01178782/frame.html.	Study type not eligible
Monier I, Ancel PY, Ego A, Guellec I, Jarreau PH, Kaminski M, et al. Gestational age at diagnosis of early-onset fetal growth restriction and impact on management and survival: a population-based cohort study. BJOG: An International Journal of Obstetrics and Gynaecology. 2017;124(12):1899-906.	Diagnostic test/monitoring strategy/intervention not relevant
Morris R, Say R, Robson S, Kleijnen J, Khan K. Systematic review and meta-analysis of middle cerebral artery Doppler to predict perinatal wellbeing (Provisional abstract). European Journal of Obstetrics and Gynecology and Reproductive Biology [Internet]. 2012; 165(2):[141-55 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/cldare/articles/DARE-12013000920/frame.html.	Study type not eligible
Morris RK, Cnossen JS, Langejans M, Robson SC, Kleijnen J, ter RG, et al. Serum screening with Down's syndrome markers to predict pre-eclampsia and small for gestational age: systematic review and meta-analysis (Structured abstract). BMC Pregnancy and Childbirth [Internet]. 2008; 8:33(2). Available from: http://cochranelibrary-wiley.com/o/cochrane/cldare/articles/DARE-12009100693/frame.html.	Study type not eligible
Morris RK, Oliver EA, Malin G, Khan KS, Meads C. Effectiveness of interventions for the prevention of small-for-gestational age fetuses and perinatal mortality: A review of systematic reviews. Acta Obstetricia et Gynecologica Scandinavica. 2013;92(2):143-51.	Diagnostic test/monitoring strategy/intervention not relevant
Morris RK, Selman TJ, Verma M, Robson SC, Kleijnen J, Khan KS. Systematic review and meta-analysis of the test accuracy of ductus venosus Doppler to predict compromise of fetal/neonatal wellbeing in high risk pregnancies with placental insufficiency. European Journal of Obstetrics Gynecology and Reproductive Biology. 2010;152(1):3-12.	Study type not eligible
Navaratnam K, Alfirevic A, Jorgensen A, Alfirevic Z. Aspirin non-responsiveness in pregnant women at high-risk of pre-eclampsia. European Journal of Obstetrics Gynecology and Reproductive Biology. 2018;221:144-50.	Diagnostic test/monitoring strategy/intervention not relevant
Neilson JP. Biochemical tests of placental function for assessment in pregnancy. Cochrane Database of Systematic Reviews [Internet].	Study type not eligible

Reference	Reason for exclusion
2012; (8). Available from: http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD000108.pub2/abstract.	
Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. Cochrane Database of Systematic Reviews. 2015;2015 (12) (no pagination)(CD000116).	Study type not eligible
Nelson-Piercy C, Powrie R, Borg JY, Rodger M, Talbot DJ, Stinson J, et al. Tinzaparin use in pregnancy: An international, retrospective study of the safety and efficacy profile. European Journal of Obstetrics Gynecology and Reproductive Biology. 2011;159(2):293-9.	Diagnostic test/monitoring strategy/intervention not relevant
Nguyen TH, Larsen T, Engholm G, Moller H. A discrepancy between gestational age estimated by last menstrual period and biparietal diameter may indicate an increased risk of fetal death and adverse pregnancy outcome. British Journal of Obstetrics and Gynaecology. 2000;107(9):1122-9.	No relevant outcomes reported
Nunes I, Ayres-de-Campos D, Ugwumadu A, Amin P, Banfield P, Nicoll A, et al. Central Fetal Monitoring With and Without Computer Analysis: A Randomized Controlled Trial. Obstet Gynecol. 2017;129(1):83-90.	Diagnostic test/monitoring strategy/intervention not relevant
Nuriyeva G, Kose S, Tuna G, Kant M, Akis M, Altunyurt S, et al. A prospective study on first trimester prediction of ischemic placental diseases. Prenatal diagnosis [Internet]. 2017; 37(4):[341-9 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/166/CN-01374166/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Odibo AO, Goetzinger KR, Odibo L, Tuuli MG. Early prediction and aspirin for prevention of pre-eclampsia (EPAPP) study: a randomized controlled trial. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;46(4):414-8.	Diagnostic test/monitoring strategy/intervention not relevant
O'Gorman N, Wright D, Rolnik DL, Nicolaides KH, Poon LC. Study protocol for the randomised controlled trial: Combined multimarker screening and randomised patient treatment with ASpirin for evidence-based PREeclampsia prevention (ASPRE). BMJ Open. 2016;6 (6) (no pagination)(e011801).	Study type not eligible
Ohana O, Holcberg G, Sergienko R, Sheiner E. Risk factors for intrauterine fetal death (19882009). Journal of Maternal-Fetal and Neonatal Medicine. 2011;24(9):1079-83.	Diagnostic test/monitoring strategy/intervention not relevant
Onalan R, Onalan G, Gunenc Z, Karabulut E. Combining 2nd-trimester maternal serum homocysteine levels and uterine artery doppler for prediction of preeclampsia and isolated intrauterine growth restriction. Gynecologic and Obstetric Investigation. 2006;61(3):142-8.	Diagnostic test/monitoring strategy/intervention not relevant
O'Neill SM, Kearney PM, Kenny LC, Khashan AS, Henriksen TB, Lutomski JE, et al. Caesarean delivery and subsequent stillbirth or miscarriage: systematic review and meta-analysis. PLoS ONE. 2013;8(1):e54588.	Study type not eligible
Onen A, Cemal Ark H. Fetal thoracic aorta doppler in cases with intrauterine growth restriction. Clinical and Experimental Obstetrics and Gynecology. 2001;28(3):168-70.	Diagnostic test/monitoring strategy/intervention not relevant
Ong GSY, Hadlow NC, Brown SJ, Lim EM, Walsh JP. Does the thyroid-stimulating hormone measured concurrently with first trimester biochemical screening tests predict adverse pregnancy outcomes occurring after 20 weeks gestation? Journal of Clinical Endocrinology and Metabolism. 2014;99(12):E2668-E72.	No relevant outcomes reported
O'Sullivan O, Stephen G, Martindale E, Heazell AEP. Predicting poor perinatal outcome in women who present with decreased fetal movements. Journal of Obstetrics and Gynaecology. 2009;29(8):705-10.	Diagnostic test/monitoring strategy/intervention not relevant
Ozyuncu O, Saygan-Karamursel B, Armangil D, Onderoglu LS, Yigit S, Velipasaoglu M, et al. Fetal arterial and venous Doppler in growth restricted fetuses for the prediction of perinatal complications. Turkish Journal of Pediatrics. 2010;52(4):384-92.	Diagnostic test/monitoring strategy/intervention not relevant
Palma-Dias RS, Fonseca MM, Brietzke E, Fritsch A, Schlatter D, Maurmann CB, et al. Screening for placental insufficiency by transvaginal uterine artery Doppler at 22-24 weeks of gestation. Fetal Diagn Ther. 2008;24(4):462-9.	Setting not relevant/study not in pregnant women
Papadakis K, Karunyam B. The role of pregnancy-associated plasma protein A (PAPP-A) in obstetrics and fetal medicine. BJOG: an international journal of obstetrics and gynaecology [Internet]. 2014; 121:[30 p.]. Available from: http://cochranelibrary-	Study type not eligible

Reference	Reason for exclusion
wiley.com/o/cochrane/clcentral/articles/251/CN-01023251/frame.html.	
Papageorghiou AT, Yu CK, Cicero S, Bower S, Nicolaides KH. Second-trimester uterine artery Doppler screening in unselected populations: a review (Structured abstract). Journal of Maternal-Fetal and Neonatal Medicine [Internet]. 2002; 12(2):[78-88 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/cldare/articles/DARE-12003003083/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Pattison N, McCowan L. Cardiotocography for antepartum fetal assessment. Cochrane database of systematic reviews (Online). 2000(2):CD001068.	Study type not eligible
Pedersen NG, Sperling L, Wojdemann KR, Larsen SO, Tabor A. First trimester growth restriction and uterine artery blood flow in the second trimester as predictors of adverse pregnancy outcome. European Journal of Obstetrics Gynecology and Reproductive Biology. 2013;168(1):20-5.	Diagnostic test/monitoring strategy/intervention not relevant
Peyronnet V, Sibiude J, Mandelbrot L, Kayem G. Impact of the detection of small for gestational age fetuses on the neonatal prognosis. Gynecologie Obstetrique Fertilite et Senologie. 2018;46(2):71-7.	Setting not relevant/study not in pregnant women
Potti S, Berghella V. ST waveform analysis versus cardiotocography alone for intrapartum fetal monitoring: A meta-analysis of randomized trials. American Journal of Perinatology. 2012;29(8):657-64.	Study type not eligible
Radoi V, Bohiltea LC. Pregnancy-associated plasma protein A and pregnancy outcomes. Ginecoro. 2009;5(1):16-20.	No relevant outcomes reported
Roberge S, Nicolaides KH, Demers S, Villa P, Bujold E. Prevention of perinatal death and adverse perinatal outcome using low-dose aspirin: a meta-analysis. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2013;41(5):491-9.	Study type not eligible
Rodger MA, Gris JC, de Vries JIP, Martinelli I, Rey E, Schleussner E, et al. Low-molecular-weight heparin and recurrent placenta-mediated pregnancy complications: a meta-analysis of individual patient data from randomised controlled trials. The Lancet. 2016;388(10060):2629-41.	Study type not eligible
Roma E, Arnau A, Berdala R, Bergos C, Montesinos J, Figueras F. Ultrasound screening for fetal growth restriction at 36 vs 32weeks' gestation: a randomized trial (ROUTE). Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;46(4):391-7.	Diagnostic test/monitoring strategy/intervention not relevant
Rossi AC, Mullin PM. Prevention of pre-eclampsia with low-dose aspirin or vitamins C and e in women at high or low risk: A systematic review with meta-analysis. European Journal of Obstetrics Gynecology and Reproductive Biology. 2011;158(1):9-16.	Study type not eligible
Saccone G, Berghella V. Induction of labor at full term in uncomplicated singleton gestations: A systematic review and metaanalysis of randomized controlled trials. American Journal of Obstetrics and Gynecology. 2015;213(5):629-36.	Not a relevant pregnant population
Saccone G, Schuit E, Amer-Wahlin I, Xodo S, Berghella V. Electrocardiogram st analysis during labor: A systematic review and meta-Analysis of randomized controlled trials. Obstetrics and Gynecology. 2016;127(1):127-35.	Study type not eligible
Saemundsson Y, Svantesson H, Gudmundsson S. Abnormal uterine artery Doppler in pregnancies suspected of a SGA fetus is related to increased risk of recurrence during next pregnancy. Acta Obstetricia et Gynecologica Scandinavica. 2009;88(7):814-7.	Diagnostic test/monitoring strategy/intervention not relevant
Salahuddin S, Wenger JB, Zhang D, Thadhani R, Karumanchi SA, Rana S. KRYPTOR-automated angiogenic factor assays and risk of preeclampsia-related adverse outcomes. Hypertension in Pregnancy. 2016;35(3):330-45.	Diagnostic test/monitoring strategy/intervention not relevant
Santoro R, Lannaccaro P, Prejano S, Muleo G. Efficacy and safety of the long-term administration of low-molecular-weight heparins in pregnancy. Blood Coagulation and Fibrinolysis. 2009;20(4):240-3.	Not a relevant pregnant population
Scala C, Bhide A, Familiari A, Pagani G, Khalil A, Papageorghiou A, et al. Number of episodes of reduced fetal movement at term: Association with adverse perinatal outcome. American Journal of Obstetrics and Gynecology. 2015;213(5):678e1-e6.	Diagnostic test/monitoring strategy/intervention not relevant

Reference	Reason for exclusion
Scandiuzzi RM, Prado CA, Araujo Junior E, Duarte G, Quintana SM, da Silva Costa F, et al. Maternal uterine artery Doppler in the first and second trimesters as screening method for hypertensive disorders and adverse perinatal outcomes in low-risk pregnancies. Obstet. 2016;59(5):347-56.	Setting not relevant/study not in pregnant women
Schleussner E, Lehmann T, Kahler C, Schneider U, Schlembach D, Groten T. Impact of the nitric oxide-donor pentaerythrityltetranitrate on perinatal outcome in risk pregnancies: A prospective, randomized, double-blinded trial. Journal of Perinatal Medicine. 2014;42(4):507-14.	Diagnostic test/monitoring strategy/intervention not relevant
Schneuer F, Nassar N, Khambalia A, Tasevski V, Guilbert C, Ashton A, et al. First trimester screening of maternal placental protein 13 for predicting preeclampsia and small for gestational age: in-house study and systematic review (Provisional abstract). Placenta [Internet]. 2012; 33(9):[735-40 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/cldare/articles/DARE-12012041894/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Schneuer FJ, Roberts CL, Ashton AW, Guilbert C, Tasevski V, Morris JM, et al. Angiopoietin 1 and 2 serum concentrations in first trimester of pregnancy as biomarkers of adverse pregnancy outcomes. American Journal of Obstetrics and Gynecology. 2014;210(4):345.e1-.e9.	No relevant outcomes reported
Schwarze A, Nelles I, Krapp M, Friedrich M, Schmidt W, Diedrich K, et al. Doppler ultrasound of the uterine artery in the prediction of severe complications during low-risk pregnancies. Archives of Gynecology and Obstetrics. 2005;271(1):46-52.	No relevant outcomes reported
Scifres CM, Stamilio D, MacOnes GA, Odibo AO. Predicting Perinatal Mortality in Preterm Intrauterine Growth Restriction. American Journal of Perinatology. 2009;26(10):723-8.	Diagnostic test/monitoring strategy/intervention not relevant
Scott F, Coates A, McLennan A. Pregnancy outcome in the setting of extremely low first trimester PAPP-A levels. Aust N Z J Obstet Gynaecol. 2009;49(3):258-62.	No relevant outcomes reported
Senat MV, Tsatsaris V. Prenatal management of isolated IUGR. [French]. Journal de Gynecologie Obstetrique et Biologie de la Reproduction. 2013;42(8):941-65.	Setting not relevant/study not in pregnant women
Sentilhes L, Bouet PE, Mezzadri M, Combaud V, Madzou S, Biquard F, et al. [Assessment of the benefit-harm balance depending on gestationnal age to induce delivery for post-term pregnancies]. J Gynecol Obstet Biol Reprod (Paris). 2011;40(8):747-66.	Setting not relevant/study not in pregnant women
Simcox LE, Myers JE, Cole TJ, Johnstone ED. Fractional fetal thigh volume in the prediction of normal and abnormal fetal growth during the third trimester of pregnancy. American Journal of Obstetrics and Gynecology. 2017;217(4):453.e1-.e12.	Diagnostic test/monitoring strategy/intervention not relevant
Skornick-Rapaport A, Maslovitz S, Kupferminc M, Lessing JB, Many A. Proposed management for reduced fetal movements: Five years' experience in one medical center. Journal of Maternal-Fetal and Neonatal Medicine. 2011;24(4):610-3.	Diagnostic test/monitoring strategy/intervention not relevant
Skrøstad RB, Eik-Nes SH, Sviggum O, Johansen OJ, Salvesen K, Romundstad PR, et al. A randomized controlled trial of third-trimester routine ultrasound in a non-selected population. Acta obstetricia ET gynecologica scandinavica. 2013;92(12):1353-1360.	No relevant outcomes reported
Smith GCS, 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. Obstetrics and Gynecology. 2006;107(1):161-6.	No relevant outcomes reported
Smith GCS, Stenhouse EJ, Crossley JA, Aitken DA, Cameron AD, Michael Connor J. Early pregnancy levels of pregnancy-associated plasma protein A and the risk of intrauterine growth restriction, premature birth, preeclampsia, and stillbirth. Journal of Clinical Endocrinology and Metabolism. 2002;87(4):1762-7.	No relevant outcomes reported
Soregaroli M, Valcamonico A, Scalvi L, Danti L, Frusca T. Late normalisation of uterine artery velocimetry in high risk pregnancy. European Journal of Obstetrics Gynecology and Reproductive Biology. 2001;95(1):42-5.	Diagnostic test/monitoring strategy/intervention not relevant
Spencer CA, Allen VM, Flowerdew G, Dooley K, Dodds L. Low levels of maternal serum PAPP-A in early pregnancy and the risk of adverse outcomes. Prenatal Diagnosis. 2008;28(11):1029-36.	Diagnostic test/monitoring strategy/intervention not relevant

Reference	Reason for exclusion
Spencer K, Cowans NJ, Avgidou K, Nicolaides KH. First-trimester ultrasound and biochemical markers of aneuploidy and the prediction of impending fetal death. Ultrasound Obstet Gynecol. 2006;28(5):637-43.	No relevant outcomes reported
Stampalija T, Gyte GM, Alfirevic Z. Utero-placental Doppler ultrasound for improving pregnancy outcome. Cochrane Database of Systematic Reviews [Internet]. 2010; (9). Available from: http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD008363.pub2/abstract.	Study type not eligible
Stock SJ, Ferguson E, Duffy A, Ford I, Chalmers J, Norman JE. Outcomes of elective induction of labour compared with expectant management: Population based study. BMJ (Online). 2012;344 (7857) (no pagination)(e2838).	Not a relevant pregnant population
Stock SJ, Ferguson E, Duffy A, Ford I, Chalmers J, Norman JE. Outcomes of Induction of Labour in Women with Previous Caesarean Delivery: A Retrospective Cohort Study Using a Population Database. PLoS ONE. 2013;8 (4) (no pagination)(e60404).	Diagnostic test/monitoring strategy/intervention not relevant
Syngelaki A, Bredaki FE, Vaikousi E, Maiz N, Nicolaides KH. Body mass index at 11-13 weeks' gestation and pregnancy complications. Fetal Diagnosis and Therapy. 2011;30(4):250-65.	No relevant outcomes reported
Tajik P, Van Wyk L, Boers KE, Le Cessie S, Zafarmand MH, Roumen F, et al. Which intrauterine growth restricted fetuses at term benefit from early labour induction? A secondary analysis of the DIGITAT randomised trial. European Journal of Obstetrics Gynecology and Reproductive Biology. 2014;172(1):20-5.	Diagnostic test/monitoring strategy/intervention not relevant
Tan KH, Sabapathy A. Maternal glucose administration for facilitating tests of fetal wellbeing. Cochrane Database of Systematic Reviews. 2009;(4) (no pagination)(CD003397).	Study type not eligible
Tan KH, Smyth RM, Wei X. Fetal vibroacoustic stimulation for facilitation of tests of fetal wellbeing. Cochrane Database of Systematic Reviews. 2013;2013 (12) (no pagination)(CD002963).	Study type not eligible
Tavor O, Shohat M, Lipitz S. The relationship between perinatal outcome of singleton pregnancies and isolated highly elevated levels of maternal serum human chorionic gonadotropin at mid-gestation. Israel Medical Association Journal. 2007;9(7):509-12.	No relevant outcomes reported
Tayyar A, Guerra L, Wright A, Wright D, Nicolaides KH. Uterine artery pulsatility index in the three trimesters of pregnancy: effects of maternal characteristics and medical history. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;45(6):689-97.	Diagnostic test/monitoring strategy/intervention not relevant
Toal M, Chaddha V, Windrim R, Kingdom J. Ultrasound Detection of Placental Insufficiency in Women With Elevated Second Trimester Serum Alpha-Fetoprotein or Human Chorionic Gonadotropin. Journal of Obstetrics and Gynaecology Canada. 2008;30(3):198-206.	No relevant outcomes reported
Triunfo S, Crispi F, Gratacos E, Figueras F. Prediction of delivery of small-for-gestational-age neonates and adverse perinatal outcome by fetoplacental Doppler at 37 weeks' gestation. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2017;49(3):364-71.	Diagnostic test/monitoring strategy/intervention not relevant
Tsiakkas A, Duvdevani N, Wright A, Wright D, Nicolaides KH. Serum placental growth factor in the three trimesters of pregnancy: effects of maternal characteristics and medical history. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;45(5):591-8.	Diagnostic test/monitoring strategy/intervention not relevant
Tsiakkas A, Duvdevani N, Wright A, Wright D, Nicolaides KH. Serum soluble fms-like tyrosine kinase-1 in the three trimesters of pregnancy: effects of maternal characteristics and medical history. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;45(5):584-90.	Diagnostic test/monitoring strategy/intervention not relevant
Tuuli MG, Cahill A, Stamilio D, MacOnes G, Odibo AO. Comparative efficiency of measures of early fetal growth restriction for predicting adverse perinatal outcomes. Obstetrics and Gynecology. 2011;117(6):1331-40.	No relevant outcomes reported
Urquhart C, Currell R, Harlow F, Callow L. Home uterine monitoring for detecting preterm labour. Cochrane Database of Systematic Reviews. 2017;2017 (2) (no pagination)(CD006172).	Study type not eligible

Reference	Reason for exclusion
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 in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2016;47(2):203-9.	No relevant outcomes reported
Van Bulck B, Kalakoutis GM, Sak P, Schneider KTM, Major T, Karpathios SE, et al. A randomised trial of timed delivery for the compromised preterm fetus: Short term outcomes and bayesian interpretation. BJOG: An International Journal of Obstetrics and Gynaecology. 2003;110(1):27-32.	Diagnostic test/monitoring strategy/intervention not relevant
Van Ravenswaaij R, Tesselaar-Van Der Goot M, De Wolf S, Van Leeuwen-Spruijt M, Visser GHA, Schielen PCJI. First-trimester serum PAPP-A and fbeta-hCG concentrations and other maternal characteristics to establish logistic regression-based predictive rules for adverse pregnancy outcome. Prenatal Diagnosis. 2011;31(1):50-7.	No relevant outcomes reported
Van 't Hooft J, Opmeer BC, Teune MJ, Versluis L, Mol BWJ. Costs and health outcomes of effectiveness studies in obstetrics: A budget impact analysis of 8 obstetric effectiveness studies. [Dutch]. Nederlands Tijdschrift voor Geneeskunde. 2014;158(3).	Diagnostic test/monitoring strategy/intervention not relevant
Vasapollo B, Novelli GP, Valensise H. Total vascular resistance and left ventricular morphology as screening tools for complications in pregnancy. Hypertension. 2008;51(4):1020-6.	Diagnostic test/monitoring strategy/intervention not relevant
Verlijsdonk JW, Winkens B, Boers K, Scherjon S, Roumen F. Suspected versus non-suspected small-for-gestational age fetuses at term: Perinatal outcomes. Journal of Maternal-Fetal and Neonatal Medicine. 2012;25(7):938-43.	Study type not eligible
Vink J, Hickey K, Ghidini A, Deering S, Mora A, Poggi S. Earlier gestational age at ultrasound evaluation predicts adverse neonatal outcomes in the preterm appropriate-for-gestational-age fetus with idiopathic oligohydramnios. American Journal of Perinatology. 2009;26(1):21-5.	Not a relevant pregnant population
Vos A, Voorst S, Waelput A, Jong-Potjer L, Bonsel G, Steegers E, et al. Effectiveness of score card-based antenatal risk selection, care pathways, and multidisciplinary consultation in the Healthy Pregnancy 4 All study (HP4ALL): study protocol for a cluster randomized controlled trial. Trials [Internet]. 2015; 16:[8 p.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/270/CN-01111270/frame.html.	Setting not relevant/study not in pregnant women
Walfisch A, Hallak M, Mazor M. Individualized risk assessment for adverse pregnancy outcome by uterine artery Doppler at 23 weeks. Obstetrics and Gynecology. 2001;98(3):369-73.	Diagnostic test/monitoring strategy/intervention not relevant
Walker MG, Hindmarsh PC, Geary M, Kingdom JCP. Sonographic Maturation of the Placenta at 30 to 34 Weeks Is Not Associated With Second Trimester Markers of Placental Insufficiency in Low-risk Pregnancies. Journal of Obstetrics and Gynaecology Canada. 2010;32(12):1134-9.	Diagnostic test/monitoring strategy/intervention not relevant
Westergaard HB, Langhoff-Roos J, Lingman G, Marsal K, Kreiner S. A critical appraisal of the use of umbilical artery Doppler ultrasound in high-risk pregnancies: Use of meta-analyses in evidence-based obstetrics. Ultrasound in Obstetrics and Gynecology. 2001;17(6):466-76.	Study type not eligible
Whitehead CL, McNamara H, Walker SP, Alexiadis M, Fuller PJ, Vickers DK, et al. Identifying late-onset fetal growth restriction by measuring circulating placental RNA in the maternal blood at 28 weeks' gestation. Am J Obstet Gynecol. 2016;214(4):521.e1-.e8.	Diagnostic test/monitoring strategy/intervention not relevant
Whitworth M, Bricker L, Mullan C. Ultrasound for fetal assessment in early pregnancy. The Cochrane database of systematic reviews. 2015;7:CD007058.	Study type not eligible
Williams K, Farquharson D, Bebbington M, Dansereau J, Galerneau F, Wilson R, et al. Screening for fetal well-being in a high-risk pregnant population comparing the nonstress test with umbilical artery Doppler velocimetry: a randomized controlled clinical trial. American journal of obstetrics and gynecology [Internet]. 2003; 188(5):[1366-71 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/453/CN-00437453/frame.html.	No relevant outcomes reported
Wilson RD, Audibert F, Allen VM, Blight C, Brock JA, Desilets VA, et al. Obstetrical Complications Associated With Abnormal Maternal	Study type not eligible

Reference	Reason for exclusion
Serum Markers Analytes. Journal of Obstetrics and Gynaecology Canada. 2008;30(10):918-32.	
Winer N, Branger B, Azria E, Tsatsaris V, Philippe H, Rozé J, et al. L-Arginine treatment for severe vascular fetal intrauterine growth restriction: a randomized double-bind controlled trial. Clinical nutrition (edinburgh, scotland) [Internet]. 2009; 28(3):[243-8 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/952/CN-00706952/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Wright D, Papadopoulos S, Silva M, Wright A, Nicolaides KH. Serum free beta-human chorionic gonadotropin in the three trimesters of pregnancy: effects of maternal characteristics and medical history. Ultrasound Obstet Gynecol. 2015;46(1):51-9.	Diagnostic test/monitoring strategy/intervention not relevant
Wright D, Silva M, Papadopoulos S, Wright A, Nicolaides KH. Serum pregnancy-associated plasma protein-A in the three trimesters of pregnancy: effects of maternal characteristics and medical history. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2015;46(1):42-50.	Diagnostic test/monitoring strategy/intervention not relevant
Wu L, Richardson ML, Dubinsky T. Predicting Adverse Neonatal Outcome Especially When Gestational Age Is Uncertain: Utility of Sonographic Measurement of Fetal Abdominal Wall Thickness. Ultrasound Quarterly. 2017;33(3):208-12.	No relevant outcomes reported
Yigiter AB, Kavak ZN, Durukan B, Isci H, Uzuner A, Uyar E, et al. Placental volume and vascularization flow indices by 3D power Doppler US using VOCAL technique and correlation with IGF-1, free beta-hCG, PAPP-A, and uterine artery Doppler at 11-14 weeks of pregnancy. Journal of Perinatal Medicine. 2011;39(2):137-41.	Diagnostic test/monitoring strategy/intervention not relevant
Zhang J, Mikolajczyk R, Grewal J, Neta G, Klebanoff M. Prenatal application of the individualized fetal growth reference. American journal of epidemiology [Internet]. 2011; 173(5):[539-43 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/175/CN-00888175/frame.html.	Diagnostic test/monitoring strategy/intervention not relevant
Zhang J, Troendle J, Meikle S, Klebanoff MA, Rayburn WF. Isolated oligohydramnios is not associated with adverse perinatal outcomes. BJOG: An International Journal of Obstetrics and Gynaecology. 2004;111(3):220-5.	Diagnostic test/monitoring strategy/intervention not relevant

Appendix 3 – Summary and appraisal of individual studies

Data extraction

Table 21. Studies relevant to criterion 4

Question 1

Study reference	Akolekar 2016a
Study Design	Design Prospective cohort study Objective To investigate whether measurement of maternal serum placental growth factor (PIGF) at 11 to 13 weeks' gestation improves the performance of screening for stillbirths that is achieved by a combination of maternal factors and PAPP-A, DV-PIV and UT-PI, and evaluate the performance of screening of this model for all stillbirths and those due to impaired placentation and unexplained or other causes. Dates March 2006 to October 2015 Country UK Setting King's College Hospital and Medway Maritime Hospital
Population Characteristics	Patient recruitment Women attending for routine pregnancy care at 11+0 to 13+6 weeks' gestation, who gave written informed consent. Pregnancies with aneuploidies, major fetal abnormalities, those ending in miscarriage, termination of pregnancy or intrapartum stillbirths were excluded. Data collection Data on pregnancy outcome was obtained from the maternity hospital records or the general practitioners of women. The hospital maternity records of all women with antepartum stillbirths were reviewed to determine if the death was associated with preeclampsia, abruption or the birthweight was <10th percentile for gestational age or it was due to other reasons or unexplained. Definition of stillbirth [or SGA<3rd centile] Unclear, but study only included pregnancies that delivered a phenotypically normal live birth or stillbirth at ≥24 weeks' gestation. Prevalence of stillbirth in the study In total there were 227 (0.49%) antepartum stillbirths out of 45,452 singleton pregnancies; 131 (58%) of these stillbirths were secondary to impaired placentation and 109 (41%) were due to other or unexplained causes. Sample size N screened/invited = NR N eligible = 45,452 N enrolled = 45,452 N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 45,452 N excluded from analysis = 0 N included in analysis = 45,452 (227 stillbirths)

Study reference	**Akolekar 2016a**
	Demographics • Maternal age, years, median (IQR) o Live births: 31.1 (26.5 to 34.8) o Stillbirths: 32.6 (26.7 to 36.1) • Parity, n (%) • *Nulliparous, n (%)* o Live births: 699 (1.5) o Stillbirths: 15 (6.6) • Smoking, n (%) o Live births: 4,335 (9.6) o Stillbirths: 25 (11.0) • Comorbidities: • *Chronic hypertension, n (%)* o Live births: 699 (1.5) o Stillbirths: 15 (6.6) • *Systemic lupus erythematosus/ antiphospholipid syndrome, n (%)* o Live births: 97 (0.2) o Stillbirths: 2 (0.9) • *Diabetes mellitus, n (%)* o Live births: 435 (1.0) o Stillbirths: 9 (4.0) • Previous pregnancy complications: • *Previous miscarriage, n (%)* o Live births: 574 (1.3) o Stillbirths: 2 (0.9) • *Previous stillbirth, n (%)* o Live births: 360 (0.8) o Stillbirths: 10 (4.4) • *Previous SGA, n (%)* o Live births: 1,494 (3.3) o Stillbirths: 9 (4.0)
Screening Method	Index test • Maternal risk factors • PAPP-A • PLGF • DV-PIV • UT-PI Risk of stillbirth was calculated using a multivariate logistic regression of maternal characteristics and history. Measured values from biochemical and US tests were transformed into multiples of median (MoMs) and univariate and bivariate logistic regression analyses were used to determine if each biochemical or biophysical marker had a contribution to stillbirth risk and if addition of PlGF improved the performance of screening. The screening performance was determined on a receiver-operating characteristics (ROC) curve analysis from the distribution of patient specific risks calculated through the logistic regression, and sensitivity at various FPR (1-specificity) levels was reported. Reference standard Unclear Data on presence of stillbirth was abstracted form hospital maternity records and it appears that stillbirths associated with PE, placental abruption or SGA were considered as "due to abnormal placentation", whereas stillbirths due to other or unexplained causes were included under "all stillbirths".
Test Accuracy	**All stillbirths (includes those with unexplained causes)** Sensitivity for maternal factors + PlGF + UT-PI + DV-PIV 31.7% (25.7 to 37.8) at 95% specificity (5% FPR) 41.9% (35.5 to 48.3) at 90% specificity (10% FPR)

Study reference	Akolekar 2016a
	Stillbirths due to abnormal placentation Sensitivity for maternal factors + PlGF 40.5% (32.1 to 48.9) at 95% specificity (5% FPR) 51.1% (42.5 to 59.7) at 90% specificity (10% FPR) Sensitivity for maternal factors + DV-PIV 28.2% (20.5 to 35.9) at 95% specificity (5% FPR) 37.4% (29.1 to 45.7) at 90% specificity (10% FPR) Sensitivity for maternal factors + UT-PI Sensitivity 35.1% (26.9 to 43.3) at 95% specificity (5% FPR) 45.8% (37.3 to 54.3) at 90% specificity (10% FPR) Sensitivity for maternal factors + PlGF + DV-PIV 42.0% (33.6 to 50.5) at 95% specificity (5% FPR) 51.4% (42.8 to 60.0) at 90% specificity (10% FPR) Sensitivity for maternal factors + UT-PI + DV-PIV 36.6% (28.1 to 44.5) at 95% specificity (5% FPR) 48.1% (39.5 to 56.7) at 90% specificity (10% FPR) Sensitivity for maternal factors + PlGF + UT-PI 47.9% (39.4 to 56.5) at 95% specificity (5% FPR) 60.8% (52.4 to 69.2) at 90% specificity (10% FPR) Sensitivity for maternal factors + PlGF + DV-PIV+UT-PI 48.1% (36.2 to 60.0) at 95% specificity (5% FPR) 61.1% (49.5 to 72.7) at 90% specificity (10% FPR)
Authors' Conclusions	The results of the study demonstrate that a high proportion of stillbirths due to impaired placentation can be effectively identified in the first trimester of pregnancy. The extent to which such stillbirths could be prevented remains to be determined.

Study reference	Akolekar 2016b, Aupont 2016
Study Characteristics	Design Prospective cohort study Objective Akolekar 2016: To evaluate the performance of screening for all stillbirths and those due to impaired placentation and unexplained or other causes by a combination of maternal factors, fetal biometry and uterine artery pulsatility index (UT-PI) at 19 to 24 weeks' gestation and compare this performance to that of screening by UT-PI alone. Aupont 2016: To investigate whether measurement of maternal serum placental growth factor (PlGF) at 19 to 24 weeks' gestation improves the performance of screening for stillbirths with a combination of maternal factors, fetal biometry and uterine artery pulsatility index (UT-PI) and evaluate the performance of screening of this model for all stillbirths and those due to impaired placentation and unexplained or other causes. Dates Not reported (though likely to be March 2006 to October 2015, suspected the same cohort as in Akolekar 2016a) Country UK Setting King's College Hospital and Medway Maritime Hospital, UK
Population Characteristics	Patient recruitment Data for this study were derived from prospective screening for adverse obstetric outcomes in women attending for routine pregnancy care at 19+0 to 24+6 weeks' gestation. The inclusion criteria were women with a singleton pregnancy who delivered a phenotypically normal live birth or stillbirth ≥24 weeks' gestation. Pregnancies with aneuploidies, major fetal abnormalities, those ending in miscarriage, termination of pregnancy or intrapartum

Study reference	Akolekar 2016b, Aupont 2016

stillbirths were excluded.

Data collection
Data on pregnancy outcome were obtained from the maternity hospital records or the general practitioners of women. The hospital maternity records of all women with antepartum stillbirths were reviewed to determine if the death was associated with preeclampsia, abruption or the birthweight was <10th percentile for gestational age or it was due to other causes or was unexplained.

Data on maternal serum PlGF was available in 9,956 pregnancies, including 86 stillbirths. In all stillbirths and subgroups of stillbirths, mean and SDs of log_{10}MoM PlGF values were estimated; the values for PlGF were then simulated in the remaining cases in study population, based on the bivariate Gaussian distributions of the marker in stillbirths and live births, defined by the mean and SD (log_{10}MoM).

Definition of stillbirth [or SGA<3rd centile]
Unclear, but study only included pregnancies that delivered a phenotypically normal live birth or stillbirth at ≥24 weeks' gestation

Prevalence of stillbirth in the study
In total there were 268 stillbirths out of 70,003 singleton pregnancies; 159 (59%) of these stillbirths were secondary to impaired placentation and 109 (41%) were due to other or unexplained causes

Sample size
N screened/invited = NR
N eligible = 70,003
N enrolled = 70,003
N excluded (with reason) = 0
N lost to follow-up = 0
N completed = 70,003
N excluded from analysis = 0
N included in analysis = 268 (100% of stillbirths)

Demographics
Maternal age, years, median (IQR)

- Live births: 30.5 (25.8 to 34.5)
- Stillbirths: 30.5 (25.8 to 35.4)

Weight (kg), median (IQR)

- Live births: 67.0 (59.2 to 78.0)
- Stillbirths: 73.4 (63.7 to 85.2)*

Nulliparous, n (%):

- Live births: 34,279 (49.2)
- Stillbirths: 132 (49.3)

Cigarette smoker, n (%):

- Live births: 7478 (10.7)
- Stillbirths: 35 (13.1)

Comorbidities
Chronic hypertension, n (%)

- Live births: 1,031 (1.5)
- Stillbirths: 17 (6.3)*

APS/SLE, n (%)

- Live births: 132 (0.2)
- Stillbirths: 4 (1.5)

Pre-existing diabetes mellitus, n (%)

- Live births: 638 (0.9)
- Stillbirths: 7 (2.6)

Previous pregnancy complications
Previous miscarriage, n (%)

- Live births: 883 (1.3)
- Stillbirths: 4 (1.5)

Previous stillbirth, n (%)

- Live births: 604 (0.9)

Study reference	Akolekar 2016b, Aupont 2016
	○ Stillbirths: 15 (5.6)* Previous SGA, n (%) ○ Live births: 2,315 (3.3) ○ Stillbirths: 12 (4.5) Comparison of stillbirth groups with live-birth group by chi-square test and Mann–Whitney *U*-test with *post-hoc* Bonferroni correction for multiple comparisons: *p<0.01
	Index test Maternal risk factors Fetal biometry UtA-PI at 19–24 weeks' gestation Maternal characteristics and medical history were recorded and ultrasound examinations were performed for measurement of fetal head circumference (HC), abdominal circumference (AC) and femur length (FL). Transvaginal colour Doppler ultrasound was used to visualize the left and right uterine arteries at the level of the internal os. Pulsed-wave Doppler was then used to obtain waveforms and when three similar consecutive waveforms are obtained the PI was measured, and the mean PI of the two vessels was calculated. Women with a mean uterine artery PI>1.6 were followed up with growth scans at 28, 32 and 36 weeks' gestation. Women with normal uterine artery Doppler received routine antenatal care. Aupont 2016 (second phase): The maternal serum concentration of PlGF at 19-24 weeks' gestation was measured using automated analysers The screening algorithm was derived from multivariable logistic regression analysis of maternal characteristics and history. Univariable and multivariable logistic regression analyses were then used to determine if *a*-priori risk for the maternal factors (HC, AC, FL and Uta-PI MoM had a significant contribution to the prediction of stillbirth. The distribution of patient-specific risks was used to determine the performance of screening by ROC curve analysis, and the DR and FPR were estimated. Reference standard Unclear Data on presence of stillbirth was abstracted form hospital maternity records and it appears that stillbirths associated with PE, placental abruption or SGA were considered as "due to abnormal placentation", whereas stillbirths due to other or unexplained causes were included under "all stillbirths".
Test Accuracy	**All stillbirths** Sensitivity for maternal factors 19.0% (14.3 to 23.7) at 95% specificity (5% FPR) 29.5% (24.0 to 34.9) at 90% specificity (10% FPR) Sensitivity for maternal factors + fetal biometry 32.2% (26.6 to 37.8) at 95% specificity (5% FPR) 42.5% (36.6 to 48.4) at 90% specificity (10% FPR) Sensitivity for maternal factors + UtA-PI 41.8% (35.9 to 47.7) at 95% specificity (5% FPR) 52.6% (46.6 to 58.6) at 90% specificity (10% FPR) Sensitivity for maternal factors + fetal biometry + UtA-PI 45.1% (39.1 to 51.0) at 95% specificity (5% FPR) 54.7% (48.7 to 60.6) at 90% specificity (10% FPR) Sensitivity for maternal factors + UtA-PI + PlGF 50.7% (44.7 to 56.7) at 95% specificity (5% FPR) 57.6% (51.7 to 63.5) at 90% specificity (10% FPR) **Stillbirth from impaired placentation**

Study reference	Akolekar 2016b, Aupont 2016
	Sensitivity for maternal factors 22.6% (16.1 to 29.1) at 95% specificity (5% FPR) 34.0% (26.6 to 41.4) at 90% specificity (10% FPR) Sensitivity for maternal factors + fetal biometry 52.8% (45.0 to 60.6) at 95% specificity (5% FPR) 63.5% (56.0 to 70.9) at 90% specificity (10% FPR) Sensitivity for maternal factors + UtA-PI 62.3% (54.8 to 69.8) at 95% specificity (5% FPR) 73.6% (66.8 to 80.5) at 90% specificity (10% FPR) Sensitivity for maternal factors + fetal biometry + UtA-PI 69.8% (62.7 to 76.9) at 95% specificity (5% FPR) 74.8 (68.1 to 81.6) at 90% specificity (10% FPR) Sensitivity for maternal factors + fetal biometry + UtA-PI + PlGF 76.1% (69.5 to 82.7) at 95% specificity (5% FPR) 83.6% (77.8 to 89.4) at 90% specificity (10% FPR)
Authors' Conclusions	The main findings demonstrate that in the study population, approximately 60% of antepartum stillbirths are due to impaired placentation and 40% are unexplained or due to other causes. The model, which combines maternal factors, UtA-PI and fetal biometry at 19–24 weeks' gestation can potentially predict about 75% of stillbirths due to impaired placentation. The performance of screening is better for stillbirth <32 weeks' gestation (88%) compared to those at term (46%).

Study reference	**Bakalis 2015 (cohort may be included in Akolekar 2016a, Akolekar 2016b, Aupont 2016, Valino 2016a and b, Yerlikaya 2016)**
Study Characteristics	Design Prospective cohort study Objective To investigate the potential value of the cerebroplacental ratio (CPR) at 30–34 weeks' gestation in the prediction of adverse perinatal outcome, by examining the relationship between CPR and birth-weight Z-score according to the rates of stillbirth, Cesarean section for fetal distress, umbilical arterial cord blood Ph <7.0, umbilical venous cord blood Ph <7.1, 5-min Apgar score <7 and admission to NNU or the neonatal intensive care unit (NICU). Dates May 2011 to August 2014 Country UK Setting King's College Hospital, London and Medway Maritime Hospital, Kent
Population Characteristics	Patient recruitment Singleton pregnancies were prospectively screened. Pregnancies with major fetal abnormalities or genetic syndromes diagnosed prenatally or postnatally, and pregnancies with no follow-up, were excluded. Data collection The screening visit at 30+0 to 34+6 weeks' gestation, included recording maternal characteristics and medical history, and estimation of fetal size from trans-abdominal ultrasound measurement of fetal head circumference, abdominal circumference and femur length. Gestational age was determined from measurement of the fetal crown-rump length at 11 to 13 weeks or the fetal head circumference at 19 to 24 weeks. Trans-abdominal colour Doppler ultrasound was used to visualise the UA and MCA. Pulsed-wave Doppler was used to assess impedance to flow, and when three similar waveforms were obtained consecutively the PI was measured. Data on pregnancy outcomes were collected from the hospital maternity records or the general medical practitioners of

Study reference	Bakalis 2015 (cohort may be included in Akolekar 2016a, Akolekar 2016b, Aupont 2016, Valino 2016a and b, Yerlikaya 2016)

the women.

Definition of stillbirth [or SGA<3rd centile]
Unclear

Prevalence of stillbirth in the study
There were 82 stillbirths out of 30,780 included pregnancies (*0.27%)*, of which 75 were antepartum and seven intrapartum

Sample size
N screened/invited = 32,370 women with singleton pregnancies were screened
N eligible = 30,780
N enrolled = 30,780
N excluded (with reason) = 213 pregnancies (0.7%) were excluded for having major fetal abnormalities or genetic syndromes diagnosed prenatally or postnatally as per exclusion criteria
N lost to follow-up = 1,377 (4.3%) were excluded for no follow-up
N completed = 30,780
N excluded from analysis = 0
N included in analysis = 30,780

Demographics

Maternal age, years, median (IQR)
- Live births: 31.3 (26.8 to 35.0)
- Stillbirths: 30.0 (25.7 to 36.2)

Weight (kg), median (IQR)
- Live births: 75.5 (67.8 to 85.7)
- Stillbirths: 83.5 (70.1 to 95.2)

Height (m), median (IQR)
- Live births: 1.65 (1.60 to 1.69)
- Stillbirths: 1.65 (1.62 to 1.68)

Parous, n (%)
- Live births: 15,332 (49.8)
- Stillbirths: 40 (48.8)

Nulliparous, n (%):
- Live births: 15,448 (50.2)
- Stillbirths: 42 (51.2)

Cigarette smoker, n (%):
- Live births: 2,791 (9.1)
- Stillbirths: 11 (13.4)

Comorbidities

Chronic hypertension, n (%)
- Live births: 413 (1.3)
- Stillbirths: 2 (2.4)

APS/SLE, n (%)
- Live births: 58 (0.2)
- Stillbirths: 0 (0.0)

Pre-existing diabetes mellitus, n (%)
Type 1:
- Live births: 107 (0.3)
- Stillbirths: 0 (0.0)

Type 2:
- Live births: 185 (0.6)
- Stillbirths: 0 (0.0)

Gestational diabetes, n (%)
- Live births: 756 (2.5)
- Stillbirths: 2 (2.4)

Pre-eclampsia

Study reference	**Bakalis 2015 (cohort may be included in Akolekar 2016a, Akolekar 2016b, Aupont 2016, Valino 2016a and b, Yerlikaya 2016)**
	○ Live births: 686 (2.2) ○ Stillbirths: 3 (3.7) Obstetric cholestasis ○ Live births: 147 (0.5) ○ Stillbirths: 0 (0.0) Spontaneous rupture of membranes (SROM) ○ Live births: 1,601 (5.2) ○ Stillbirths: 1 (1.2)
Screening Method	Index test Screening for low cerebroplacental ratio (<5th centile) in the prediction of stillbirth The CPR was calculated by dividing MCA-PI MoM by UA-PI MoM. Regression analysis was used to examine the association between log_{10}MoM CPR and birth-weight Z-score in the study population as well as within each weekly interval from the time of assessment to delivery. Univariable and multivariable logistic regression analyses were used to determine if the log_{10}MoM CPR had a significant additional contribution to maternal characteristics, medical history and obstetric factors in predicting adverse outcome. DR, FPR and PPV of screening by CPR were estimated for each adverse outcome Reference standard Unclear
Test Accuracy	**All pregnancies** CPR <5th percentile for all stillbirths Sensitivity: 8.5% (7/82) Specificity 94.8% (FPR 5.2 [1609/30,698]) CPR <5th percentile for stillbirths with birthweight <10th centile Sensitivity: 16.7% (4/24) Specificity: 90.1% (FPR 9.9 [340/3422]) CPR <5th percentile for stillbirths with birthweight >10th centile Sensitivity: 5.2% (3/58) Specificity: 95.3% (FPR 4.7 [1269/27,276]) **Pregnancies delivering ≤2 weeks following assessment** CPR <5th percentile for all stillbirths Sensitivity 20.0% (1/5) Specificity: 76.7% (FPR 23.3 [84/360]) CPR <5th percentile for stillbirths with birthweight <10th centile Sensitivity: 50.0% (1/2) Specificity: 50.4% (FPR 49.6 [56/113]) CPR <5th percentile for stillbirths with birthweight >10th centile: Sensitivity: 0% (0/3) Specificity: 88.7% (FPR 11.3 [28/247]) **Pregnancies delivering >2 weeks following assessment** CPR <5th percentile for all stillbirths Sensitivity: 7.8% (6/77) Specificity: 95.3% (FPR 5.0 [1525/30,338]) CPR <5th percentile for stillbirths with birthweight <10th centile Sensitivity: 13.6% (3/22) Specificity: 91.4% (FPR 8.6 [284/3309]) CPR <5th percentile for stillbirths with birthweight >10th centile

Study reference	**Bakalis 2015 (cohort may be included in Akolekar 2016a, Akolekar 2016b, Aupont 2016, Valino 2016a and b, Yerlikaya 2016)**
	Sensitivity: 5.5% (3/55) Specificity: 90.1% (FPR 4.6 [1241/27,029])
Authors' Conclusions	Assessment by CPR contributed significantly (in addition to maternal characteristics, medical history and obstetric factors) in the prediction of multiple perinatal outcomes but not in the prediction of stillbirth. In general, the PPV of low CPR in the prediction of adverse outcome was higher in SGA than in non-SGA fetuses, particularly in those delivering within two weeks of assessment.

Study reference	**Chaiworapongsa 2017**
Study Characteristics	Design Case-cohort study Objective To determine if maternal plasma concentrations of angiogenic and antiangiogenic factors measured at 24–28 weeks of gestation can predict subsequent fetal death Dates NR Country USA Setting NR
Population Characteristics	Patient recruitment One thousand subjects were randomly selected from a cohort of 4,006 pregnant women enrolled in a previous longitudinal study. The remaining women in the original cohort who had a fetal death, but were not selected in the random sample of 1000 women, were subsequently added to the case cohort. Women who had multiple gestations or any of the following conditions at the time of enrolment were excluded: active vaginal bleeding, severe maternal morbidity (i.e. renal insufficiency, congestive heart disease, chronic respiratory insufficiency), chronic hypertension requiring medication, asthma requiring systemic steroids, requirement of antiplatelet or nonsteroidal antiinflammtory drugs, active hepatitis, or fetal chromosomal abnormalities and congenital abnormalities. Data collection Patients donated maternal plasma in EDTA tubes at time of enrolment, then every 4 weeks until the 24th week of gestation, and biweekly thereafter until delivery. Maternal plasma concentrations of sVEGFR-1 PlGF and sEng were measured by immunoassays. Placentas were examined according to standardised protocols by perinatal pathologists blinded to clinical diagnoses and obstetrical outcomes. Placental lesions consistent with maternal vascular lesions of under-perfusion were diagnosed using criteria established by the Perinatal Section of the Society for Pediatric Pathology, and were classified as 1) villous changes and 2) vascular lesions. Definition of stillbirth [or SGA<3rd centile] Fetal death was diagnosed as the death of the fetus >20 weeks of gestation Prevalence of stillbirth in the study The prevalence of fetal death >24 weeks in the tested cohort was 1.3% (11/840) Sample size N screened/invited = 4,006 pregnant women from a previous longitudinal screening study N eligible = 40,06 pregnant women from a previous longitudinal screening study N enrolled = 1,018 (random sample + 18 further cases of fetal death from the prior study) N excluded (with reason) = NR N lost to follow-up = NR N completed = 1,018 N excluded from analysis = 178 (13 cases and 165 controls) N included in analysis = 840

Study reference	**Chaiworapongsa 2017**
	Demographics Maternal age, years, median (IQR) o No fetal death: 23 (20 to 27) o Fetal death: 22.5 (20 to 29.5) Nulliparity, n (%) o No fetal death: 381 (38.7) o Fetal death: 8 (33.3) Cigarette smoker, n (%): o No fetal death: 206 (20.8) o Fetal death: 5 (20.8) Pre-pregnancy BMI, median (IQR) o No fetal death: 26.6 (22.5 to 32.5) o Fetal death: 26.8 (23.2 to 33.3) Gestational age at delivery, weeks, median (IQR) o No fetal death: 39.1 (37.9 to 40.1) o Fetal death: 28.3 (23 to 31.5) African-American o No fetal death: 921 (92.7) o Fetal death: 23 (95.8) Birthweight, g o No fetal death: 3,172.5 (2,800 to 3,485) o Fetal death: 924.5 (493.5 to 1,400)
Screening Method	Index test Maternal plasma angiogenic and antiangiogenic factor concentrations and ratios (at 24–28 weeks of gestation) for subsequent fetal death (from 24–37.6 weeks of gestation): • sVEGFR-1 • sEng • PlGF • PlGF/sVEGFR-1 • PlGF/sEng Using quantile regression, the percentiles of maternal plasma concentration of analytes and their ratios in a subset of controls were estimated. Linear quantile regression was used iteratively to estimate the quantiles at discrete values of gestation, in narrow windows of gestational age, over which the linear assumption holds. Positive tests were defined as analyte concentrations (or their ratios) <2.5th and 10th centiles (PlGF, PlGF/sVEGFR-1, and PlGF/sEng) or >90th and 97.5th centiles (sVEGFR-1 and sEng). Data from the 11 fetal deaths and 829 controls with available measurements were used to construct receiver operating characteristic (ROC) curves and to determine the sensitivity, specificity, and likelihood ratios (positive and negative). Likelihood ratios for a positive test result >10 and likelihood ratios for a negative test results <0.1 were taken as strong predictive evidence under most circumstances, with 5 to 10 and 0.1 to 0.2 (moderate prediction), <5 and >0.2 minimally predictive, respectively. Reference standard Fetal death, diagnosed as death of the fetus >20 weeks of gestation and confirmed by ultrasound examination.

Test Accuracy

	Cut-off	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
sVEGFR-1	>90	0.55 (0.23 to 0.83)	0.89 (0.87 to 0.91)	4.9 (2.8 to 8.7)	0.51 (0.27 to 0.98)
sVEGFR-1	>97.5	0.27 (0.06 to 0.61)	0.97 (0.95 to 0.98)	8.4 (3 to 23.5)	0.75 (0.52 to 1.08)
sEng	>90	0.64 (0.31 to 0.89)	0.89 (0.87 to 0.91)	5.9 (3.6 to 9.5)	0.41 (0.19 to 0.89)
sEng	>97.5	0.55 (0.23 to 0.83)	0.96 (0.94 to 0.97)	13.7 (7.3 to 25.8)	0.47 (0.25 to 0.90)
PlGF	<10	0.55 (0.23 to	0.87 (0.84 to	4.1 (2.3 to 7.2)	0.52 (0.27 to

Study reference	**Chaiworapongsa 2017**					
			0.83)	0.89)		1.00)
	PlGF	<2.5	0.45 (0.17 to 0.77)	0.95 (0.94 to 0.97)	9.9 (4.8 to 20.3)	0.57 (0.33 to 0.98)
	PlGF/sVEGFR-1	<10	0.64 (0.31 to 0.89)	0.89 (0.87 to 0.91)	5.7 (3.5 to 9.3)	0.41 (0.19 to 0.89)
	PlGF/sVEGFR-1	<2.5	0.55 (0.23 to 0.83)	0.96 (0.95 to 0.97)	14.6 (7.7 to 27.7)	0.47 (0.25 to 0.90)
	PlGF/sEng	<10	0.64 (0.31 to 0.89)	0.87 (0.85 to 0.90)	5.1 (3.1 to 8.2)	0.42 (0.19 to 0.91)
	PlGF/sEng	<2.5	0.55 (0.23 to 0.83)	0.96 (0.94 to 0.97)	13.7 (7.3 to 25.8)	0.47 (0.25 to 0.90)
Authors' Conclusions	A maternal plasma angiogenic index-1 value of <2.5th centile at 24 to 28 weeks of gestation carries a 29-fold increase for the risk of subsequent fetal death, and identifies 55% of these patients as a very low (3.7%) false-positive rate.					

Study reference	**Chaiworapongsa 2013a**
Study Characteristics	Design Prospective longitudinal cohort study Objective To determine whether maternal plasma concentrations of PlGF, sEng, sVEGFR-1, and their ratios at 30–34 weeks of gestation could be used to identify patients at risk for stillbirth, late pre-eclampsia, severe late pre-eclampsia, or delivery of SGA neonates. Dates November 2003 to August 2006 Country Chile Setting Prenatal clinic, Sotero del Rio Hospital (tertiary care centre)
Population Characteristics	Patient recruitment Patients were enrolled in the prenatal clinic of the study hospital and followed until delivery. Inclusion criteria were singleton gestation and 6 to 22 weeks of gestation. Exclusion criteria were: 1) preterm labour, preterm prelabour rupture of membranes, pre-eclampsia or impaired fetal growth at the time of recruitment; 2) known major fetal anomaly or fetal death; 3) active vaginal bleeding; 4) serious medical illness (renal insufficiency, congestive heart disease, chronic respiratory insufficiency or active hepatitis). Data collection At enrolment and each subsequent visit, patients underwent a venepuncture for the collection of maternal blood. Samples were collected every 4 weeks until 24 weeks of gestation and every 2 weeks thereafter until delivery. Gestational age at venepuncture and at delivery were based on best obstetrical estimates with the use of the last menstrual period and the earliest fetal biometric parameters, which were performed at ≤20 weeks of gestation in 98.2% of cases. Customised case report forms and a perinatal database was generated. Data were extracted from medical records by trained research nurses. To account for misclassification, abstracters were trained, the data collection methods were verified, and data logic was monitored. Cases of uncertainty were resolved by iterative discussion among 3 of the authors. Definition of stillbirth [or SGA<3rd centile] Death of a fetus before delivery that was not a consequence of an induced termination of pregnancy (including intrapartum and antepartum stillbirth) Prevalence of stillbirth in the study There were 5 cases (0.4%) of stillbirth Sample size N screened/invited = NR N eligible = NR

Study reference	**Chaiworapongsa 2013a**
	N enrolled = 2,998 N excluded (with reason) = 1,638 (n=503 did not have sample collected at 6–15 weeks, n=578 did not have sample collected at 20–25 weeks, n=204 had no UADV results in the second trimester (20–25 weeks), n=27 delivered ≤34 weeks, n=326 did not have a sample at 30–34 weeks) N lost to follow-up = 91 N completed = 1,269 N excluded from analysis = 0 N included in analysis = 1,269 Demographics for uncomplicated pregnancies (n=886) and stillbirth (n=5) Maternal age, years, mean (SD) o Uncomplicated pregnancy: 26.2 (5.9) o Stillbirth: 27.0 (10.0) Nulliparity, n (%) o Uncomplicated pregnancy: 356 (40.2) o Stillbirth: 2 (40.0) Tobacco use, n (%): o Uncomplicated pregnancy: 93 (10.5) o Stillbirth: 0 BMI, mean (SD) o Uncomplicated pregnancy: 24.6 (4.2) o Stillbirth: 22.4 (1.6) Gestational age at venipuncture, weeks, mean (SD) o Uncomplicated pregnancy: 32.2 (1.1) o Stillbirth: 32 (0.9) Birthweight, g o Uncomplicated pregnancy: 3,505 (399) o Stillbirth: 2,896 (642) Previous complications Pre-eclampsia, n (%) o Uncomplicated pregnancy: 19 (2.1) o Stillbirth: 0
Screening Method	Index test Maternal plasma concentrations and their ratios (PlGF/sVEGFR-1 and PlGF/sEng) for predicting stillbirth Maternal plasma concentrations of sVEGFR-1, PlGF, and sEng were determined by sensitive and specific immunoassays. The laboratory personnel who performed the assays were blinded to the clinical information. Quantile regression was used to calculate median analyte ratio concentrations (PlGF/sVEGFR-1, PlGF/sEng) that were conditional on gestational age among uncomplicated pregnancies (n=886). MoM values were calculated for both analyte ratios for each patient. MoM cutoffs were determined based on inspection of ROC curves for stillbirth. Predictive performance metrics were also calculated. Paired sample nonparametric statistical techniques were used to compare area under the receiver operating characteristic curves (AUC) of models that were constructed with logistic regression for the identification of selected pregnancy outcomes. A McNemar's test was also used to test for differences in sensitivity at a fixed false-positive rate of 15%. A 5% threshold for type I error was used to determine statistical significance. Reference standard Stillbirth, defined as death of a fetus before delivery. Data were extracted from medical records by trained research nurses, but method of diagnosis e.g. postnatal and the involved personnel are unclear.
Test Accuracy	**PlGF/sVEGFR-1 for stillbirth** Sensitivity 80 (28 to 100) Specificity 94 (93 to 96) PPV
	PlGF/sEng for stillbirth Sensitivity 60 (15 to 95) Specificity 89 (87 to 91) PPV

Study reference	Chaiworapongsa 2013a	
	5 (1 to 13)	2 (0.4 to 6)
	NPV	NPV
	100 (99 to 100)	99 (99 to 100)
	FPR	FPR
	6 (4 to 7)	11 (9 to 13)
	FNR	FNR
	20 (0.5 to 72)	40 (5 to 85)
	LR+	LR+
	14.2 (8.7 to 23.3)	5.5 (2.7 to 11.5)
	LR−	LR−
	0.2 (0.04 to 1.22)	0.4 (0.2 to 1.3)

Authors' Conclusions	A maternal plasma concentration of PlGF/sVEGFR-1 ratio <0.12 MoM at 30–34 weeks' gestation was associated with a subsequent stillbirth (aOR, 23). This cutoff had a sensitivity of 80%, a specificity of 94% and a likelihood ratio of a positive result of 14 for the identification of patients destined to have a stillbirth.

Study reference	Chaiworapongsa 2013b
Study Characteristics	Design Case-control Objective To determine whether maternal plasma concentrations of PlGF, sEng, sVEGFR-1, and their ratios at 30–34 weeks of gestation could be used to identify patients at risk for stillbirth, late pre-eclampsia, severe late pre-eclampsia, or delivery of SGA neonates. Dates 2007 to 2009 Country USA or Chile Setting Unclear, but it appears that the study was conducted at Hutzel Women's Hospital, Detroit and/or Sotero del Rio Hospital, Chile
Population Characteristics	Patient recruitment Participants were identified from a cohort of 5,828 singleton pregnancies who were either enrolled in a prospective cohort study in Chile or another cross-sectional protocol from 2007–2009 at Hutzel Women's Hospital in Detroit, US. Patients were enrolled when they presented to the labour and delivery unit with a suspicion of spontaneous preterm labour or medically-indicated preterm birth. There were 31 cases of stillbirth at ≥34 weeks of gestation, five of which were included because they had a plasma sample collected at 30–34 weeks of gestation. Control subjects were identified from uncomplicated pregnancies who both delivered an appropriate weight for gestational age neonate at term and had a plasma sample collected at 30–34 weeks of gestation. Controls were matched to cases 6:1 on gestational age at venepuncture, parity, ethnicity, tobacco use, and BMI. All women provided written informed consent before participation in the study. Data collection Maternal plasma concentrations of sVEGFR-1, sEng, and PlGF were determined by sensitive and specific immunoassays Definition of stillbirth [or SGA<3rd centile] Stillbirth was defined as death of a fetus before delivery (which was not a consequence of an induced termination of pregnancy) Prevalence of stillbirth in the study

Study reference	**Chaiworapongsa 2013b**
	NR Sample size N screened/invited = 5,828 singleton pregnancies N eligible = 31 cases of stillbirth occurred during the study period N enrolled = 5 cases of stillbirth (those who had plasma samples collected at 30–34 weeks of gestation) and 30 controls (matched 6:1) N excluded (with reason) = NR N lost to follow-up = NR N completed = 35 (5 stillbirth and 30 controls) N excluded from analysis = 0 N included in analysis = 35 (5 stillbirth and 30 controls) Demographics Maternal age, years, median (IQR) o Control: 21.5 (19.8 to 23.2) o Fetal death: 26 (21.5 to 36.0) Nulliparity, n (%) o Control: 16 (53.3) o Fetal death: 2 (60) Tobacco use, n (%): o Control: 0 (0) o Fetal death: 0 (0) BMI, median (IQR) o Control: 26.1 (21.1 to 35.8) o Fetal death: 24.6 (17.9 to 44.5) Gestational age at venipuncture, weeks, median (IQR) o Control: 32.9 (32.1 to 33.6) o Fetal death: 33.4 (32 to 33.7) African-American, n (%) o Control: 24 (80) o Fetal death: 4 (80) Birthweight, g o Control: 32,73 (3,165 to 3,478) o Fetal death: 2,305 (1,635 to 3,360)
Screening Method	Index test Maternal plasma concentrations and their ratios (PlGF/sVEGFR-1 and PlGF/sEng) for predicting stillbirth Differences among cases and controls were tested with the Chi-squared test, Fisher exact test, or the Mann-Whitney *U* test where appropriate. AUC was calculated, and sensitivities and specificities were determined with the use of absolute value thresholds for each biomarker ratio that were derived from the inspection of receiver operating characteristic curves Reference standard Stillbirth, defined as death of a fetus before delivery. Method of diagnosis and the involved personnel are unclear.
Test Accuracy	A maternal plasma concentration of PlGF/sVEGFR-1 ≤0.046 or PlGF/sEng ratio ≤11.7 pg/ng at 30–34 weeks had a sensitivity of 80% and a specificity of 93% for the identification of subsequent stillbirth.
Authors' Conclusions	Risk assessment for stillbirth (and severe late preeclampsia) in the third trimester is possible with the determination of maternal plasma concentrations of angiogenic and antiangiogenic factors at 30–34 weeks of gestation.

Study reference	**Conde-Agudelo 2015**
Study Characteristics	Design Systematic literature review and meta-analysis Objective To determine the accuracy of tests performed during the first and/or second trimester of pregnancy to predict

Study reference	**Conde-Agudelo 2015**
	stillbirth in unselected women with singleton, structurally and chromosomally normal fetuses. Dates Inception – 31 May 2014 Country Most (93%) of included studies were conducted in high-income countries. Setting N/A
Population Characteristics	Patient recruitment NA Data collection NA Definition of stillbirth [or SGA<3rd centile] It was noted that a wide variety of definitions for stillbirth exist, but in the review, stillbirth was to include a structurally and chromosomally normal fetus whose death occurred at or after 20 weeks of gestation or with a birth weight ≥350 g before the onset (antepartum), or after the start (intrapartum) of labour, and before the complete expulsion or extraction from its mother. Prevalence of stillbirth in the study The prevalence for stillbirth in cohort studies ranged from 0.6 and 16.4 per 1,000 births (median 4.6 per 1,000 births) Sample size The sample size in cohort studies ranged from 110 to 718,011 (median 9672) women. The number of cases and controlled enrolled in case-control studies ranged from 33 to 701, and 100 to 15,869, respectively. Demographics NA
Screening Method	Index tests Fetoplacental proteins/hormone-related tests, ultrasound imaging-related tests, combinations of tests, and others. Reference standard Various
Test Accuracy	Relevant measures of test accuracy are presented in the main report.
Authors' Conclusions	Currently, there is no clinically useful first-trimester or second-trimester test to predict stillbirth as a sole category. Uterine artery pulsatility index and maternal serum PAPP-A levels appeared to be good predictors of stillbirth related to placental dysfunction disorders.

Study reference	**Dugoff 2004**
Study Characteristics	Design Prospective cohort study Objective To determine whether first-trimester measurements of maternal serum PAPP-A, fβhCG, or nuchal translucency were associated with obstetric complications Dates Unclear Country USA Setting

Study reference	Dugoff 2004
	Multicenter
Population Characteristics	Patient recruitment Women who participated in the study were enrolled in the First and Second Trimester Evaluation of Risk (FASTER) trial, which was a multicentre study to compare the diagnostic performance of several first and second trimester screening markers for Down syndrome. All potential subjects underwent an ultrasound examination to confirm gestational age and to measure nuchal translucency, and women who were confirmed to have a viable singleton pregnancy (with fetal crown-rump length between 36 and 79 mm inclusive) were eligible to participate. Women whose fetuses were diagnosed with anencephaly or a septated cystic hygroma were ineligible for enrolment in the FASTER trial but were followed separately. Data collection Relevant patient history, demographic data, and obstetric history were collected at the time of enrolment in the FASTER trial. After delivery, the clinical coordinators or physicians at each FASTER site obtained pregnancy outcome data by medical record review and/or patient interviews, with an extensive follow-up protocol and predefined criteria. Any abnormal outcome that was reported by a patient was followed up by a formal medical record review. The outcomes examined included spontaneous loss at ≤24 weeks of gestation, fetal loss at >24 weeks of gestation, neonatal loss, preterm delivery, gestational hypertension, preeclampsia, low birth weight, preterm premature rupture of membranes, placenta previa, placental abruption, gestational diabetes mellitus, and macrosomnia. Definition of stillbirth [or SGA<3rd centile] Fetal loss at >24 weeks of gestation Prevalence of stillbirth in the study 0.28% (95 cases of intrauterine fetal death at >24 weeks of gestation) Sample size N screened/invited = NR N eligible = NR N enrolled = 34,411 N excluded (with reason) = 1,016 (fetus with a chromosomal n=221 or structural abnormality n=463; women with insulin-dependent diabetes mellitus n=332) N lost to follow-up = NR N completed = 33,395. The study population that was available for the analysis included all women in the FASTER trial for whom PAPP-A level, fβhCG, nuchal translucency, and pregnancy outcome data were available N excluded from analysis = 0 (whilst no enrolled women were excluded from the analysis, only women with complete outcome data from thr FaSTER trial were included in the study) N included in analysis = 33,395 Demographics (entire study population) Maternal age at delivery, years, mean (SD, range) ○ 30.1 (5.77, 16 to 53) Parity, mean (SD) ○ 0.9 (1.07) Nulliparity, n (%) ○ 15,077 (45.2) Race (%) ○ White (67%), Hispanic (23%), Black (5%), Asian/Pacific Islanders (4%), Other (1%) Comorbidities and pregnancy complications Gestational hypertension, n (%) ○ 1,484 (4.47) Gestational diabetes, n (%) ○ 1,012 (3.05) Pre-eclampsia, n (%) ○ 764 (2.30) Birth weight at <10th percentile, n (%) ○ 2,994 (9.20) Birth weight at ≤5th percentile, n (%) ○ 1,300 (3.99) Preterm premature rupture of the membranes, n (%) ○ 521 (1.57)

Study reference	**Dugoff 2004**
	Placenta previa, n (%) ○ 191 (0.58) Placenta abruption, n (%) ○ 229 (0.69) Macrosomia, n (%) ○ 397 (1.21)
Screening Method	Index test Various levels of PAPP-A for fetal death >24 weeks of gestation For each pairwise combination of outcome and marker, a receiver operating characteristic curve analysis was performed to examine the tradeoff between sensitivity and specificity at every cutoff at the given marker, to determine the most appropriate thresholds. Values were classified as abnormally low if they were at ≤5th percentile and as abnormally high if they were at ≥95th percentile Reference standard: Intrauterine fetal death was defined as fetal death >24 weeks of gestation. The clinical coordinators or physicians obtained pregnancy outcome data by medical record review and/or patient interviews, with an extensive follow-up protocol and predefined criteria. Any abnormal outcome that was reported by a patient was followed up by a formal medical record review

Test Accuracy

PAPP-A levels for intrauterine fetal death >24 weeks of gestation

PAPP-A Level	Sensitivity	False positive rate	Positive predictive value	Negative predictive value
PAPP-A ≤10th percentile	15.79	10.09	0.44	99.73
PAPP-A ≤5th percentile	10.53	5.19	0.58	99.73
PAPP-A ≤1st percentile	3.16	1.06	0.84	99.72

Authors' Conclusions	Some evidence was found that low PAPP-A levels are associated with intrauterine fetal death at >24 weeks of gestation.

Study reference	**Dugoff 2005 (this study reports on the same study population as Dugoff 2004**
Study Characteristics	Design Prospective cohort study Objective To estimate the predictive relationship between second-trimester levels of maternal serum AFP, hCG, unconjugated estriol (uE_3), and inhibin A on obstetrical complications Dates 1999 to 2002 Country USA

Study reference	**Dugoff 2005 (this study reports on the same study population as Dugoff 2004**
	Setting NR
Population Characteristics	Patient recruitment Women who participated in the study were enrolled in the First and Second Trimester Evaluation of Risk (FASTER) trial, which was a multicentre study to compare the diagnostic performance of several first and second trimester screening markers for Down syndrome. Women aged ≥16 confirmed to have a singleton gestation between 10 3/7 and 13 6/7 weeks of gestation, as defined by the Hadlock criteria, were eligible to be included in the FASTER trial. Women who were confirmed to have a viable singleton pregnancy (with fetal crown-rump length between 36 and 79 mm inclusive) were eligible to participate. Women whose fetuses were diagnosed with anencephaly or a septated cystic hygroma at the initial ultrasound examination were excluded. Data collection Relevant patient history, demographic data, and obstetric history were collected at the time of enrolment in the FASTER trial. After delivery, the clinical coordinators or physicians at each FASTER site obtained pregnancy outcome data by medical record review and/or patient interviews, with an extensive follow-up protocol and predefined criteria. Any abnormal outcome that was reported by a patient was followed up by a formal medical record review. The outcomes examined included spontaneous loss at ≤24 weeks of gestation, fetal loss at >24 weeks of gestation, neonatal loss, preterm delivery, gestational hypertension, preeclampsia, low birth weight, preterm premature rupture of membranes, placenta previa, placental abruption, gestational diabetes mellitus, and macrosomnia. Definition of stillbirth [or SGA<3rd centile] Fetal loss at >24 weeks of gestation Prevalence of stillbirth in the study 0.28% (95 cases of intrauterine fetal death at >24 weeks of gestation) Sample size N screened/invited = NR N eligible = NR N enrolled = NR N excluded (with reason) = Women whose fetus had a chromosomal (n=221) or structural abnormality (n=462) N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 33,145 Demographics (entire study population) **Maternal characteristics of this study population are also reported by the Dugoff 2004 study.** Maternal age at delivery, years, mean (SD, range) ○ 30.2 (5.71, 16 to 53) Nulliparity, n (%) ○ 15,127 (45.7) Race, % ○ White (68.7%), Hispanic (21.7%), African American (4.7%), Asian/Pacific Islanders (4.0%), Other (0.9%)
Screening Method	Index test Maternal serum AFP, hCG, uE_3, and inhibin A for fetal death >24 weeks of gestation Blood samples for measurement of maternal serum AFP, hCG, Ue_3, and inhibin A were obtained between 15 and 18 6/7 weeks of gestation. For each pairwise combination of outcome and marker, a receiver operating characteristic analysis was performed to examine the trade-off between sensitivity and specificity at every cutoff at the given marker, to determine the most appropriate thresholds. Values were classified as abnormally low if they were at or below 0.5 MoM and as abnormally high if they were at or above 2.0 MoMs (specifically analysed the effects of AFP, hCG, and inhibin A at or above 2.0 MoMs, and the effect of uE_3, at or below 0.5 MoMs. Reference standard Unclear. Complete pregnancy and pediatric outcome data were available on all participants, and a perinataologist and paediatric geneticist reviewed the maternal and paediatric medical records of all subjects who had abnormal first- or second- trimester screening or adverse paediatric outcome.

Study reference	**Dugoff 2005 (this study reports on the same study population as Dugoff 2004**
Test Accuracy	Multiple abnormal markers (≥2) versus single or no abnormal markers (<2) for fetal femise >24 weeks (N=33,145) Sensitivity: 13.4 (95% CI 6.62 to 20.18) Specificity: 96.83 (95% CI 96.64 to 97.02) PPV: 1.23 (95% CI 0.57 to 1.89) NPV: 99.74 (95% CI 99.68 to 99.8) LR+: 4.23 (95% CI 4.01 to 4.45) LR−: 0.89 (95% CI 0.79 to 0.99)
Authors' Conclusions	These data suggest that components of the screening test may prove useful in predicting adverse obstetric outcomes.

Study reference	**Dugoff 2008**
Study Characteristics	Design Prospective cohort study Objective To develop and evaluate a method of estimating patient-specific risk for fetal loss by combining maternal characteristics with serum markers Dates October 1999 to December 2002 Country USA Setting 15 US centres
Population Characteristics	Patient recruitment The women analysed in this study were enrolled in the FaSTER Trial (First- and Second-Trimester Evaluation of Risk), a multicentre, prospective study on combined ultrasound and biochemical screening for Down syndrome. At the time of enrolment, all women had viable singleton pregnancies with fetal crown-rump length between 36 mm and 79 mm inclusive (corresponding to 10 3/7 and 13 6/7 weeks' GA by Hadlock criteria). Data collection Relevant patient history, demographic data, and obstetric history were collected at the time of enrolment in the FaSTER trial. This information was entered on the study data collection forms and sent to the centralised data coordinating centre Definition of stillbirth [or SGA<3rd centile] Fetal loss occurring at 24 weeks or later Prevalence of stillbirth in the study There were 103 cases of fetal loss at 24 weeks or later Sample size N screened/invited = NR N eligible = NR N enrolled = 36,014 (all women with pregnancy and outcome data available) N excluded (with reason) = chromosomal abnormality (n=219), structural abnormality (n=467), elective termination of pregnancy (n=75) N lost to follow-up = NR N completed = NR N excluded from analysis = 0 N included in analysis = 35,253 Demographics NR
Screening Method	Index test The quad screen markers (AFP, total hCG, uE_3 [unconjugated estriol] and inhibin A) were measured between 15

Study reference	**Dugoff 2008**
	0/7 and 18 6/7 weeks' gestation. Reference standard After delivery the clinical coordinators or physicians at each FaSTER site obtained pregnancy outcome data by medical record review and/or patient interviews using an extensive follow-up protocol and predefined criteria. Any abnormal outcome reported by a patient was followed up by a formal medical record review
Test Accuracy	**Inhibin A** 8% sensitivity for a false positive rate of 1% (specificity 99%) 17% sensitivity for a false positive rate of 5% (specificity 95%) 27% sensitivity for a false positive rate of 10% (specificity 90%) The mean multiples of the mean (MoM) was estimated by the observed median and the standard deviation from the 10th to the 90th percentile range divided by 2.563. The extent of separation between MoM for fetal losses and controls was 0.23
Authors' Conclusions	Patient-specific risk assessment for late fetal loss has low detection rates.

Study reference	**Familiari 2016**
Study Characteristics	Design Retrospective cohort study Objective To investigate the potential value of screening integrating mid-pregnancy maternal demographics, fetal biometry and UtA pulsatility index (PI) in the prediction of IUD and to examine the potential value of such assessment in identifying women who may benefit from increased antenatal surveillance. Dates 2000 to 2014 Country UK Setting A single maternity centre
Population Characteristics	Patient recruitment Morphologically normal singleton pregnancy women attending their routine, scheduled ultrasound examination between 19 and 24 weeks of gestation. Pregnancies complicated by fetal abnormality, maternal medical disorders, previous adverse obstetric outcome, aneuploidy or infection were excluded from the analysis, as were all women who delivered elsewhere or were referred from other hospitals. Only those classified as low-risk after first trimester scan were included Data collection Data on ultrasound examinations were obtained from computerised records and only the anomaly scan examination (1/fetus) was included in the analysis. Qualified and experienced sonographers performed all scans. Definition of stillbirth [or SGA<3rd centile] Stillbirth was defined as the death of a fetus with a birthweight ≥500 g or gestation age >23 (+6) weeks of gestation Prevalence of stillbirth in the study There were 90 stillbirths out of 23,894 pregnancies (*0.38%).* There were 38 stillbirths at term and 52 were preterm Sample size N screened/invited = NR

Study reference	**Familiari 2016**
	N eligible = NR N enrolled = 23,894 N excluded (with reason) = NR N lost to follow-up = NR N completed = 23,894 N excluded from analysis = 0 N included in analysis = 23,894 Demographics Maternal age (years), mean (SD): o Live births: 29.99 (5.6) o Stillbirths: 30.6 (6.6) Gestational age at delivery (weeks) o Live births: 39.7 (2) o Stillbirths: 35.4 (4) BMI, median (IQR) o Live births: 24.55 (4.8) o Stillbirths: 25.24 (5.1) Ethnicity, n (%): o Live births: Caucasian 13,401 (56.3); Asian 4942 (20.8); Afro-Caribbean 3501 (14.7); Mixed-others 1960 (8.2) o Stillbirths: Caucasian 51 (56.8); Asian 13 (14.4); Afro-Caribbean 13 (14.4); Mixed-others 13 (14.4) Small for gestational age, n (%) o Live births: 1,563 (6.6) o Stillbirths: 52 (57.8) Parity/gravida: NR Smoking status: NR Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test Maternal characteristics, second trimester fetal biometry (femur length) and Doppler measurements (UtA-PI) for stillbirth Measurements were taken with pulsed-wave Doppler at the lowest insonation angle achievable and when uniform waveforms with high signal-to-noise ratio were obtained, the PI was measured. Ultrasound data, including head circumference, abdominal circumference and femur length (FL), were measured according to a standard protocol and were converted to percentiles using reference values derived from low-risk pregnancies with documented normal outcome. The risk for each of the pregnancy outcomes was calculated. The performance of screening was determined by receiver operating characteristic curves. Reference standard Unclear
Test Accuracy	**All stillbirths** Sensitivity for maternal factors 19% at a specificity (1-false positive rate) of 90% Sensitivity for femur length centile 26% at a specificity (1-false positive rate) of 90% Sensitivity for Uterine artery Doppler 28% at a specificity (1-false positive rate) of 90% Sensitivity for maternal factors + femur length centile + uterine artery Doppler 31% at a specificity (1-false positive rate) of 90% Area under the operator curve graphs are available to estimate test parameters at other levels of specificity.

Study reference	**Familiari 2016**
	Term stillbirths Sensitivity for maternal factors 12% at a specificity (1-false positive rate) of 90% Sensitivity for femur length centile 27% at a specificity (1-false positive rate) of 90% Sensitivity for Uterine artery Doppler 24% at a specificity (1-false positive rate) of 90% Sensitivity for maternal factors + femur length centile + uterine artery Doppler 27% at a specificity (1-false positive rate) of 90% **Preterm births** Sensitivity for maternal factors 14% at a specificity (1-false positive rate) of 90% Sensitivity for femur length centile 23% at a specificity (1-false positive rate) of 90% Sensitivity for Uterine artery Doppler 31% at a specificity (1-false positive rate) of 90% Sensitivity for maternal factors + femur length centile + uterine artery Doppler 35% at a specificity (1-false positive rate) of 90% Area under the operator curve graphs are available to estimate test parameters at other levels of specificity.
Authors' Conclusions	The findings of the study demonstrate that second trimester assessment does not have a high sensitivity in detecting term stillbirths (consistent with the literature). The performance improved to a sensitivity of 35% for a 10% FPR when examining the rate of preterm stillbirths, which show a strong association with severe intrauterine growth restriction, explaining why the UtA Doppler indices are the best single predictor. The study also shows that the fetal biometry parameters, only the FL centile was an independent predictor of IUD in the study population of low-risk woman. The performance of screening was only marginally improved by the addition of FL to the UtA Doppler assessment.

Study reference	**Hemming 2011**
Study Characteristics	Design Retrospective cohort study Objective To discuss different methods for evaluating fetal growth and population-based birthweight standards relevant to different uses: either in antenatal care or in epidemiology Dates 1980 to 2003 Country UK (Scotland) Setting Routinely collected data in Scotland

Study reference	**Hemming 2011**
Population Characteristics	Patient recruitment All singleton pregnancies ≥24 weeks between 1980 and 2003 were included Data collection Data was routinely collected as part of the antenatal care pathway Definition of stillbirth [or SGA<3rd centile] Unclear, but study only includes fetuses ≥24 weeks is conducted in the UK and specifically refers to stillbirths, so assume UK definition. SGA was defined as <10th percentile. Prevalence of stillbirth in the study 2,702 stillbirths out of 540,849 deliveries Sample size N screened/invited = NR N eligible = 540,849 N enrolled = 540,849 N excluded (with reason) = 0 N lost to follow-up = NR N completed = NR N excluded from analysis = 0 N included in analysis = 540,849 Demographics Gestational age at birth, n (%): ○ Term: 507,716 (93.9) ○ 34 to 36 weeks: 22,818 (4.2) ○ 32 to 33 weeks: 4,491 (0.8) ○ 24 to 31 weeks: 5,824 (1.1) Parity/gravida: NR Smoking status: NR Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test **Fetal growth standard** and a **population-based birthweight** standard were compared for the accuracy of estimated risks of stillbirth. For the fetal growth chart or birthweight standard, sensitivity is the percentage of stillborn infants who were classified as SGA by the standard. Specificity is the percentage of liveborn not classified as SGA. Reference standard Unclear
Test Accuracy	**Fetal growth standard**

	Sensitivity	Specificity	PPV	NPV	LR+
All	0.43	0.90	2.11	99.68	4.29
Term	0.30	0.91	0.57	99.86	3.17
34–36 weeks	0.34	0.90	0.88	99.81	3.49
32–33 weeks	0.36	0.90	1.09	99.79	3.67
24–31 weeks	0.43	0.90	2.11	99.68	4.29

Birthweight standard

	Sensitivity	Specificity	PPV	NPV	LR+
All	0.29	0.90	1.45	99.61	2.93

Study reference	**Hemming 2011**

Term	0.29	0.90	0.52	99.86	2.93
34–36 weeks	0.30	0.90	0.76	99.80	3.00
32–33 weeks	0.30	0.90	0.90	99.77	3.02
24–31 weeks	0.29	0.90	1.45	99.61	2.94

The authors stated that fetal growth standard had better sensitivity, but they had no preference between tests based on specificity.

Authors' Conclusions	In clinical care, the the evidence points towards using fetal growth standards: sensitivity at term is about 30%, increasing to 43% for preterm births (24–31 weeks of gestation), compared with about 29% across all ages under the birthweight standard. The results indicate that for screening purposes during pregnancy, the fetal growth standard performs better than the birthweight standard. Despite this, neither standard performs particularly well: at early gestations it is necessary to screen around 500 pregnancies (rising to 2000 pregnancies at term) to identify one fetus that will be classified as SGA and will have a poor perinatal outcome. Importantly, identifying one fetus at risk does not mean that the life will be saved.

Study reference	**Marttala 2010**
Study Characteristics	Design Retrospective population-based register study Objective To determine whether low maternal serum PAPP-A levels in the first trimester are associated with SGA and stillbirth Dates 1st January 2005 to 31st December 2008 Country Finland Setting Oulu University Hospital
Population Characteristics	Patient recruitment Low-risk women with singleton pregnancies attending their first trimester combined Down's syndrome screening during the study period who volunteered to participate. Data collection Serum samples were analysed at the accredited laboratory at Oulu University Hospital. The data pertaining to all SGA cases, including all stillbirths, as well as data for smoking habits, were extracted from the Finnish Birth Register and Hospital Discharge Register of the National Institute for Health and Welfare Definition of stillbirth [or SGA<3rd centile] Stillbirth was defined as fetal death during or after the 22nd gestational week or birth weight under 500g SGA was defined as birthweight <2 SD related to gestational age on the basis of the national sex-specific standards Prevalence of stillbirth in the study There were 9 stillbirths in the study group (1.0%) and 51 stillbirths (0.3%) in the control group Sample size N screened/invited = 19,536 N eligible = 19,536 N enrolled = 19,536 N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 19,536 N excluded from analysis = 0 N included in analysis = 19,536 (921 women with low PAPP-A [study group] and 18,615 control women)

Study reference	**Marttala 2010**
	Demographics Maternal age, mean years o Study group: 30.3 o Control group: 29.2 Smoking during pregnancy, n (%): o Study group: 202 (21.9) o Control group: 2,923 (15.7) Parity/gravida: NR Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test **PAPP-A to predict SGA and stillbirth** PAPP-A was measured from the serum of women attending the Down's syndrome screening between 9+0 and 13+6 weeks' gestation. PAPP-A was given as MoMs with corrections for maternal weight, diabetic status and smoking The lowest 5% of PAPP-A corresponded to <0.3 MoM. The study groups consisted of women with <0.3 MoM PAPP-A and the control group were women with ≥0.3 MoM PAPP-A levels Reference standard Unclear
Test Accuracy	Sensitivity 15.0% Specificity 95.3% PPV 1.0% NPV 99.7%
Authors' Conclusions	The study showed significant linkage between low maternal first trimester PAPP-A serum levels and SGA newborns and stillbirths. A low PAPP-A level can be seen as a marker for SGA newborns and stillbirth. In this study, the lowest 5% PAPP-A level corresponded to 0.3 MoM.

Study reference	**Mastrodima 2016 (same cohort as Akolekar 2016, Akolekar 2016b, Aupont 2016 with crossover with Bakalis 2011)**
Study Characteristics	Design Prospective cohort study Objective To develop a model for prediction of stillbirth based on a combination of maternal characteristics and medical history with first trimester biochemical and biophysical markers and evaluate the performance of screening of this model for all stillbirths and those due to impaired placentation and unexplained causes Dates March 2006 to October 2015 Country UK Setting King's College Hospital and Medway Maritime Hospital

Study reference	**Mastrodima 2016 (same cohort as Akolekar 2016, Akolekar 2016b, Aupont 2016 with crossover with Bakalis 2011)**
Population Characteristics	Patient recruitment Data for this study were derived from prospective screening for adverse obstetric outcomes in women attending for routine pregnancy care at 11+0 to 13+6 weeks' gestation Data collection Data on pregnancy outcome were obtained from the maternity hospital records or the general practitioners of women. The hospital maternity records of all women with antepartum stillbirths were reviewed to determine if the death was associated with preeclampsia, abruption or the birthweight was <10th percentile for gestational age or it was unexplained Definition of stillbirth [or SGA<3rd centile] Stillbirth: Unclear, but study only included pregnancies that delivered a phenotypically normal live birth or stillbirth at ⩾24 weeks' gestation Prevalence of stillbirth in the study 268 antepartum stillbirths out of 76,897 singleton pregnancies, with 157 (59%) secondary to impaired placentation and 111 (41%) due to other or unexplained causes. Sample size N screened/invited = NR N eligible =NR N enrolled = 76,897 N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = 0 N included in analysis = 76,897 Demographics • Maternal age, median (IQR) o Live births: 31.3 (26.7 to 35.1) o Stillbirths: 31.6 (26.4 to 35.6) • Chronic hypertension, n (%) o Live births 1,075 (1.4) o Stillbirths 18 (6.7) • Systemic lupus erythematosus/ antiphospholipid syndrome, n (%) o Live births 154 (0.2) o Stillbirths 4 (1.5) • Diabetes mellitus, n (%) o Live births 695 (0.9) o Stillbirths 10 (3.7) • Smoking during pregnancy o Live births: 7,125 (9.3) o Stillbirths: 31 (11.6)
Screening Method	Index test PAPP-A, UT-PI and DV-PIV for prediction of stillbirth Maternal characteristics and medical history were recorded and combined screening to estimate risks for fetal aneuploidies based on maternal age, fetal nuchal translucency (NT) thickness and measurement of maternal serum pregnancy associated plasma protein-A (PAPP-A) and free ß- hCG was performed. Transabdominal colour Doppler ultrasound was performed to measure ductus venosus pulsatility index for veins (DV-PIV) and uterine artery pulsatility index (UT-PI) Reference standard Unclear
Test Accuracy	**All stillbirths** Sensitivity for maternal factors 19.7% (15.0 to 24.6) for a specificity (1-false positive rate) of 95%

Study reference	Mastrodima 2016 (same cohort as Akolekar 2016, Akolekar 2016b, Aupont 2016 with crossover with Bakalis 2011)
	31.3% (25.6 to 36.9) for a specificity (1-false positive rate) of 90% Sensitivity for maternal factors + PAPP-A + UT-PI + DV PIV 32.5% (26.9 to 38.1) for a specificity (1-false positive rate) of 95% 39.9% (34.0 to 45.8) for a specificity (1-false positive rate) of 90% **Stillbirths from impaired placentation** Sensitivity for maternal factors 23.6% (17.0 to 30.2) for a specificity (1-false positive rate) of 95% 36.3% (28.8 to 43.8) for a specificity (1-false positive rate) of 90% Sensitivity for maternal factors + PAPP-A + UT-PI + DV PIV 47.8% (40.0 to 55.6) for a specificity (1-false positive rate) of 95% 54.8% (47.0 to 62.6) for a specificity (1-false positive rate) of 90%
Authors' Conclusions	The main findings of the study demonstrate that more than half of stillbirths that are due to impaired placentation can be predicted in the first trimester of pregnancy by a combination of maternal factors and biomarkers. The performance of screening is better for stillbirths that are secondary to impaired placentation than for those that are unexplained. Among stillbirths due to impaired placentation, the DR is higher for those that occur at term.

Study reference	Odibo 2012
Study Characteristics	Design Retrospective cohort study Objective To estimate the impact of adding ultrasound biometric parameters to the customised chart to refine the prediction of intrauterine fetal death in fetuses classified as small for gestational age. Dates 1990 to 2009 Country USA Setting Washington University School of Medicine (a major referral institution)
Population Characteristics	Patient recruitment The study population was derived from a prospectively collected ultrasound and genetics database of unselected singleton pregnancies seen between 16 and 20 weeks' gestation. Pregnancies with congenital anomalies, spontaneous losses or termination prior to 20 weeks, multifetal pregnancies and stillbirths were excluded, only from derivation data. The eligible births were split into derivation (34,832) and validation (24,184) samples. Data collection All pregnancies were followed prior to or at 20 weeks' gestation and gestational age was assigned using ultrasound criteria if discrepant with ultrasound age. Each patient seen in the prenatal diagnosis centre was given a standardised form requesting pregnancy outcome to be returned following delivery. When a form was not returned within 4 weeks of the expected date of delivery, the patient received a phone call from the coordinator. In cases where the patient cannot be contacted, her referring physician was contacted for the outcome information. For patients delivering in the study healthcare system, outcome data were extracted from the perinatal computerised database. Definition of stillbirth [or SGA<3rd centile] Unclear, but IUFD was defined as fetal death after 20 weeks' gestation. SGA was defined as the 10th weight-for-gestational-age centile limit Prevalence of stillbirth in the study

Study reference	**Odibo 2012**
	Of the 24,184 pregnancies in the validation sample, IUFD was noted in 169 (0.7%). Sample size N screened/invited = NR N eligible = NR N enrolled = 59,016 (derivation sample = 34,832, validation sample = 24,184) N excluded (with reason) = NR N lost to follow-up = NR N completed = 59,016 (derivation sample = 34,832, validation sample = 24,184) N excluded from analysis = 0 N included in analysis = 59,016 (derivation sample = 34,832, validation sample = 24,184) Demographics • Maternal age, mean (SD) o Derivation population: 29.2 (6.4) o Validation population: 30.6 (6.4) • BMI (kg/m^2), mean (SD) o Derivation population: 25.8 (6.1) o Validation population: 25.7 (6.2) o Parity, n (%)Para 0 (Derivation population): 11,842 (34.0) o Para 0 (Validation population): 7,836 (32.4) o Para 1 (Derivation population): 11,878 (34.1) o Para 1 (Validation population): 8,634 (35.7) o Para 2 (Derivation population): 6,514 (18.7) o Para 2 (Validation population): 4,764 (19.7) o Para 3 (Derivation population): 2,996 (8.6) o Para 3 (Validation population): 1,838 (7.6) o Para ≥4 (Derivation population): 1,602 (4.6) o Para ≥4 (Validation population): 1,112 (4.6) • Smoking, n (%) o Derivation population: 3,553 (10.2) o Validation population: 3,313 (13.7) • Pre-gestational diabetes, n (%) o Derivation population: 522 (1.5) o Validation population: 339 (1.4) • Hypertensive disorders of pregnancy, n (%) o Derivation population: 2,995 (8.6) o Validation population: 1,620 (6.7) • Placental abruption, n (%) o Derivation population: 209 (0.6) o Validation population: 121 (0.6)
Screening Method	Index test Screening for IUFD associated with SGA defined using customised charts on adjusted physiologic variable only (Cust-chart) or with addition of ultrasound biometry (Cust-plus-USS-chart), compared with population-based growth chart (Pop-chart) Pregnancies in the validation model were used to estimate the association between SGA and IUFD. The sensitivity, specificity, and positive and negative predictive values for the association between SGA defined by each growth chart and IUFD were calculated. Reference standard Unclear. SGA was defined by the customised and population-based charts, and outcome data were collected from standardised forms completed by pregnant women following delivery, or was obtained from the referring physician or perinatal computerised database.
Test Accuracy	**Customised chart** Sensitivity (95% CI) 55.5 (52.6 to 58.6) Specificity (95% CI) 90.4 (88.4 to 92.1)

Study reference	Odibo 2012
	PPV (95% CI) 3.8 (3.1 to 4.6) NPV (95% CI) 99.6 (99.5 to 99.7) **Cust-plus-USS chart** Sensitivity (95% CI) 54.9 (51.7 to 58.0) Specificity (95% CI) 90.3 (88.2 to 92.1) PPV (95% CI) 3.5 (2.8 to 4.3) NPV (95% CI) 99.6 (99.5 to 99.7) **Population-based growth chart** Sensitivity (95% CI) 19.5 (15.5 to 24.4) Specificity (95% CI) 89.2 (88.8 to 89.6) PPV (95% CI) 1.8 (1.4 to 2.4) NPV (95% CI) 99.1 (98.9 to 99.2)
Authors' Conclusions	In this study, we found that the cases identified as SGA by the Cust-Chart had an increased risk of IUFD compared with those identified by our Pop-chart. The addition of second-trimester ultrasound biometry did not improve the identification of pregnancies at risk for IUFD.

Study reference	Poon 2013
Study Characteristics	Design Retrospective cohort (pooled data from a retrospective analysis of an RCT and a prospective screening study) Objective To examine the role of second-trimester uterine artery Doppler in the prediction of stillbirths Dates 1999 to 2002 and 2006 to 2011 Country UK Setting 9 hospitals in and around London
Population Characteristics	Patient recruitment Women with singleton pregnancies undergoing routine ultrasound examination at 20 to 24 weeks' gestation Data collection The first (data set A) consisted of 30,566 pregnancies examined between 1999 and 2002 at 7 hospitals in and around London. The second (data set B) consisted of 35,253 pregnancies in which the scan at 20 to 24 weeks was preceded by combined screening for aneuploidies at 11 to 13 weeks' gestation between 2006 and 2011 at 3

Study reference	**Poon 2013**
	hospitals in and around London. Data on pregnancy outcomes were collected from the hospital maternity records or from general medical practitioners Definition of stillbirth [or SGA<3rd centile] Dead infant delivered at or after 24 weeks' gestation. Stillbirths were divided into three groups: antepartum, intrapartum and placental abruption Prevalence of stillbirth in the study In a total population of 65,819 singleton pregnancies, there were 65,513 live births and 306 stillbirths. Sample size N screened/invited = NR N eligible = NR N enrolled = NR N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = NR N included in analysis = 65,819 Demographics Maternal age, median (IQR): o Livebirths 30.4 (26.1 to 34.3) years o Stillbirths 30.9 (26.0 to 34.9) years Smoking, n (%) o Live births: 7,834 (12.0) o Stillbirths: 41 (13.4) Parity, n (%) Nulliparous o Livebirths 34,035 (52.0) o Stillbirths 153 (50.0) Parous o Livebirths 31,478 (58.0) o Stillbirths 153 (50.0) GA at birth – reported on graph Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test Uterine artery PI was measured by transvaginal Doppler sonography. The sonographers who performed the Doppler studies had received the Certificate of Competence in Doppler of the Fetal Medicine Foundation Reference standard Antepartum stillbirth was defined as fetal death before the onset of labour and in such cases the diagnosis was essentially made by ultrasonography in women presenting with reduced or absent fetal movements. Intrapartum stillbirth was defined as fetal death after the onset of labour and before birth and in these cases there was ultrasonographic or cardiotocographic evidence that the fetus was alive at the onset of labour. Placental abruption was defined retrospectively by the presence of a retroplacental clot at the delivery of a stillbirth that was either alive or dead at presentation with abdominal pain either with or without vaginal bleeding
Test Accuracy	Sensitivity for UtA-PI >90th percentile MoM 39.9% (122/306)
Authors' Conclusions	The measurement of uterine artery PI at 20–24 weeks can identify a high proportion of subsequent stillbirths. However, in most of these cases the stillbirth may not be avoidable because it is associated with severe early-onset fetal growth restriction and it should, therefore, be aimed at identifying high-risk pregnancies in the first trimester when therapeutic interventions may improve placentation.

Study reference	**Singh 2012**
Study Characteristics	Design Retrospective cohort study Objective To estimate the association between uterine artery Doppler indices and stillbirth in routinely screened populations Dates 2000 to 2008 Country UK Setting Single obstetric unit serving an inner city multi-ethnic population
Population Characteristics	Patient recruitment Women with singleton pregnancies who were nulliparous or had a previous pregnancy with a history of placental syndromes (fetal growth restriction, preeclampsia, or stillbirth) Data collection Pregnancy outcomes were obtained from the hospital computerised maternity database, the general practitioner, or the patient. Data for all stillbirths were obtained from the hospital national reporting register, even for women who delivered in other health districts. Cases of congenital abnormalities and unknown pregnancy outcome because of out-of-date contact details, as well as stillbirth cases with unknown gestation of death were excluded from the analysis Definition of stillbirth [or SGA<3rd centile] None given, but assume UK definition and antepartum stillbirth Prevalence of stillbirth in the study 144 antepartum stillbirths out of 15,796 pregnancies Sample size N screened/invited = 15,835 N eligible =15,796 N enrolled =15,796 N excluded (with reason) = NR N lost to follow-up = NR N completed = NR N excluded from analysis = 10 N included in analysis = 15,786 Demographics Age: NR Parity/gravida: NR Smoking status: NR BMI, median (IQR): o Live births: 22.9 (20.9 to 25.9) kg/m^2 o Stillbirths: 24.8 (21.9 to 28.5) kg/m^2 Smoking o Live births: 11.4% o Stillbirths 15.3% GA at birth, median (IQR) o Live births: 40+0 (38+6 to 41+0) weeks o Stillbirths: 28+3 (24+2 to 37+5) weeks Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test Uterine artery Doppler-RI >90th percentile Doppler assessment was performed at the time of routine anomaly scan between 19 and 23 weeks of gestation. The flow velocity waveform of the uterine arteries was identified as previously described and the resistance index

Study reference	**Singh 2012**
	(RI) recorded on Viewpoint, a dedicated ultrasound database Reference standard Gestation of birth was used as a proxy for the gestation of stillbirth. The latter is based on an average 2-day time interval between fetal death and delivery in the third trimester. When the degree of tissue autolysis indicated a protracted interval between fetal demise and birth, the gestational age at death was classified as unknown
Test Accuracy	Uterine artery Doppler >90th percentile Sensitivity 46.2% PPV 0.46% NPV 95.73% Sensitivity of uterine artery Doppler >95th percentile 35.4% Sensitivity of uterine artery Doppler >99th percentile 15.4%
Authors' Conclusions	Elevated second-trimester Doppler indices, a proxy for impaired placentation, are more strongly associated with stillbirth than conventional risk factors. Risk factors such as ethnicity, maternal age, BMI, and smoking contribute to risk of term stillbirth through uteroplacental dysfunction.

Study reference	**Smith 2007a**
Study Characteristics	Design Retrospective cohort study Objective To assess the relationship between maternal characteristics and serum screening data and the risk of stillbirth Dates 1992 to 2001 Country UK Setting Scottish Morbidity Record (collects data for all women discharged from Scottish maternity hospitals), the Scottish Stillbirth and Infant Death Enquiry national register, and a database for the West of Scotland antenatal screening programme in the Institute of Medical Genetics in Glasgow
Population Characteristics	Patient recruitment A probability-based matching approach was employed using maternal identifiers to link the Scottish Morbidity Record, the Scottish Stillbirth and Infant Death Enquiry and the antenatal screening database in the Institute of Medical Genetics. Multiple births and births outside the range of 24 to 43 weeks of gestation were excluded. Data collection Maternal age, parity, postcode of residence and all outcome data were obtained solely from the Scottish Morbidity record. Definition of stillbirth [or SGA<3rd centile] Unclear, but the study only included pregnancies between 24 and 43 weeks of gestation (in line with the UK definition of stillbirth). Prevalence of stillbirth in the study Of 84,769 pregnancies, there were 406 antepartum stillbirths Sample size N screened/invited = 216,563

Study reference

Smith 2007a

N eligible = 97,264
N enrolled = 97,062
N excluded (with reason) = 202 (69 deliveries were outside 24 to 43 gestational weeks and 133 deaths caused by fetal abnormality or rhesus isoimmunisation)
N lost to follow-up = 12,293 (missing data)
N completed = 84,769
N excluded from analysis = 0
N included in analysis = 84,769

Demographics
Maternal age, median (IQR):
- No antepartum stillbirth: 26 (22 to 30)
- Antepartum stillbirth: 27 (22 to 31)

BMI, median (IQR):
- No antepartum stillbirth: 23.5 (21.4 to 26.3)
- Antepartum stillbirth: 24.0 (21.8 to 27.9)

Parity/gravida: NR
Smoking status, n (%):
- No antepartum stillbirth: Never: 52,202 (61.9); Current: 24,258 (28.8); Former: 7,903 (9.4)
- Antepartum stillbirth: Never: 185 (45.6); Current: 185 (45.6); Former: 36 (8.9)

Comorbidities: NR
Previous pregnancy complications: NR

Screening Method

Index test
Maternal characteristics and maternal serum levels of AFP and hCG to predict stillbirth

Biochemical data were categorised into quintiles, with the last quintile split into the top 5, 6–10 and 11–20%. Univariate and multivariate Cox regression were performed using gestational age as the timescale, antepartum stillbirth as the event, and all other births as censored. The performance of the model in each gestational window was assessed via the area under the ROC curve, and measures of test accuracy were estimated using different thresholds of predicted risk as screen positive, specifically, the top 5%, 10% and 20%.

Reference standard
Stillbirths were defined as antepartum or intrapartum, and the cause of stillbirth was classified according to direct obstetric causes: congenital abnormality, pre-eclampsia, haemorrhage (antepartum), mechanical, maternal, miscellaneous and unexplained. Classification was performed by a single, medically-qualified individual, with the results of the post-mortem investigations, where obtained.

Test Accuracy

Gestational age and predicted risk	PPV (%)	Sensitivity (%)	Specificity (%)	LR+	LR−
24 to 28 weeks					
Top 5%	0.97	36.68	95.04	7.80	0.65
Top 10%	0.52	41.51	90.04	4.17	0.65
Top 20%	0.34	54.72	80.04	2.74	0.57
29–32 weeks					
Top 5%	0.33	18.67	95.01	3.74	0.86
Top 10%	0.28	32.00	90.02	3.21	0.76
Top 20%	0.20	44.00	80.02	2.20	0.70
33–36 weeks					
Top 5%	0.47	25.00	95.13	5.13	0.79
Top 10%	0.32	34.21	90.23	3.50	0.73
Top 20%	0.21	46.05	80.45	2.36	0.67
37–43 weeks					
Top 5%	0.63	16.78	95.02	3.38	0.88
Top 10%	0.42	22.15	90.03	2.22	0.86

Combination of maternal and biochemical data

Study reference	**Smith 2007a**					
	Top 20%	0.33	35.57	80.03	1.78	0.81
Authors' Conclusions	The main finding is that women with elevated serum levels of α-FP and hCG in the second trimester of pregnancy were at increased risk of stillbirth and that these associations were strongest for stillbirth at extreme preterm gestations. A model combining maternal characteristics and biochemical data performed reasonably well in predicting stillbirth risk at extreme preterm gestations (24-28 weeks). In contrast, a model combining maternal characteristics and biochemical data performed poorly at predicting stillbirth at term.					

Study reference	**Smith 2007b**
Study Characteristics	Design Prospective cohort study Objective The aims of the study were 1) to relate the risk of stillbirth to indices of Doppler flow velocimetry of the uterine arteries performed at 22–24 weeks of gestation in relation to presumed placental and nonplacental causes of stillbirth and the timing of stillbirth, and 2) to characterize the properties of uterine artery Doppler as a screening test for stillbirth in an unselected population. Dates October 1999 to August 2002 Country UK Setting Seven hospitals in and around London (Basildon Hospital; Greenwich Hospital; Harold Wood Hospital, Romford; King George Hospital, Ilford; King's College Hospital, London; Queen Mary's Hospital, Sidcup; University Hospital Lewisham, London).
Population Characteristics	Patient recruitment The data analysed were obtained as part of a multicentre study of screening and intervention in pregnancy involving seven hospitals in and around London. Women were recruited to the study had a scan at 22–24 weeks of gestation, which included assessment of cervical length and uterine artery Doppler flow velocimetry. The current analysis in this cohort excluded women recruited to the trial of cervical cerclage. Stillbirths due to congenital abnormalities were excluded. All analyses for the present study focussed on antepartum stillbirths. Data collection Maternal details and past medical and obstetric history were obtained using a questionnaire. Outcome was ascertained by computerized databases in each of the centres. Maternal history and Doppler findings were recorded in a computer database at the time of the Doppler studies in each participating centre. In all cases of adverse outcome (abruption, preterm birth, preeclampsia, and stillbirth), the clinical case record (and stillbirth autopsy, where performed) was reviewed by a medically qualified individual, the diagnosis confirmed, and further details were obtained, as necessary. Definition of stillbirth [or SGA<3rd centile] Stillbirths were defined as delivery of an infant which showed no signs of life. Stillbirth was defined as all intrauterine fetal deaths subsequent to the measurement of uterine artery Doppler. Stillbirths were divided into those where the fetus was thought to have died before the onset of labour (antepartum) and those where the fetus was thought to have been alive at the start of labour. The presumed cause of stillbirth was obtained from the case notes. Prevalence of stillbirth in the study There were 109 antepartum stillbirths in the cohort. There were 5 (2.0%) stillbirths among 255 women randomly assigned to aspirin and 4 (1.6%) stillbirths among 256 women assigned to placebo. Sample size N screened/invited = NR N eligible = 30,755 N enrolled = NR N excluded (with reason) = 256 with a short cervix or who had cervical cerclage (or both) N lost to follow-up = NR N completed = NR N excluded from analysis = Unclear

Study reference

Smith 2007b

N included in analysis = 30,519

Demographics

Maternal age, years, median (IQR)
- No stillbirth: 30 (26 to 34)
- Stillbirth: 31 (26 to 34)

Primigravidy, n (%)
- No stillbirth: 15,462 (50.8)
- Stillbirth: 52 (47.7)

Parous (uncomplicated), n (%):
- No stillbirth: 12,019 (39.5)
- Stillbirth: 36 (33.0)

Smoking status

Current smoker, n (%):
- No stillbirth: 4,513 (14.8)
- Stillbirth: 21 (19.3)

Nonsmoker or ex-smoker:
- No stillbirth: 25,897 (85.2)
- Stillbirth: 88 (80.7)

Comorbidities: NR

Previous pregnancy complications

Previous preterm birth
- No stillbirth: 2,291 (7.5)
- Stillbirth: 12 (11.0)

Previous IUFD and stillbirth
- No stillbirth: 638 (2.1)
- Stillbirth: 9 (8.3)

Screening Method

Index test

Doppler (mean pulsatility index, unilateral notch and bilateral notch), maternal characteristics (age, height, BMI, smoking, ethnicity, and past obstetric history).

Uterine artery Doppler studies were performed by using transvaginal ultrasonography by ultrasonographers trained in this method.

Each uterine artery was identified using colour flow mapping, and three similar consecutive waveforms were obtained using pulsed wave Doppler. The pulsatility index was measured, and the mean pulsatility index of the two uterine arteries was calculated. Quality control of screening, handling of data, and verification of adherence to protocols at the different centres was performed on a regular basis by the trial coordinators. Women with a mean pulsatility index greater than 1.6, which in an earlier study represented the 95th centile, were followed up with growth scans, blood pressure measurements, and urinalysis for protein at 28, 32, and 36 weeks. Women with normal uterine artery Doppler received routine antenatal care. For the purposes of dichotomizing uterine mean pulsatility index, those women with values in the top 10% were considered to have elevated mean pulsatility. For predictive models, variables were selected using backward stepwise logistic regression, with the threshold for removal being P=0.05. Logistic regression models were converted to tables of likelihood ratios.

Reference standard

Outcome was ascertained by computerised databases in each of the centers. In all cases of adverse outcome (abruption, preterm birth, preeclampsia, and stillbirth), the clinical case record (and stillbirth autopsy, where performed) was reviewed by a medically-qualified individual, the diagnosis confirmed, and further details were obtained, as necessary. The presumed cause of stillbirth was obtained from the case notes. Where the women suffered severe preeclampsia or had an abruption, these were assumed to have caused the stillbirth.

Test Accuracy

Outcome and Predictors	Top 5% Screen Positive*			
	Sensitivity	Specificity	LR+	LR-
All-cause stillbirth at ≤32 weeks				
Doppler† alone	58.5 (42.1–73.7)	95.2 (94.9–95.4)	12.1	0.44
Doppler† and maternal‡	53.7 (37.4–69.3)	95.1 (94.8–95.3)	10.9	0.49

Study reference	**Smith 2007b**

Maternal‡ alone	31.7 (18.1–48.1)	95.1 (94.8–95.3)	6.4	0.72
All cause stillbirth at ≥33 weeks				
Doppler† alone	6.6 (1.8–15.9)	95.1 (94.8–95.3)	1.3	0.98
Doppler† and maternal‡	21.3 (11.9–33.7)	95.0 (94.8–95.3)	4.3	0.83
Maternal‡ alone	18.0 (9.4–30.0)	95.1 (94.8–95.3)	3.7	0.86

LR, likelihood ratio. Data in parentheses are 95% confidence intervals.

*The sensitivity, specificity, and likelihood ratios were estimated when women in the top 5% of predicted risk were deemed to have screened positive.

† Doppler scores are mean pulsatility index (expressed as a continuous variable), unilateral notch and bilateral notch. In all models, mean pulsatility index was linear in the log odds scale on the basis of fractional polynomial analysis.

‡ Maternal characteristics in the model were age, height, body mass index, smoking, ethnicity, and past obstetric history.

	The study shows that a high resistance pattern of flow in the uterine artery is associated with an increased risk of stillbirth. The association is strongest for stillbirths due to placental dysfunction and, since these tend to occur at earlier gestations, is strongest for stillbirths occurring at extreme preterm gestations. Uterine artery Doppler is a relatively poor predictor of unexplained stillbirth unrelated to fetal growth restriction.

Study reference	**Smith 2014**
Study Characteristics	Design Retrospective case-control study Objective To compare the performance of three fetal growth curves in identifying abnormally grown fetuses at risk of stillbirth Dates 1st January 2000 to 31st December 2010 Country USA Setting Single institution
Population Characteristics	Patient recruitment Stillbirths (cases) delivered at a single institution were identified by death certificates and were included if an ultrasound examination had been performed within 1 month before delivery and if the fetuses were alive at that time and were beyond 24+0 weeks' gestation (because estimated weights were not reported at the authors' institution before 24+0 weeks). Four live-born controls for each case were randomly selected from birth logs of 85,968 singleton deliveries over the same 10-year period. Selection was performed by computer random-number generation. Inclusion and exclusion criteria were identical to those in the stillbirth group. Controls were not matched by gestational age Data collection For both groups, maternal and fetal characteristics were reviewed, as were fetal biometry and indications for ultrasound. Maternal height and ethnicity were self-reported. Early pregnancy weight was obtained from prenatal records or from the self-report at admission to Labor and Delivery. Definition of stillbirth [or SGA<3rd centile] Not specified Prevalence of stillbirth in the study

Study reference	**Smith 2014**
	N/A (case-control) Sample size N screened/invited = 223 stillbirths and 85,968 live births N eligible = 49 stillbirths (and four liveborn controls for each case) N enrolled = 49 stillbirths and 197 live births N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 49 N excluded from analysis = 0 N included in analysis = 49 stillbirths and 147 live births Demographics Maternal age, mean (SD) o Live births: 33.8 (6.0) years o Stillbirths: 32.8 (6.2) years BMI o Live births: 26.9 (7.2) kg/m2 o Stillbirths: 28.5 (6.7) kg/m2 Parity/gravida: NR Smoking status: NR Comorbidities: NR Previous pregnancy complications: NR
Screening Method	Index test Fetal growth curves for the identification of SGA as a predictor of stillbirth All ultrasound examinations were performed by trained sonographers and were reviewed in real time by a maternal–fetal medicine specialist or a radiologist. Estimated fetal weight was calculated using the Hadlock formula. Growth percentile of each fetus was assessed in three ways: Hadlock ultrasound norms; customised norms; and a birth-weight-based population norm Reference standard Unclear, appears that it was only ascertained by death certificate
Test Accuracy	**SGA by population growth norm** Sensitivity 14% (95% CI: 4 to 25) **SGA by ultrasound growth norm (Hadlock)** Sensitivity 33% (95% CI: 19 to 47) **SGA by customised growth norm** Sensitivity 39% (95% CI: 24 to 54)
Authors' Conclusions	Among fetuses destined to be stillborn, customised and ultrasound norms identified a greater proportion of both SGA and LGA estimated fetal weights. The customized norms performed best in identifying death among SGA fetuses. These results should be interpreted within the limitations of the study design.

Study reference	**Sutan 2010**
Study Characteristics	Design Retrospective database study Objective

Study reference

Sutan 2010

To determine the risk factors of unexplained antepartum stillbirth in Scotland from 1994 to 2003 and assess their value as a screening tool

Dates
1994 to 2003

Country
UK

Setting
Scottish Stillbirth and Neonatal Death Enquiry Form with linkage to the Maternity Inpatient and Day Case Record

Population Characteristics

Patient recruitment
The data from the retrospective databases contained all stillborn and live born infants irrespective of gestational age and all women receiving care in the obstetric specialities/health professions listed as either in-patient or day cases or also referral cases, but excluded home births and births in non-NHS hospitals. Analyses were performed for singleton pregnancies occurring at 20 completed weeks of gestation and more or occurring after the fetus reached a body mass of 200 g or more. Multiple births were excluded.

Data collection
Gestational age was calculated based on the estimated date of delivery.

Definition of stillbirth [or SGA<3rd centile]
Unclear but the perinatal data collected included late fetal deaths from 20 weeks gestation. Unexplained fetal deaths were defined as deaths occurring before labour with no evident fetal, maternal or placental abnormality sufficient to be considered as the cause of death.

Prevalence of stillbirth in the study
There were 1,452 unexplained antepartum stillbirths out of a total of 2,822 stillbirths. The prevalence of unexplained antepartum stillbirth with the presence of three risk factors (maternal age, smoking during pregnancy and maternal height) was very low at 0.2% of the population.

Sample size
N screened/invited = 541,811 births including 1452 unexplained antepartum stillbirths out of 2822 stillbirths
N eligible = NR
N enrolled = NR
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis = 1452 stillbirths and 541,811 births

Demographics

	Unexplained antepartum stillbirth (N=1452)	Others (N=541,811)
Maternal age group (years), %		
10–19	9.5	8.2
20–24	20.0	18.4
25–29	26.6	29.9
30–34	27.0	29.4
35–39	14.4	12.1
40–	2.5	1.9
Smoking during pregnancy, %		
No	44.9	57.0
Yes	33.6	26.0
Unknown	21.5	17.0
Smoking history, %		
Never	50.8	64.5
Current	35.8	27.0
Former	2.5	2.7
Parity, %		
0	39.6	35.0

Study reference	Sutan 2010		
	1	28.6	32.0
	2	16.3	18.0
	3	8.8	8.5
	4	3.7	3.7
	≥5	3.0	2.9
	Previous caesarean		
	0	90.2	91.0
	1	8.3	7.7
	2	1.4	1.1
	≥3	0.2	0.2
	Diabetes status, %		
	Pre-existing	0.3	0.5
	Gestational	0.1	0.4
	No diabetes during pregnancy	99.6	99.1
	Previous history of spontaneous abortion, %		
	0	78.8	79.7
	1	16.5	15.6
	2	3.6	3.4
	≥3	1.1	1.3
	Previous stillbirth (data from 1997 to 2003 only), %		
	0	97.0	99.1
	1	2.8	0.8
	≥2	0.2	0
	Previous neonatal death (data from 1997 to 2003 only), %		
	0	98.9	99.4
	1	0.9	0.6
	≥2	0.2	0

Screening Method	Index test Maternal characteristics (maternal age, smoking during pregnancy and maternal height) to predict stillbirth Chi-squared tests were used for comparisons of categorical data. P-values for all tests were two-sided and significance was set at P-value <0.05. Crude ORs were calculated for each explanatory variable. To control for the possibility that factors might be confounded, all factors were entered simultaneously into multivariable logistic regression models, and adjusted ORs were obtained. Logistic regression analyses were performed to predict or estimate the probability that an individual would have an unexplained antepartum stillbirth and to make a comparison with other outcomes. Bayes' theorem of conditional probabilities was used to calculate the predictive values. Reference standard Unclear. Unexplained antepartum stillbirth as recorded by the retrospective databases; the data contained all stillbirth and live born infants, and the perinatal data collected included late fetal deaths from 20 weeks gestation. It was reported that the cause of death was assigned and those deaths classified as antepartum stillbirths were not categorised further according to the Obstetric (Aberdeen) classification as had been done under the Confidential Enquiry into Maternal and Child Health reporting system. Method of diagnosis of stillbirth at the time is unclear.
Test Accuracy	In multivariable analysis, only maternal age (aOR 1.8, 95% CI 1.1 to 3.0), smoking during pregnancy (aOR 2.0, 95% CI 1.1 to 3.5), and maternal height (aOR 1.4, 95% CI 1.1 to 1.8) were significantly associated with unexplained antepartum stillbirth. The prevalence of unexplained antepartum stillbirth with the presence of three risk factors was very low at 0.2% of the population. **Maternal age, smoking during pregnancy and maternal height to predict unexplained antepartum stillbirth**

Study reference	Sutan 2010
	Sensitivity 4.2% Specificity 99.4% PPV 1.2% NPV 99.8%
Authors' Conclusions	Advanced maternal age, maternal smoking and shorter maternal height were associated with increased risk for unexplained antepartum stillbirth, but screening based on these factors would be of limited value.

Study reference	Tancrede 2015
Study Characteristics	Design Prospective cohort study Objective To evaluate the predictive values of mid-trimester serum AFP and hCG for preterm and term placenta-mediated adverse pregnancy outcomes (PMAPOs) Dates 2005 to 2010 Country Canada Setting Two hospitals in Quebec City
Population Characteristics	Patient recruitment Women were recruited at their first prenatal visit if they were at least 18 years old and had no chronic renal disease. Each participant completed a self-administered questionnaire regarding sociodemographic and biomedical information. Only nulliparous women with singleton pregnancies were included in the analyses. Women with a fetus with aneuploidy or lethal anomalies or with a pregnancy that did not exceed 20 weeks of gestation were excluded Data collection After delivery, a research nurse reviewed all medical records to collect delivery data, and any cases with suspected PMAPO, including all cases of gestational hypertension of pregnancy, were reviewed by a physician blinded to the AFP and hCG results to confirm the presence or absence of PMAPO. Maternal serum AFP and hCG levels were obtained from the hospitals' biochemical laboratory records. Definition of stillbirth [or SGA<3rd centile] Not reported; intrauterine fetal death is reported and likely to mean stillborn fetuses after 20 weeks of gestation (pregnancies lost before 20 weeks excluded) Prevalence of stillbirth in the study There were 5 intrauterine fetal deaths. Sample size N screened/invited = 7,929 N eligible = 3,466 N enrolled = 3,466 N excluded (with reason) = 0 N lost to follow-up = 0 N completed = N excluded from analysis = Unclear

Study reference	**Tancrede 2015**
	N included in analysis = 3,466 women were enrolled but it was reported that 2,110 and 2,125 of these women had data available for serum AFP and serum hCG Demographics • Maternal age, years o Term with PMAPO: 28.2 o Preterm with PMAPO: 28.8 o Preterm without PMAPO: 28.1 o Term without PMAPO: 27.6 • BMI, kg m^{-2} o Term with PMAPO: 22.4 o Preterm with PMAPO: 22.9 o Preterm without PMAPO: 22.7 o Term without PMAPO: 22.5 • GA at birth o Term with PMAPO: 39.4 o Preterm with PMAPO: 35.6 o Preterm without PMAPO: 35.4 o Term without PMAPO: 39.9 • Parity/gravida: NR • Smoking status: NR • Comorbidities: NR • Previous pregnancy complications: NR
Screening Method	Index test Maternal serum AFP and hCG Levels were expressed as multiples of the median (adjusted for gestational age) and collected between 13 and 17 weeks of gestation as part of the provincial Down syndrome screening program. Reference standard PMAPOs, which included intrauterine fetal death, were confirmed by a physician
Test Accuracy	**AFP >2.0 MoM (n=2110)[a]** Sensitivity o Preterm: 0% o Term: - Specificity o Preterm: 95.9% o Term: - PPV o Preterm: 0/87 o Term: 0/87 NPV o Preterm: 99.8% o Term: - **hCG >2.0 MoM (n=2125)[a]** Sensitivity o Preterm: 40% o Term: - Specificity

Study reference	Tancrede 2015
	○ Preterm: 89.9% ○ Term: - PPV ○ Preterm: 0.9% ○ Term: 0/218 NPV ○ Preterm: 99.8% Term: - [a] 3466 women were enrolled but it was reported that 2110 and 2125 of these women had data available for serum AFP and serum hCG
Authors' Conclusions	Maternal serum AFP or hCG >2.0 MoM increases the risk of preterm PMAPO but not term PMAPO in the study population. The authors suggest that women with elevated serum AFP or hCG should receive standard pregnancy care once they have reached 37 weeks of gestation if fetal growth is in the normal range.

Study reference	Trudell 2015
Study Characteristics	Design Retrospective cohort study Objective To determine if the use of a sex-specific standard to define small-for-gestational age (SGA) will improve prediction of stillbirth Dates January 1990 to December 2009 Country USA Setting Perinatal database of the Washington University in St Louis, Missouri (an academic tertiary care centre)
Population Characteristics	Patient recruitment All singleton pregnancies were included, except those complicated by fetal anomalies or where birthweight or sex status data was missing Data collection The perinatal database is a large validated system updated and maintained daily by a dedicated perinatal research nurse. Maternal demographics, medical and obstetric history were entered into the database using a self-report questionnaire at the initial ultrasound visit. Follow-up information history was entered into the database from the medical record or through contact with the patient and referring physician if the patient delivered outside the authors' hospital system Definition of stillbirth [or SGA<3rd centile] Intrauterine fetal death ≥20 weeks' gestation Prevalence of stillbirth in the study 319 stillbirths out of 57,170 (*0.56%)* Sample size N screened/invited = NR N eligible = 57,170 N enrolled = 57,170 N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 57,170 N excluded from analysis = 0 N included in analysis = 57,170 Demographics

Study reference	**Trudell 2015**

	• Maternal age, median (IQR)
	  ○ Stillbirths: 29 (23 to 35) years
	  ○ Live births: 31 (26 to 35) years
	• BMI, median (IQR)
	  ○ Stillbirths: 25.9 (22.7 to 33.0) kg/m^2
	  ○ Live births: 24.8 (21.7 to 29.5) kg/m^2
	• Chronic hypertension, n (%)
	  ○ Stillbirths: 18 (5.6)
	  ○ Live births: 1,406 (5.7)
	• Preeclampsia, n (%)
	  ○ Stillbirths: 21 (6.7)
	  ○ Live births: 4,561 (8.0)
	• Pre-gestational diabetes, n (%)
	  ○ Stillbirths: 13 (4.1)
	  ○ Live births: 1,094 (1.9)
	• Gestational diabetes, n (%)
	  ○ Stillbirths: 15 (4.8)
	  ○ Live births: 3,003 (5.3)
	• Smoking, n (%)
	  ○ Stillbirths: 48 (15.1)
	  ○ Live births: 6,190 (10.9)

Screening Method	Index test Growth standards Ultrasound examinations were performed by certified obstetric sonographers. Gestational age was confirmed by ultrasound criteria. Gestational age was reassigned if there was a discrepancy of ±5 days in the first trimester or ±10 days in the second trimester. To compare the predictive abilities of the two growth standards, receiver operating characteristic curves were developed for the prediction of stillbirth among SGA pregnancies as defined by the non-sex-specific standard and the sex-specific standard Reference standard Final diagnoses were made by board-certified Maternal-Fetal Medicine specialists
Test Accuracy	**Non-sex-specific growth standard** Sensitivity 102/319 (32%) Specificity 92.8% PPV 2.4% NPV 99.6% LR+ 4.43 (95% CI 3.8 to 5.2) LR− 0.73 (95% CI 0.7 to 0.8) Detection rate (sensitivity) of 44% for a fixed false-positive (1-specificity) rate of 10% **Sex-specific growth standard** Sensitivity

Study reference	**Trudell 2015**
	204/319 (64%) Specificity 92% PPV 4.3% NPV 99.8% LR+ 7.96 (95% CI 7.3 to 8.7) LR- 0.39 (95% CI 0.3 to 0.5) Detection rate (sensitivity) of 70% for a fixed false-positive (1-specificity) rate of 10%
Authors' Conclusions	The sex specific growth standard was better at discriminating the SGA fetuses at risk for stillbirth than the non-sex specific growth standard, and was associated with a higher detection of stillbirth for a fixed false-positive rate. The findings demonstrate the superiority of a sex specific growth standard for the prediction of stillbirth among SGA fetuses and support the clinical utility of sex specific standards of fetal growth.

Study reference	**Trudell 2017**
Study Characteristics	Design Retrospective cohort study Objective To generate a clinical prediction tool for stillbirth that combines individual maternal risk factors to provide an evidence-based approach to the estimation of stillbirth risk with the goal of identifying women who may benefit most from antenatal testing but would otherwise not have met generally accepted indications for testing. Dates 1999 to 2009 Country USA Setting Washington University School of Medicine (quaternary referral centre for a large catchment area in Midwest US)
Population Characteristics	Patient recruitment The study database is a validated perinatal database consisting of prospectively collected information on singleton pregnancies presenting for routine second trimester anatomic screening. After exclusion of 12,280 women with missing delivery information, a cohort of 64,173 women were included for analysis. Data collection At the time of inclusion, patients were given a survey to return following delivery (and received a phone call from the research coordinator if not returned within 4 weeks of expected date of delivery). If unable to be contacted, the primary obstetrician was contacted for follow-up. Outcome data was also gathered from the electronic medical record if delivery occurred with the study institution's healthcare system. Definition of stillbirth [or SGA<3rd centile] Stillbirth at or beyond 20 weeks gestation Prevalence of stillbirth in the study There were 464 stillbirths out of 64,173 women, for a stillbirth rate of 7.2/1,000 total births. Sample size N screened/invited = NR N eligible = 64,173

Study reference

Trudell 2017

N enrolled = NR
N excluded (with reason) = 12,280 (below)
N lost to follow-up = 12,280 had missing delivery information
N completed = NR
N excluded from analysis = 0
N included in analysis = 64,173

Demographics

- Maternal age (years), median (IQR): 31 (26 to 35)
- BMI (N=51,669), median (IQR): 25.23 (22.36 to 30.02)
- Parity, median (IQR): 1 (0 to 2)
- Nulliparous, n (%): 22,121 (38.59)
- Current smoker, n (%): 6,265 (10.96)

Comorbidities

- Chronic hypertension (N=57,326), n (%): 1,430 (2.49)
- Preeclampsia, n (%): 4,587 (8.08)
- Pre-gestational diabetes (N=57,326), n (%): 1,112 (1.94)
- Gestational diabetes, n (%): 3,020 (5.32)

Previous pregnancy complications

- Stillbirth: 330 (0.58%). This is assumed to be previous stillbirth as this value does not match the reported number of stillbirths in the current cohort.

Screening Method

Index test

A stillbirth risk calculator and stillbirth risk score based on maternal factors were developed for the prediction of stillbirth at or beyond 32 weeks excluding fetal anomalies and aneuoploidy.

Logistic regression was used to develop models for the prediction of stillbirth. Maternal risk factors for stillbirth were identified from the literature in combination with univariate analysis for stillbirth at or beyond 20 weeks gestation. Starting with the most comprehensive model including all maternal risk factors, a backward stepwise selection process was utilised. Variables with non-significant p-values (>0.05) were identified first and elimination began with the variable with an OR closest to 1. A variable was considered significant and kept in the model if there was a reduction in the discriminative ability of the model as determined by the AUC or if the AUC did not change but the beta-coefficients for the remaining variables changed by greater than 10%. If a continuous variable had a significant impact on the model then it was further explored in categorical and dichotomous forms. The final variable format used in the model was determined by the rules stated above for retention of significant variables. At initial model development, multiple models were evaluated for the prediction of stillbirth at or beyond the gestational ages of 20, 24, 28 and 32 weeks, to explore model discrimination for the prediction of stillbirth throughout gestation. Each model excluded pregnancies ending prior to the gestational age specified.

The sensitivities and specificities were examined over a range of clinically relevant probabilities. Statistical cut-points were determined using the Youden Index and Liu test. Clinical cut-points were explored through multiple case scenarios with the use of a stillbirth risk prediction calculator that was generated using the beta-coefficients from the final model.

Reference standard

The method of diagnosis of antepartum stillbirth was unclear. At the time of inclusion, patients were given a survey to return following delivery (and received a phone call from the research coordinator if not returned within 4 weeks of expected date of delivery). If unable to be contacted, the primary obstetrician was contacted for follow-up. Outcome data was also gathered from the electronic medical record if delivery occurred with the study institution's healthcare system.

Test Accuracy

Stillbirth calculator for the prediction of stillbirth at or beyond 32 weeks gestation, excluding fetal anomalies and aneuploidy (used to determine clinical cut–points):

Cut-point: no. of stillbirths/10,000 ongoing pregnancies	Sensitivity %	Specificity %	Correctly classified	LR+
5	100	0.11	0.30	1.00
12	87.5	23.6	23.7	1.14

Study reference	**Trudell 2017**

17	60.4	62.8	62.8	1.62
18	55.2	67.4	67.4	1.69
27	33.3	83.7	83.6	2.04
34	25.0	91.7	91.6	3.03
73	4.2	99.1	98.9	4.37

Stillbirth risk score to predict stillbirth at or beyond 32 weeks excluding fetal anomalies and aneuploidy:

The clinical cut-point as determined by the scenario of a 25-year-old white multiparous female with a BMI of 24 kg/m2 with pre-gestational diabetes who does not smoke and does not have chronic hypertension is 3 points. The statistical cut-point as determined by the methods of Youden and Liu were determined to be 1.5 points and 2.5 points, respectively.

Score cut-point	Sensitivity %	Specificity %	Correctly classified	LR+
0	100	0	0.2	1.0
1	91.7	12.1	12.2	1.04
2	78.1	40.9	41.0	1.32
3	**53.1**	**65.4**	**65.3**	**1.54**
4	34.4	82.6	82.5	1.98
5	22.9	93.1	93.0	3.32
6	11.5	97.5	97.4	4.64
7	3.13	99.1	98.9	3.48
8	0	99.6	99.4	0
9	0	99.9	99.7	0
10	0	100	99.8	0
11	0	100	99.8	0

Authors' Conclusions	Using maternal risk factors, a stillbirth risk calculator and a simplified stillbirth risk score were developed and internally validated to predict the risk of stillbirth at or beyond 32 weeks gestation. The stillbirth calculator and simple risk score demonstrated modest discrimination but clinical significant performance with no difference in overall performance between the tools [(AUC 0.66 95% CI 0.60 to 0.72) and (AUC 0.64 95% CI 0.58 to 0.70)].

Study reference	**Valino 2016a (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016 b, Yerlikaya 2016)**
Study Characteristics	Design Prospective cohort study Objective To investigate the potential value of UtA-PI at 30 to 34 weeks' gestation in the prediction of adverse perinatal outcome, by examining the relationship between UtA-PI and the rates of PE, delivery of a SGA neonate, stillbirth, Caesarean section for fetal distress, umbilical arterial cord blood pH ≤7.0 or umbilical venous cord blood pH ≤7.1 and 5-min Apgar score <7. Dates May 2011 to August 2014 Country UK Setting King's College Hospital, University College London Hospital, and Medway Maritime Hospital, Kent
Population	Patient recruitment

Study reference	**Valino 2016a (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016 b, Yerlikaya 2016)**
Characteristics	Women were prospectively screened whilst attending their routine hospital visit at 30+0 to 34+6 weeks' gestation in the third trimester at the study hospitals. Pregnancies with major fetal abnormalities or genetic syndromes diagnosed prenatally or postnatally, or those with no follow-up were excluded. Data collection The visit was attended at 30+0 to 34+6 weeks' gestation, and included the recording of maternal characteristics and medical history, estimation of fetal size from transabdominal ultrasound measurement of fetal head circumference, abdominal circumference and femur length. Gestational age was determined by the measurement of fetal crown-rump length at 11 to 13 weeks or the fetal head circumference at 19 to 24 weeks. Transabdominal colour Doppler ultrasound was used to visualise the left and right UtAs at their apparent crossover with the external iliac arteries. Pulsed-wave Doppler was then used to assess impedance to follow; when three similar waveforms were obtained consecutively the PI was measured, and the mean PI of the two vessels calculated. Data on pregnancy outcomes were collected from the hospital maternity records or the general practitioners of the women. Definition of stillbirth [or SGA<3rd centile] Unclear Prevalence of stillbirth in the study There were 80 stillbirths out of 30,261 pregnancies, including 73 antepartum stillbirths and seven intrapartum. Sample size N screened/invited = 31,804 N eligible = NR N enrolled = 30,261 N excluded (with reason) = 1,534 (major fetal abnormalities or genetic syndromes diagnosed prenatally or postnatally [=206, 0.6%], no follow- up [n=1,337, 4.2%]) N lost to follow-up = 1,337 (4.2%) N completed = 30,261 N excluded from analysis = 0 N included in analysis = 30,261 Demographics • Maternal age, years, median (IQR) o Total population: 31.3 (26.8 to 35.0) o Stillbirths: 30.0 (25.2 to 35.9) Parous, n (%) o Total population: 15,076 (49.8) o Stillbirths: 38 (47.5) Nulliparous, n (%) o Total population: 15,185 (50.2) o Stillbirths: 42 (52.5) • Smoking, n (%) o Total population: 2,741 (9.1) o Stillbirths: 11 (13.8) Comorbidities • Chronic hypertension, n (%) o Total population: 404 (1.3) o Stillbirths: 2 (2.5) • Systemic lupus erythematosus/ antiphospholipid syndrome, n (%) o Total population: 58 (0.2) o Stillbirths: 0 (0.0) • Diabetes mellitus, n (%) o Total population: 281 (0.9) o Stillbirths: 0 (0.0)
Screening Method	Index test UtA-PI (with adjustment for maternal characteristics and medical history) for the prediction of stillbirth The measured UtA-PI value was expressed as MoM after adjustment for variables relating to maternal characteristics and medical history that affect this measurement. The association between $\log_{10}$ MoM UtA-PI and

Study reference	**Valino 2016a (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016 b, Yerlikaya 2016)**
	birth-weight Z score in each of the adverse perinatal-outcome groups and those without an adverse outcome was examined in scatterplots. Univariable and multivariable logistic regression analysis was used to determine if log10 MoM UtA-PI had a significant additional contribution to that of maternal characteristics, medical history and obstetric factors in predicting adverse outcome. The detection rate (DR) and false-positive rate (FPR) of screening by UtA-PI were estimated for each adverse outcome. Reference standard Unclear, stillbirth is not defined and so it is not clear how stillbirth was diagnosed; data on pregnancy outcomes were collected from the hospital maternity records or the general practitioners of the women.
Test Accuracy	**UtA-PI >95th percentile for all stillbirth** Sensitivity 16.3% (13/80) FPR 5.5% (1,645/30,181) **UtA-PI >95th percentile for stillbirth (birth weight <10th centile)** Sensitivity 24.0% (6/25) FPR 12.7% (428/3,379) **UtA-PI >95th percentile for stillbirth (birth weight ≥10th centile)** Sensitivity 12.7% (7/55) FPR 4.5% (1,217/26,802)
Authors' Conclusions	The findings demonstrate that high UtA-PI at 30 to 34 weeks' gestation is associated with subsequent development of PE, delivery of SGA neonates and stillbirth. The rationale for the study was that, if adverse outcome is the consequence of impaired placentation, prenatal care should be directed at identifying and monitoring pregnancies with high UtA-PI rather than only those with small fetuses. The findings confirm that, although the incidence of adverse perinatal outcome was higher in SGA than in non-SGA fetuses, the majority of cases for each adverse outcome were in the appropriately-grown group, including about 70% of stillbirths and more than 80% of cases that underwent Cesarean section for fetal distress in labour, low cord blood pH and low 5-min Apgar score.

Study reference	**Valino 2016b (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016a, Yerlikaya 2016)**
Study Characteristics	Design Prospective cohort study Objective To investigate the potential value of UtA-PI, UA-PI, MCA-PI, MAP and serum levels of PlGF and sFlt-1 at 30 to 34 weeks' gestation in the prediction of adverse perinatal outcome, including development of PE, delivery of a SGA neonate, stillbirth, Caesarean section for fetal distress before or during labour, umbilical arterial cord blood pH ≤7.0 or umbilical venous cord blood pH ≤7.1, 5-min Apgar score <7 and admission to the neonatal unit (NNU). Dates May 2011 to August 2014 Country

Study reference	**Valino 2016b (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016a, Yerlikaya 2016)**
	UK Setting King's College Hospital, University College London Hospital, and Medway Maritime Hospital, Kent
	Patient recruitment Women attending for their routine hospital visit in the third trimester of pregnancy at the study hospitals at 30+0 to 34+6 weeks' gestation. The pregnancies included in the study were those with data available on all eight biomarkers and resulted in the live birth or stillbirth of a phenotypically normal baby at ≥24 weeks' gestation. Data collection Data on pregnancy outcome were collected from the hospital maternity records or the general medical practitioners of the women. Definition of stillbirth [or SGA<3rd centile] Unclear, although only pregnancies that resulted in the live birth or stillbirth of a phenotypically normal baby at ≥24 weeks' gestation. Prevalence of stillbirth in the study There were 23 cases of stillbirth (0.28%) Sample size N screened/invited = 8,268 N eligible = 8,268 N enrolled = 8,268 N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 8,268 N excluded from analysis = 0 N included in analysis = 8,268
Population Characteristics	Demographics • Maternal age, years, median (IQR) o Total population: 31.0 (26.6 to 34.7) o Stillbirths: 32.9 (23.2 to 37.2) Parous, n (%) o Total population: 4,198 (50.8) o Stillbirths: 10 (43.5) Nulliparous, (%) o Total population: 4,070 (49.2) o Stillbirths: 13 (56.5) • Smoking, n (%) o Total population: 827 (10.0) o Stillbirths: 3 (13.0) Comorbidities • Chronic hypertension, n (%) o Total population: 121 (1.5) o Stillbirths: 1 (4.3) • Systemic lupus erythematosus/ antiphospholipid syndrome, n (%) o Total population: 15 (0.2) o Stillbirths: 0 (0.0) • Diabetes mellitus, n (%) o Total population: 80 (1.0) o Stillbirths: 0 (0.0) • Previous miscarriage: NR • Previous stillbirth: NR
Screening Method	Index test UtA-PI, UA-PI, MCA-PI, MAP, PlGF and sFlt-1 levels at 30 to 34 weeks' gestation to predict stillbirth At 30+0 to 34+6 weeks' gestation, maternal characteristics and medical history were recorded, with estimation of fetal size from transabdominal ultrasound measurement of fetal head circumference, abdominal circumference

Study reference	**Valino 2016b (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011, Mastrodima 2016, Valino 2016a, Yerlikaya 2016)**
	and femur length. Gestational age was determined by the measurement of fetal crown–rump length at 11–13 weeks or the fetal head circumference at 19–24 weeks. Transabdominal colour-flow mapping was used to visualize the UtA, UA and fetal MCA. Pulsed-wave Doppler was then used to obtain waveforms; when three similar waveforms were obtained consecutively the PI was measured. The measured values of UtA-PI, UA-PI, MCA-PI, MAP, PlGF and sFlt-1 were expressed as MoMs after adjusting for variables from maternal characteristics and medical history that affect these measurements. Univariable and multivariable logistic regression analysis was used to determine if log10MoM of each biomarker had a significant contribution in predicting each adverse outcome. The DR and FPR of screening were estimated for each adverse outcome. The performance of screening was determined by ROC curves analysis. Reference standard Unclear
Test Accuracy	**Sensitivity for stillbirth by combined screening** 30.4% (FPR 10%)
Authors' Conclusions	Multivariable logistic regression analysis demonstrated that significant contributions for prediction of stillbirth were provided by estimated fetal weight, UtA-PI and MCA-PI (area under ROC curve (AUC), 0.683 (95% CI 0.568 to 0.797). Impaired placentation/placental dysfunction and fetal hypoxemia were observed in some of the pregnancies resulting in stillbirth, in those developing fetal distress in labour necessitating delivery by Cesarean section and in those requiring admission to NNU. However, the performance of screening with biomarkers at 32 weeks' gestation for these complications is poor with respective DRs of 30%, 16% and 25%, at a FPR of 10%.

Study reference	**Yerlikaya 2016 (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011 and Mastrodima 2016)**
Study characteristics	Design Prospective cohort study Objective To examine the accuracy of a previously published model in a population of 79,559 pregnancies screened after development of the model, and to derive an updated model using the total screened population of 113,415 pregnancies. The third objective was to evaluate the performance of the new model in screening for all stillbirths and for subgroups of stillbirths that occurred due to impaired placentation and to unexplained or other causes. Dates March 2006 to October 2015 Country UK Setting King's College Hospital and Medway Maritime Hospital, UK
Population Characteristics	Patient recruitment Women attending routine pregnancy care at 11+0 to 13+6 and at 19+0 to 24+6 weeks' gestation at King's College Hospital and Medway Maritime Hospital were recruited. Women were included in the study if they had a singleton pregnancy who delivered a phenotypically normal live birth or stillbirth ≥24 weeks' gestation. Pregnancies with aneuploidy, major fetal abnormality, those ending in a miscarriage or termination of pregnancy or stillbirths due to intrapartum causes were excluded. Data collection Data on pregnancy outcome were obtained from the maternity hospital records or the general practitioners of the participating women. The hospital maternity records of all women with antepartum stillbirths were reviewed to determine whether the death was associated with pre-eclampsia, placental abruption, a birth weight <10th percentile for gestational age or it was due to other/unexplained reasons. Combined screening for fetal aneuploidy (at the first visit) and assessment of fetal growth and anatomy (at the second visit) were performed. Definition of stillbirth [or SGA<3rd centile]

Study reference	Yerlikaya 2016 (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011 and Mastrodima 2016)
	Stillbirth was defined as a pregnancy loss occurring ≥24 weeks Sample size and demographics A total of 119,622 pregnancies were prospectively screened, and 113,415 pregnancies met the study entry criteria. The 6207 cases were excluded because they had missing outcome data (n=3517), the pregnancies resulted in miscarriage or termination, there were major fetal chromosomal abnormalities, babies with major fetal defects were born (n=2649) or stillbirth occurred due to intrapartum factors (n=41). Prevalence of stillbirth in the study There were 396 stillbirths (0.35%) antepartum stillbirths; 230 (58%) were secondary to impaired placentation and 166 (42%) were due to other or unexplained causes. Sample size N screened/invited = 119,622 N eligible = 113,415 N enrolled = 113,415 N excluded (with reason) = 6,207 (missing outcome data n=3,517; miscarriage, termination, major fetal chromosomal abnormalities or fetal defects n=2,649; stillbirth due to intrapartum factors n=41) N lost to follow-up = 0 N completed = 113,415 N excluded from analysis = 0 N included in analysis = 113,415 Demographics • Maternal age, years, median (IQR) o Live births: 30.9 (26.3 to 34.5) o Stillbirths: 30.4 (25.5 to 35.5) • Nulliparous, n (%) o Live births: 54,206 (48.0) o Stillbirths: 30.4 (25.5 to 35.5) • Smoking, n (%) o Live births: 12,089 (10.7) o Stillbirths: 60 (15.2)* • Chronic hypertension, n (%) o Live births: 1,438 (1.3) o Stillbirths: 22 (5.6)† • Systemic lupus erythematosus/ antiphospholipid syndrome, n (%) o Live births: 209 (0.2) o Stillbirths: 4 (1.0)* • Diabetes mellitus, n (%) o Live births: 996 (0.9) o Stillbirths: 13 (3.3)† • Previous miscarriage o Live births: 1,306 (1.2) o Stillbirths: 5 (1.3) • Previous stillbirth o Live births: 882 (0.8) o Stillbirths: 20 (5.1)† *P<0.01; †P<0.001
Screening Method	Index test The population was screened for all stillbirths, unexplained stillbirths and those due to abnormal placentation using an algorithm based on maternal factors. The distribution of patient-specific risks was used to determine the performance of screening by receiver–operating characteristics (ROC) curves analysis and the detection rates (DR) and FPR were estimated. Reference standard Unclear. The authors report that the screening model was derived and tested using the same dataset

Study reference	**Yerlikaya 2016 (Cohort may include women from Akekolar 2016a and b, Aupoint 2016, Bakalis 2011 and Mastrodima 2016)**
Test Accuracy	All stillbirth (N=396) Sensitivity (95% CI) at 95% specificity (5% FPR) ○ 18.4 (14.6 to 22.2) Sensitivity (95% CI) at 90% specificity (10% FPR) ○ 29.0 (24.5 to 33.4) Unexplained stillbirths (N=166) Sensitivity (95% CI) at 95% specificity (5% FPR) ○ 16.3 (10.7 to 21.9) Sensitivity (95% CI) at 90% specificity (10% FPR) ○ 25.9 (19.2 to 32.6) Due to impaired placentation at any gestational age (N=230) Sensitivity (95% CI) at 95% specificity (5% FPR) ○ 20.0 (14.8 to 25.2) Sensitivity (95% CI) at 90% specificity (10% FPR) ○ 31.3 (25.3 to 37.2)
Authors' Conclusions	A model based on maternal characteristics and medical history recorded in early pregnancy can potentially predict one-third of subsequent stillbirths. The extent to which such stillbirths could be prevented remains to be determined.

Table 22. Studies relevant to criterion 11

Question 2

Study reference	**Lees 2015 (TRUFFLE)**
Study Design	Design RCT Objective To assess whether changes in the fetal ductus venosus Doppler waveform (DV) could be used as indications for delivery instead of cardiotocography (CTG) short-term variation (STV) Dates 1st January 2005 to 1st October 2010 Countries Germany, Italy, the Netherlands, Austria, UK Setting 20 tertiary care centres with a fetal medicine unit
Population Characteristics	Patient recruitment Women were eligible for inclusion if they were admitted to hospital with singleton pregnancies and were diagnosed with fetal growth restriction. Pregnancies had a gestational age assigned from crown rump length before 14 weeks or biparietal diameter between 14.0 weeks and 22.0 weeks Data collection CTG was recommended at least once a week, but could be more frequent according to local policy. Safety net criteria for delivery applied to all women irrespective of randomised group if the cut-off rescue value of STV for delivery based on CTG at 26.0 to 28.9 weeks less than 2.6 ms; if short-term variation less than 3 ms at 29.0 to

Study reference	Lees 2015 (TRUFFLE)

32.0; or if, irrespective of STV, spontaneous repeated persistent unprovoked decelerations on CTG occurred. Note that these STV limits are lower than those set for delivery in the CTG STV group. For Doppler measurements, investigators from each unit submitted ductus venosus images for blinded scoring by two members of the quality control group. CTG monitoring was done using equipment that allows waveform analysis with Oxford Sonicaid 8002 or equivalent Dawes-Redman software-based algorithm. The recordings were at least 45 min in duration.

Allocation methods

Participants were randomly assigned to one of three groups in a 1:1:1 ratio to establish the timing of delivery of their fetus with severe early onset fetal growth restriction. Baseline maternal and fetal data were collected with a secure internet data entry page. Eligible patients were allocated through the study website. Allocation to groups was done with randomly sized blocks, stratified for gestational age (<29 weeks vs ≥29 weeks) and for participating centres. Concealment of the allocated monitoring regime was not possible, and clinicians responsible for the care of the women entered in the study and women themselves were aware of the treatment allocation. However, the paediatrician doing the follow-up examination was masked to follow-up assessment and data entry allocation

Definition of stillbirth [or SGA<3rd centile]

Fetal death not defined but instead split into 'fetal death no intervention', where parents in this category refused the study intervention (delivery as indicated) and 'unexpected death' defined as fetal death not anticipated between scheduled follow-up appointments, which appears to show stillbirth occurring despite parents following the study intervention as planned. However, gestational age at inclusion was between 26 weeks and 31.9 weeks so it can be assumed that 'fetal death' occurred after 26 weeks. Neonatal deaths are reportedly separately.

Sample size

N screened/invited = NR

N eligible = 542

N enrolled = 511

N excluded (with reason) = 31 were not enrolled (21 refused participation, 2 not approached, 2 language problems, 6 organisational problems) and 8 women from two centres were entered into the study but were not randomly allocated treatment because the centres had not reported any outcomes

N lost to follow-up = 59 (CTG STV N=21; DV p95 N=25; DV no A N=13)

N completed = NR

N excluded from analysis = 1 (no neonatal data). 0 excluded for analysis of fetal death

N included in analysis = For primary outcome: 443 (CTG STV N=144; DV p95 N=142; DV no A N=157); for fetal death outcome: 503 (CTG STV N=166; DV p95 N=167; DV no A N=170)

Demographics

	CTG STV N=166	DV p95 N=167	DV no A N=170
Age	31 (5)	31 (6)	31 (6)
Parity/gravida	101 (61%) nulliparous	103 (62%) nulliparous	115 (68%) nulliparous
Smoking status (smoking during pregnancy)	31 (19)	24 (12)	22 (13)
Comorbidities			
Renal morbidity	5 (3)	5 (3)	1 (1)
Chronic hypertension	14 (8)	21 (14)	19 (11)
Diabetes	3 (2)	3 (2)	3 (2)
Other medical disease	33 (20)	29 (17)	29 (17)
Previous pregnancy complications	NR	NR	NR

Data reported as n (%), mean (standard deviation) or median (range).

Monitoring Strategy	**Intervention** DV p95 (ductus venosus pulsatility index >95th percentile, n=167) Women delivered on the basis of early ductus venosus changes DV no A (DV A wave with no or reversed flow, n=170) Women delivered on the basis of late ductus venosus changes

Study reference	**Lees 2015 (TRUFFLE)**
	Normal care CTG STV (cardiotocograph short-term variation, n=166) Timing of delivery was dependent on presence of reduced STV (STV <3.5 ms at <29 weeks of gestation or STV <4 ms at ≥29 weeks of gestation). In cases where corticosteroids had been given for fetal lung maturity, no decision regarding delivery was made on the grounds of reduced variation from 24 h to 72 h after the first intramuscular dose because corticosteroid administration is known to lead to transient reduced STV. Umbilical artery Doppler measurements were taken in this group, but no waveform measurements of the ductus venosus were recorded
Outcomes measured	The primary outcome for this trial was survival without neurodevelopmental impairment at 2 years of age, corrected for prematurity. The secondary outcome was a composite of adverse neonatal outcome defined as fetal or postnatal death (between trial entry in-utero and discharge home from neonatal services) or one or more of the following severe morbidities: bronchopulmonary dysplasia, severe germinal matrix cerebral haemorrhage (GMH; intraventricular haemorrhage with dilation of the lateral ventricles [grade 3] or intraparenchymal haemorrhage [grade 4]), cystic periventricular leukomalacia, proven neonatal sepsis, or necrotising enterocolitis.
Effectiveness of Monitoring	Stillbirths • CTG STV: 0/166 (0%) • DV p95: 3/167 (2%) • DV no A: 4/170 (2%) No statistical analysis was performed (fetal death was not a pre-specified outcome) 59 randomised women were lost to follow-up, 1 had no neonatal data and 8 women were excluded (as no delivery or outcome data could be obtained from the centres they were enrolled in)
Authors' Conclusions	Although survival without neuroimpairment did not differ between groups, neuroimpairment at 2 years in the study was less frequent in the infants of women randomly assigned to delivery based on late ductus venosus changes compared with those randomly assigned to delivery based on computerised cardiotocograph (CTG) changes. Previous observational and retrospective studies have suggested that a worse outcome is associated with late ductus venosus changes and these studies have informed management. By contrast, these findings support waiting for late ductus venosus changes before delivery because no increase in hypoxia mediated deaths occurred and neuroimpairment is less frequent than when delivery is based on computerised CTG changes. The neonatal outcomes (as death or severe neonatal morbidity or combined) did not differ between the three groups.

Study reference	**Spaggiari 2013**
Study Design	Design Retrospective cohort study Objective To assess maternal-fetal outcomes in pregnancies associated with persistently elevated second-trimester maternal serum alpha-fetoprotein. Dates 2004 to 2008 Countries France Setting Eight hospitals
Population Characteristics	Patient recruitment Unselected women were eligible for inclusion if they underwent screening for Down Syndrome using maternal

Study reference	Spaggiari 2013
	serum in singleton pregnancies over the period 2004–2008. Data collection Second trimester (14 to 18 weeks) maternal serum markers used in screening for Down syndrome in singleton pregnancies, including hCGb and AFP, were assayed. Maternal sera from an unselected population were routinely sent from 8 hospitals. Gestational age was determined by crown-rump length at first trimester ultrasound examination (expressed in weeks and days). Patients with a high MSAFP (defined as AFP ≥ 2.5multiple of median [MoM]) were offered a second AFP assay as part of the diagnostic work-up. This second assay was performed at least 18 days later (which represents approximately 3 times the half-life of AFP in maternal serum). Reasons for not performing a second assay were as follows: (1) diagnosis of a severe fetal malformation at ultrasound examination or an IUFD before 18 days, and cases of vanishing twin or selective fetal reduction because of the artificial increase in AFP. (2) Physician or maternal failure to adhere to the procedure. Patient management consisted of standard ultrasound surveillance and, if necessary, amniocentesis for fetal karyotyping, amniotic fluid AFP assay, and cholinesterase electrophoresis (AChE). Allocation methods Patients who underwent a second MSAFP assay were classified into 2 groups: (1) group with MSAFP remaining ≥2.5 MoM and (2) group with MSAFP returning to a normal level. Clinical data, ultrasonographic reports, laboratory findings, outcome, and final diagnosis were recorded. Pregnancy complications were classified in 5 groups: severe fetal anomalies, IUFD, IUGR<3rd percentile, preeclampsia, and spontaneous premature birth. Definition of stillbirth [or SGA<3rd centile] IUFDs occurred between 15-30 weeks, with a median of 17 weeks. IUGR was defined as <3rd percentile. Sample size N screened/invited = 658 N eligible = 614 N enrolled = 614 N excluded (with reason) = 44 (vanishing twin, selective fetal reduction) N lost to follow-up = 0 N completed = 614 N excluded from analysis = 0 N included in analysis = 614 (341 patients with a second AFP assay) Demographics Age (years): 29 (IQR, 25–34) Parity/Gravidity: NR Smoking status: NR Comorbidities: NR Previous pregnancy complications: NR
Monitoring Strategy	Intervention Women with maternal serum alpha-fetoprotein ≥2.5 MoM receiving a second AFP assay. Normal care Women with maternal serum alpha-fetoprotein ≥2.5 MoM not receiving a second AFP assay.
Outcomes measured	Pregnancy outcome was documented in all cases. Outcomes were split into pregnancies without complication and complicated pregnancies, including severe fetal anomalies, IUFD, IUGR, Preeclampsia and spontaneous pre-term delivery (<34 weeks).
Effectiveness of Monitoring	Stillbirths and SGA (IUFD and IUGR)

	Outcomes of the pregnancies with AFP >2.5 MoM	Outcomes of the pregnancies without a second AFP assay	Outcomes according to AFP evolution, in the population with a second AFP assay and without clear ultrasound explanation		
	n=614 (%)	n=273	AFP remained high n= 92 (%)	AFP returned to normal n= 226 (%)	*P* value
Pregnancies without	372 (60.6)	*128 (46.9)*	55 (59.8)	189 (83.6)	<0.001

Study reference	**Spaggiari 2013**					
	complication					
	IUFD	60 (9.8)	46 (16.8)	9 (9.8)	5 (2.1)	0.005
	IUGR (<3rd centile)	36 (5.8)	18 (6.6)	10 (10.9)	8 (3.5)	0.01
Authors' Conclusions	In conclusion, high MSAFP level is associated with a high rate of pregnancy complications. When maternal serum alpha-fetoprotein returns to a normal level on a second assay, the risk of adverse outcome significantly decreases, but these pregnancies are still at risk of complications and therefore need close surveillance. Repeat maternal serum alphafetoprotein assay allows identification of patients who should be offered amniocentesis to evaluate the risk of nephrotic syndrome and epidermolysis bullosa. Alpha-fetoprotein should be monitored in pregnancies associated with unexplained high maternal serum alpha-fetoprotein. A management strategy based on ultrasound examination, second maternal serum alpha-fetoprotein assay and amniocentesis is proposed to improve prenatal counselling and management of such pregnancies. However, a prospective study remains necessary to evaluate it.					

Study reference	**Tveit 2009**
Study Design	Design Prospective study Objective To examine two cohorts of women with DFM before and during two consensus-based interventions aiming to improve care through: 1) written information to women about fetal activity and DFM, including an invitation to monitor fetal movements, 2) guidelines for management of DFM for health-care professionals. Dates Interventions implemented: November 2005 to March 2007 (Quality assessment period: April to October 2005) Countries Norway Setting Fourteen delivery units in eastern Norway and the city of Bergen
Population Characteristics	Patient recruitment All singleton third trimester pregnancies of at least 28 weeks gestation or more who reported a concern for DFM (either by spontaneous reporting or upon questioning), were included in the study. Recurrent visits for DFM in already registered pregnancies were excluded as we intended to report the number of women newly reporting DFM. Data from women with a stillborn infant were obtained separately, to ensure completeness of mortality data, but stillbirths not initially identified by DFM were subsequently excluded, as were pregnancies with a gestational age under 28 weeks and multiple pregnancies. To ensure unbiased registrations for quality-assurance of clinical practice at the individual hospital, maternal consent was not sought. Data collection The registration period included 7 months of baseline observation followed by 17 months of intervention: from April 1, 2005 to March 31, 2007. Women fulfilling the inclusion criteria were registered prospectively by the caregiver at the time the woman presented to the hospital. Pregnancy outcome were collected independently after delivery from the medical files by study coordinator at each hospital. Data were anonymized and submitted to the study-coordinating centre. DFM was defined as any woman presenting with concerns for DFM, irrespective of whether this was based on her subjective opinion or it emerged during an antenatal visit for other reasons. In addition to the registrations by our study protocol, the numbers of births and stillbirths from our population were obtained from the Medical Birth Registry in Norway to assess overall trends in stillbirth, for the most updated period available: April 2005 to December 2006. Allocation methods No allocation methods provided. Definition of stillbirth [or SGA<3rd centile] As there was only one neonatal death, all deaths are called "stillbirths". Women with pregnancies of at least 28 weeks gestation were included in this study so it is assumed that the study classifies stillbirth as death at or after 28 weeks of gestation. Sample size

Study reference | **Tveit 2009**

N screened/invited = 65,550

N eligible = 1,370 (baseline cohort); 3,534 (intervention cohort)

N enrolled = 1,370 (baseline cohort); 3,534 (intervention cohort)
N excluded (with reason) = 128 (baseline cohort); 439 (intervention cohort) (Recurrent visits, gestational age < 28 weeks, fetal deaths not initially identified by DFM and multiple pregnancies)

N lost to follow-up = Baseline cohort: n=27 (2.2%); Intervention cohort: n=57 (1.9%)

N completed = 4,253

N excluded from analysis = 0

N included in analysis = 4,253

Demographics

Characteristics	Women with DFM before the intervention*	Women with DFM during the intervention*	P value
Age, years mean (SD)	29.6 (4.9)	29.6 (5.1)	0.625
Primiparity	559 (51)	1414 (52)	0.490
Parity			
Para 0	559 (51)	1414 (52)	0.601
Para 1	372 (34)	878 (33)	
Para 2+	163 (15)	409 (15)	
Smoking status (smoker)	104 (8.8)	259 (8.9)	0.924
Comorbidities	NR	NR	NR
Previous pregnancy complications	NR	NR	NR

* Data are reported as n (%) unless otherwise noted.

Monitoring Strategy

Intervention, n=3038 women with DFM
Two consensus-based interventions aiming to improve care through:
1) written information to women about fetal activity and DFM, including an invitation to monitor fetal movements
2) guidelines for management of DFM for health-care professionals.

Normal care, n=1215 women with DFM
Baseline cohort – without consensus-based interventions

Outcomes measured

The main outcome measures were all antepartum, intrapartum and neonatal death in the delivery room (i.e., the death occurred immediately after completion of delivery) from 28 completed weeks of gestation in women who were previously registered as having one or more episodes of DFM.

Secondary outcomes for women with DFM were: severe neonatal depression, defined as Apgar score of < 3 at 5 minutes postpartum; symptoms of multisystem organ failure and pH < 7 in the umbilical artery or fetal capillary scalp, if obtained; pre-term birth (28– 36 weeks); FGR (< 10th percentile of birthweight adjusted for gender and mother's height, weight, parity, and ethnicity); fetal heart rate tracings judged clinically as nonreassuring and leading to intervention in labour; oligohydramnios defined as an amniotic fluid index of < 5 cm or at < 2.5th percentile; polyhydramnios defined as an amniotic fluid index of > 25 cm or at > 97.5th percentile; investigations undertaken for reduced FM; and examinations of DFM resulting in immediate admission for induction of labour or caesarean section.

Outcomes related to maternal behaviour were: the number of women waiting more than 24 hours with an absence of FM or more than 48 hours with a decrease of FM before contacting health-care professionals.

Effectiveness of Monitoring

Stillbirths

Pregnancy Outcomes	Baseline % (n)	Intervention % (n)	Crude OR	95% CI	P Value	Adjusted OR†	95% CI	P Value
Stillbirths (DFM)	4.2 (50)	2.4 (73)	0.58	0.41–0.84	0.004	0.51	0.32–0.81	0.004
Normally formed stillbirths (DFM)	3.9 (46)	2.2 (65)	0.57	0.39–0.83	0.004	0.50	0.31–0.81	0.005
Stillbirths (rate in total	3.0/1000	2.0/1000	0.67	0.48–0.93	0.02	Not available		

Study reference	Tveit 2009
	population) | | | | | | | | | Normally formed stillbirths (rate in total population) | 2.8/1000 | 1.8/1000 | 0.60 | 0.42–0.85 | 0.004 | Not available | | | * Univariate and multivariate logistic regression showing crude (unadjusted) and adjusted odds ratios (OR) with their 95% confidence intervals (CI). † OR adjusted for maternal weight, age, parity, smoking habits and ethnicity (considered as potential confounding factors). DFM: cases of decreased fetal movements.
Authors' Conclusions	Improved quality of management of DFM and uniform information to improve the value of the existing "selfscreening" of fetal activity was associated with a reduction in stillbirth rates in our population. For further improvements, new and individually adjusted definitions of DFM are needed, as well as randomized controlled trials to determine the optimal management and information to pregnant women with DFM. Further research is required to identify optimal methods for detecting important reductions in FM if DFM is to be an effective screening tool for adverse pregnancy outcomes.

Table 23. Studies relevant to criteria 9 and 10

Question 3

Study reference	Ayala 2012
Study Characteristics	Design RCT Objective To investigate whether bedtime treatment with low-dose ASA (100 mg/d, a dose assumingly affecting both maternal and placental thromboxane; Walsh & Wang, 1998) exerts significantly better BP control during gestation and reduction of the risk of pre-eclampsia, IUGR, and pre-term delivery than ASA upon awakening or placebo in high-risk pregnant women who entered the study protocol at ≤16 wks of gestation. Dates NR Country Spain Setting High-risk hospital unit
Population Characteristics	Patient recruitment Women were recruited from a high-risk hospital unit where they were estimated to be at a 3.5-fold increased risk of gestational hypertension or pre-eclampsia. Reasons for why women may have been treated at the unit included: familial or personal history of either gestational hypertension or pre-eclampsia; chronic hypertension; cardiovascular, endocrine, bleeding, or metabolic disease; personal history of spontaneous abortion; multiple pregnancy; obesity; and adolescent or middle-aged nulliparous pregnancy (<18 or >35 yrs). Eligibility criteria Inclusion criteria for this trial were gestational age ≤16 wks at randomisation and maternal age ≥18 yrs. Exclusion criteria were multiple pregnancy, chronic hypertension or any other condition requiring the use of BP-lowering medication, cardiovascular disorders (unstable angina pectoris, heart failure, life-threatening arrhythmia, atrial fibrillation, kidney failure, and grade III–IV retinopathy), chronic liver disease, any disease requiring the use of anti-inflammatory medication, diabetes or any other endocrine disease such as hyperthyroidism, history of drug/alcohol abuse, night/shift work employment, acquired immunodeficiency syndrome (AIDS), intolerance to ABPM, and inability to communicate and comply with all of the study requirements.

Study reference	**Ayala 2012**

Allocation methods
Participants were randomly assigned at the time of their first visit to the hospital to one of six groups, defined according to treatment (placebo or ASA, 100 mg/d) and to the timing of daily administration of ASA or placebo: upon awakening (Time 1), 8 h after awakening (Time 2), or at bedtime (Time 3). Randomisation of boxes containing either ASA or placebo followed an allocation table constructed by a computerised random-number generator. Concealed assignment of participants to the six treatment-time regimens was done according to the order of recruitment.

Blinding
Trial is reported to have been double-blind

Data collection
Women were assessed at the time of recruitment, and then scheduled every 4 wks until the 7th month of gestation and every 2 wks thereafter until delivery.

Definition of stillbirth [or SGA <3rd centile]
NR

Prevalence of stillbirth [SGA <3rd centile]
NR

Sample size
N screened/invited = NR
N eligible = NR
N enrolled = 350 total, 174 (placebo), 176 (ASA)
N excluded (with reason) = NR
N lost to follow-up = NR
N completed = NR
N excluded from analysis = NR
N included in analysis = 174 in the placebo group and 176 in the ASA group

Maternal characteristics

Characteristic	Placebo				ASA			
	Time 1 (n=59)	Time 2 (N=57)	Time 3 (N=58)	All (N=174)	Time 1 (N=58)	Time 2 (N=59)	Time 3 (N=59)	All (N=176)
Maternal age, years (SD)	31.5 (5.8)	32.0 (4.5)	30.0 (5.2)	31.1 (5.2)	31.0 (5.8)	30.4 (5.3)	29.7 (4.8)	30.3 (5.3)
BMI at randomisation, kg/m^2 (IQR)	24.8 (3.9)	25.5 (4.3)	26.3 (4.4)	25.5 (4.2)	25.8 (3.9)	24.9 (4.6)	25.4 (4.3)	25.4 (4.3)
Nulliparous, %	59.3	52.6	53.4	55.1	41.4	59.3	47.5	49.4
Gestational age at randomisation, wk mean (SD)	13.6 (1.6)	13.6 (1.4)	13.6 (1.4)	13.6 (1.5)	13.6 (1.4)	13.4 (1.5)	13.4 (1.4)	13.5 (1.4)
Previous abortion, %	32.2	28.1	25.9	30.5	29.3	32.2	32.2	31.3
Comorbidities	NR	NR	NR	NR	NR	NR	NR	NR
Smoking status	NR	NR	NR	NR	NR	NR	NR	NR

Time 1: Women randomized to ingest aspirin or placebo upon awakening. Time 2: Women randomized to ingest aspirin or placebo 8 h after awakening. Time 3: Women randomized to ingest aspirin or placebo at bedtime

Intervention

Intervention group (N=176)
Daily 100 mg aspirin
Placebo group (N=174)
Matching placebo taken daily.
Within each intervention, women were also randomised to 3 times of intervention ingestion: **Time 1**: Women randomized to ingest aspirin or placebo upon awakening. **Time 2**: Women randomized to ingest aspirin or placebo 8 h after awakening. **Time 3**: Women randomized to ingest aspirin or placebo at bedtime

Study reference	**Ayala 2012**
Outcomes Measured	The primary outcome study endpoint was total serious adverse events, which included pre-eclampsia, pre-term delivery, IUGR, and stillbirth. An additional endpoint used consisted of the composite outcome plus gestational hypertension.

Effectiveness of the Intervention

	Placebo (N=174)				ASA (N=176)			
	Time 1 (n=59)	Time 2 (N=57)	Time 3 (N=58)	All (N=174)	Time 1 (N=58)	Time 2 (N=59)	Time 3 (N=59)	All (N=176)
Stillbirth, rate (95% CI)	5.1 (-0.5 to 10.7)	3.5 (-1.3 to 8.3	0	2.9 (0.4 to 5.4)	1.7 (-1.6 to 5.1)	1.7 (-1.6 to 5.0)	0	1.1 (-0.4 to 2.7)

p=0.246

Authors' Conclusions	(i) 100 mg/d ASA should be the recommended minimum dose to be used for prevention of complications in pregnancy; (ii) ingestion of low-dose ASA for prevention of complications in pregnancy should start at ≤16 wks of gestation; and (iii) low-dose ASA ingested at bedtime, but not upon awakening, significantly lowers ambulatory BP and reduces the incidence of pre-eclampsia, gestational hypertension, pre-term delivery, and IUGR. For such clinical recommendations to be practical, however, one must be able to properly identify, among the general obstetric population, the women at higher risk for hypertension in pregnancy who might thus benefit most from daily ASA ingestion.

Study reference	**Haddad 2016**
Study Characteristics	Design RCT Objective To evaluate whether daily enoxaparin, added to low-dose aspirin, started before 14 weeks of gestation reduces placenta-mediated complications in pregnant women with previous severe pre-eclampsia diagnosed before 34 weeks of gestation. Dates 14 Nov 2009 to 21 Feb 2015 (enrolment) Country France Setting 16 secondary or tertiary care centres
Population Characteristics	Patient recruitment NR Eligibility criteria

Inclusion criteria	Exclusion criteria
1. History of severe pre-eclampsia diagnosed before 34 weeks of gestation 2. Pregnancy confirmed by ultrasonography at less than 14 weeks of gestation 3. Randomization at less than 14 weeks of gestation 4. Singleton pregnancy 5. 18 years of age or older at randomization 6. Social Security affiliation 7. Signed informed consent	1. Contraindication to heparin or aspirin therapy, including: • History of heparin-induced thrombocytopenia • Platelet count lower than 100,000 cells/µL • History of osteoporosis (potential increased risk of osteoporosis and osteoporotic fracture with heparin therapy) • Active bleeding • Documented peptic ulcer within 6 weeks (contraindication to anticoagulation) • Heparin or aspirin allergy 2. Women with serum creatinine level greater than 2.05 dL 3. Geographic inaccessibility (less likely to comply with necessary follow-up visits and care) 4. Need for anticoagulants as judged by the local investigator, including (but not limited to): • Women with recurrent pregnancy loss with antiphospholipid antibody syndrome

Study reference	Haddad 2016

	• Antiphospholipid antibody confirmed by a positive test for one or more of the following: anticardiolipin IgG (40 units/mL or greater), anticardiolipin IgM (40 units/mL or greater), anti-b2 glycoprotein IgG (20 units/mL or greater), anti-b2 glycoprotein IgM (20 units/mL or greater), or positive lupus anticoagulant • Women with previous venous thromboembolism, pulmonary embolism, or deep vein thrombosis • Women with mechanical heart valves • Women on long-term anticoagulants before pregnancy 5. Previous participation in the Heparin-Pre-eclampsia (HEPEPE) trial 6. Human immunodeficiency virus, or hepatitis B virus, or hepatitis C virus-positive serum 7. Pre-pregnancy weight of 100 kg or more

Allocation methods

Centralised random allocation using a computer-based list, with allocation concealed from investigators and stratified according to centre.

Blinding

Patients or investigators were not blind to treatment allocation. Primary outcome assessors were blind to treatment allocation.

Data collection

Patients attended clinic follow-up visits at 12, 16, 20, 24, 28, 32 and 36 weeks of gestation. Outcomes measured were: adverse events, weight, blood pressure and study drug compliance. Additional follow-up visits were performed at 22, 26, 30, 34, 37, 38, 39, 40, and 41 weeks, depending on local protocol surveillance.

All patients had at least three prenatal ultrasound examinations (at the first trimester, at 22 to 24 weeks of gestation, and at 32 to 34 weeks of gestation) according to the recommendations of the French College of Obstetrics and Gynecology, and additional examinations based on local practices and pregnancy evolution.

Definition of stillbirth [or SGA <3rd centile]

Fetal death from 22 weeks to delivery

Prevalence of stillbirth [or SGA <3rd centile]

A total of 4 stillbirths occurred in 244 women with available outcome data

Sample size
N invited = 397
N eligible = NR
N excluded (with reason) = 140 (declined n=79, gestational age ≥14 weeks n=50, previous inclusion HEPEPE n=9, pulmonary embolism n=1, other reason =1)
N randomised = 257 (total), 130 (enoxaparin + ASA), 127 (ASA)
N lost to follow-up = 2 (enoxaparin + ASA), 3 (ASA)
N completed = 122 (each arm, reported as "outcome data available")
N excluded from analysis = 6 (enoxaparin + ASA, withdrew consent), 2 (ASA, 1 included after termination, 1 previously participated)
N included in analysis = 124 (enoxaparin + ASA), 125 (ASA) included in baseline analysis; 122 (enoxaparin + ASA) and 122 (ASA) included in analysis of primary and secondary outcomes

Demographics:

Characteristic	Enoxoparin + ASA (n=124)	ASA (n=125)
Maternal age, years (SD)	31.5 (4.4)	31.7 (4.9)
Smoking during pregnancy, n (%)	2/123 (1.6)	9 (7.2)
BMI at randomisation, kg/m^2 (IQR)	25.1 (22.9 to 29.7) (n=122)	24.8 (22.0 to 28.3)
Gravidity, median (IQR)	2 (2 to 3.5)	2 (2 to 3)
Parity, median (IQR)	1 (1 to 2)	1 (1 to 2)
Gestational age, wk median (IQR)	12 (10.3 to 12.7) (n=123)	12 (10.7 to 12.7)
Chronic hypertension, n (%)	24/123 (19.5)	27 (21.6)
Pre-existing diabetes, n (%)	4 (3.2)	2 (1.6)
Prior fetal loss (≥22 wk)	23/123 (18.7)	20 (16)
Prior SGA, n (%)	81/123 (65.9)	93 (74.4)

Study reference	Haddad 2016

Severe pre-eclampsia 34 week of gestation or greater	4/123 (3.3)	7 (5.6)

N indicated where data not available for all women randomised to study arm. Also reported: ethnic origin, alcohol, drug abuse, n at specific GA, baseline blood pressure, prior fetal loss (up to 21 weeks), prior severe PE at ≥34 wk GA, prior placental abruption, prior HELLP syndrome, prior termination for PE, nephropathy, autoimmune disease, protein C deficiency, protein S deficiency, antithrombin deficiency, factor V Leiden, prothrombin gene mutation.

Intervention

Intervention group (n=124)

Enoxoparin (4,000 international units of antepartum daily by subcutaneous injection) + **low dose aspirin** (ASA, 100 mg)

Enoxoparin was continued until delivery, ASA was stopped at 35 weeks of gestation

Placebo group (n=125)

Low dose aspirin (ASA, 100 mg)

ASA was stopped at 35 weeks of gestation

Outcomes Measured

Primary endpoint

A composite outcome including any of: maternal death, perinatal death, pre-eclampsia, placental abruption, SGA (<10th centile)

Secondary endpoints:

Fetal death (from 22 weeks to delivery), severe pre-eclampsia, severe pre-eclampsia before 34 weeks of gestation, birth weight less than the fifth percentile for gestational age, neonatal death (from delivery to 28 days of life), fetal loss (15–21 weeks of gestation), neonatal morbidity (transfer to neonatal care, length of hospitalisation, mechanical ventilation greater than 24 hours, respiratory distress syndrome, necrotizing enterocolitis, periventricular leukomalacia, bronchopulmonary dysplasia, intraventricular haemorrhage grade III–IV), and enoxaparin toxicity (haemorrhage, skin reaction, thrombocytopenia [less than 100,000 cells/microliter] related to heparin, bone fracture).

Effectiveness of the Intervention

Efficacy

Outcome	Enoxoparin + ASA (n=122)	ASA (n=122)	RR (95% CI)	aRR (95% CI)	p value
Stillbirth	**1 (0.8)**	**3 (2.5)**	**0.33 (0.04 to 3.16)**	**1.6 (–1.5 to 4.8)**	**0.62**
Severe SGA (<5th centile)	15 (12.3)	21 (19.0)	0.65 (0.36 to 1.18)	7.2 (–2.7 to 17.1)	0.35

Safety

There were no statistically significant differences between occurrence of adverse events. There were two discontinuations for headaches and injection site itching in the enoxaparin+ASA arm and 1 discontinuation for epistaxis in the ASA arm.

Authors' Conclusions

In women receiving low-dose aspirin for a history of severe pre-eclampsia before 34 weeks of gestation, the administration of a daily prophylactic dose of enoxaparin in the antepartum period does not significantly reduce placenta-mediated complications.

Study reference	Kingdom 2011

Study Characteristics

Design
RCT

Objective
To conduct a pilot randomised controlled trial of unfractionated heparin (UFH) in women considered at high risk of placental insufficiency in the second trimester

Dates
March 2007 to May 2010

Country
Canada

Setting

Study reference	**Kingdom 2011**
	Mount Sinai Hospital, Toronto
Population Characteristics	Patient recruitment Eligible women (n=41) were approached in the high-risk Placenta Clinic, with 32 women providing written informed consent to participate. All women had prior pregnancy dating by ultrasound between 8 and 13 weeks' gestation using crown rump length measurements. Eligibility criteria (1) Singleton pregnancy between 18+0 and 23+6 weeks' gestation (2) Negative thrombophilia screen (3) Evidence of placental dysfunction through: (i) abnormal biochemical markers on first trimester, second trimester or integrated maternal serum screening tests for trisomy 21 and neural tube defects; (ii) sonographic evidence of abnormal placental morphology; and (iii) abnormal uterine artery Doppler waveforms. Abnormal placental biochemistry was defined as one or more of: PAPP-A<0.35 MoM, AFP>2.0 MoM, inhibin>3.0 MoM or total hCG>4.0 MoM. Allocation methods Eligible women who provided written informed consent were randomised using a central telephone randomisation service. The randomisation was computer generated, used balanced variable blocks, and was prepared by a statistician not involved with recruitment or clinical care. Blinding Women and their caregivers were not blinded to treatment allocation. Data collection All women were followed with clinic visits that included obstetric ultrasound examinations, either as full care and delivery at Mount Sinai Hospital (n=31) or shared care with delivery at the referral community hospital (n=1). A maternal complete blood count was checked 1 week after starting UFH, then every 2 weeks thereafter (and at 26 weeks in both arms at the time of the 50 g glucose challenge test. UFH was discontinued if the platelet count was < 100 x 10-9 L-1. Ultrasound examinations for fetal growth (HC, BPD, AC and FL) were performed as a minimum every 4 weeks, including the amniotic fluid index, umbilical artery Doppler PI, and biophysical profile. If the estimated fetal weight was <10th percentile, ultrasound examinations were increased to a minimum of weekly and incorporated Doppler studies of the fetal middle cerebral artery and ductus venosus. Serial ultrasound examinations also included imaging of the placenta and uterine artery Doppler, but were not used for clinical management. UFH was discontinued following the diagnosis of severe pre-eclampsia due to the associated risk of thrombocytopenia, but was otherwise continued until 34 weeks of gestation or delivery (whichever occurred first). Three maternal-fetal medicine specialists managed all 32 patients in a uniform manner Definition of stillbirth [or SGA <3rd centile] Intrauterine death defined as fetal death prior to birth and after trial entry (18 weeks' gestation earliest). However, 'previous stillbirth' was defined as >20 weeks for the analysis of baseline characteristics, and one stillbirth is reported and appears to be defined as >20 weeks. Birthweight centiles were ascertained based on sex and GA-specific Canadian data. Prevalence of stillbirth [SGA<3rd centile] NR Sample size N screened/invited = NR N eligible = 41 N enrolled = 32 (total), 16 (UFH), 16 (standard care) N excluded (with reason) = 0 N lost to follow-up = 0 N completed = 32 (total), 16 (UFH), 16 (standard care) N excluded from analysis = 0 N included in analysis = 32 (total), 16 (UFH), 16 (standard care) Maternal characteristics

	UFH, N=16	Standard care, N=16
Maternal age, years median (range)	33.5 (25 to 42)	35 (25 to 42)
GA at birth, weeks median (range)	34.3 (27.9 to 38.4 weeks)	35.6 (28.4 to 38.3 weeks)

Study reference	**Kingdom 2011**		
	Nulliparity, n (%)	6 (37.5)	6 (37.5)
	Smoking or other narcotic use, n (%)	1 (6.3)	0
	BMI kg/m^2, median (range)	28.3 (20.0 to 45.5)	27.0 (20.4 to 41.7)
	Comorbidities		
	Chronic renal disease, n (%)	0	1 (6.3)
	Chronic hypertension, n (%)	0	2 (12.5)
	Previous pregnancy complications		
	Previous stillbirth (>20 weeks)	4 (25)	4 (25)
	≥ 2 pregnancy losses <20 weeks	1 (6.3)	1 (6.3)
	Previous pre-eclampsia	1 (6.3)	2 (12.5)
	Previous SGA infant	1 (6.3)	6 (37.5)
Intervention	Intervention group (N=16) Women randomised to the heparin group received specific written instructions and instructions from a trial nurse (supervised by a haematologist) in the self-administration of subcutaneous unfractionated heparin (UFH). They injected 7500 IU, given as 0.3 cc of UFH subcutaneously using an insulin syringe twice a day from randomisation until 34 completed weeks of gestation or delivery (whichever occurred first). They were instructed to inject into the lower abdomen or lateral thighs on a rotating basis. The drug and syringes were paid for by the grant Placebo group (N=16) Women randomised to the standard care group received ongoing antenatal surveillance provided through the antenatal clinic, but were not administered medication. Low-dose aspirin was not subsequently taken by women in the heparin group. Eight women in the standard care group took aspirin until 34 completed weeks of gestation or delivery (whichever occurred first)		
Outcomes Measured	The primary outcome was mean maternal anxiety during pregnancy as measured by the Spielberger State-Trait Inventory Self Evaluation Questionnaire. Secondary outcomes were (i) Adverse outcomes for the infant: intrauterine fetal death (fetal death prior to birth and after trial entry); neonatal death (death of a live born infant prior to hospital discharge, and excluding lethal congenital anomalies); infant birth weight <10th centile for gestational age and infant sex; 5 min Apgar score <7; and a composite neonatal morbidity rate between groups; (ii) Adverse outcomes for the woman: ultrasound diagnosis of IUGR (defined as absent or reversed end diastolic flow in the umbilical arteries and an estimated fetal weight< 10th percentile); pre-term birth at <32 weeks' gestation; vaginal bleeding, pre-eclampsia, eclampsia or HELLP syndrome.		
Effectiveness of the Intervention	Stillbirth: UFH: 0/16 Standard care: 1/16 *No statistical analysis provided* Birthweight <3rd centile Standard care: 5/16 (31.3%) UFH: 4/16 (25%) RR: 1.25 (95% CI 0.41 to 3.82), p=1.000		
Authors' Conclusions	Prenatal screening of placental function may aid future trials of heparin by defining the type of underlying placental insufficiency so as to focus on those at most risk of infarction. The deployment of a strategy of "placental function screening" to reduce trial entry to women with a high positive predictive value of extreme pre-term delivery in association with placental developmental and vascular pathology, is an appropriate clinical strategy. The study design and findings, including placental pathology data, challenge the widely-prevailing view		

Study reference	Kingdom 2011
	that pregnant women should receive prophylactic heparin to improve perinatal outcomes based solely on clinical risk factors for placental insufficiency.

Study reference	Sharp 2018 (STRIDER)
Study Characteristics	Design RCT Objective To report the results of the first study from the Global Obstetric Network (GONet) initiative – a randomised trial hypothesising that sildenafil can delay the birth of the severely growth restricted fetus by at least 1 week by increasing blood flow to the placental bed with subsequent improvement in fetal growth and wellbeing in utero. Dates November 2014 to July 2016 Country UK Setting 19 fetal medicine units
Population Characteristics	Patient recruitment All women had a singleton pregnancy between 22 weeks and 0 days' gestation and 29 weeks and 6 days' gestation with a diagnosis of fetal growth restriction and had agreed to expectant management. For the purposes of the study, fetal growth restriction was defined as a fetus with abdominal circumference or estimated fetal weight below the tenth percentile using local charts and absent or reversed end diastolic flow in the umbilical artery on Doppler velocimetry. Women were excluded from the study if: • They were younger than 16 years old, • Had a known contraindication or allergy to sildenafil, • Had known or suspected significant chromosomal or structural anomaly, • Reported current cocaine use, or • Had a condition which was likely to require delivery within 72 h (such as severe pre-eclampsia). Allocation methods A web-based application was used to allocate treatment (1:1) with randomisation stratified by site and gestation (<26 weeks and 0 days and ≥26 weeks and 0 days). Randomisation lists were pre-generated using randomly permuted blocks of size two and four. Blinding Women, clinicians, and pharmacists were masked to the study drug. Data collection Gestational age was confirmed by first trimester ultrasound and, in each case, the diagnosis of severe early-onset fetal growth restriction was confirmed by a fetal medicine expert having excluded fetal anatomical abnormalities. In addition, a full history was taken, measurements of maternal cardiovascular parameters (pulse and blood pressure), fetal biometry, and Doppler velocimetry were taken, and maternal venepuncture for angiogenic biomarkers was carried out at randomisation. Women were followed up within 3–4 days and at weekly intervals thereafter, or earlier when clinically indicated. The rest of clinical care was at the discretion of the local fetal medicine experts and included regular ultrasound assessment of growth and Doppler blood flow and antenatal cardiotocography. Data on pregnancy outcome were collected prospectively from the clinical maternity notes and entered them on a secure electronic database. Definition of stillbirth [or SGA <3rd centile] NR; data was reported for fetal death <26 weeks' gestation and ≥26 weeks' gestation. Prevalence of stillbirth [SGA <3rd centile]

Study reference	**Sharp 2018 (STRIDER)**

NR

Sample size
N screened/invited = 149
N eligible = NR
N enrolled = 135
N excluded (with reason) = 0
N lost to follow-up = 0
N completed = 135
N excluded from analysis = 0
N included in analysis = 135 (sildenafil n=70, placebo n=65)

Maternal characteristics

	Sildenafil group, N=70	Placebo group, N=65
Maternal age, years median (range)	29 (26 to 34)	33 (28 to 36)
Nulliparity, n (%)	35 (50%)	25 (38%)
Current smoker, n (%)	12 (17%)	2 (3%)
BMI kg/m^2, median (range)	25 (23 to 32)	26 (23 to 31)
Comorbidities		
Pre-eclampsia	13 (19%)	11 (17%)
Gestational hypertension	12 (17%)	23 (35%)
Current antihypertensive treatment	25 (36%)	27 (42%)
Gestational diabetes	2 (3%)	3 (5%)

Intervention

All participants received oral medication, sildenafil 25 mg or placebo, three times a day. Medication was dispensed in 10-day supplies with a new supply being provided every week to ensure there was no period when medication was missed. We used pharmacy logs to monitor adherence. We stopped treatment at 32 weeks and 0 days or delivery, whichever came first.

Outcomes Measured

The primary efficacy outcome was the time from randomisation to delivery, measured in days.

Secondary outcomes included livebirths, fetal and neonatal deaths, birthweight, neonatal morbidity (any intraventricular haemorrhage, oxygen dependency at 28 days and 36 weeks corrected gestational age, necrotising enterocolitis, or retinopathy of prematurity), use of surfactant, ventilator dependency, admission to neonatal intensive care unit, time to newborn discharge, and maternal side-effects.

Adverse events and adherence were assessed and recorded at weekly clinical visits from recruitment to delivery.

Effectiveness of the Intervention

	Sildenafil group, N=70	Placebo group, N=65	Relative risk (95% CI)	P value
Livebirths	49 (70%)	43 (66%)	1.06 (0.84 to 1.33)	0.62
<26 weeks' gestation	22 (31%)	15 (23%)	1.28 (0.80 to 2.06)	0.31
≥26 weeks' gestation	27 (39%)	28 (43%)	0.96 (0.83 to 1.12)	0.59
Fetal death	21 (30%)	22 (34%)	0.89 (0.54 to 1.45)	0.64
<26 weeks' gestation	18 (26%)	20 (31%)	0.79 (0.50 to 1.23)	0.31
≥26 weeks' gestation	3 (4%)	2 (3%)	1.50 (0.27 to 8.34)	0.64

Authors' Conclusions

Sildenafil did not prolong pregnancy or improve pregnancy outcomes in severe early-onset fetal growth restriction and therefore it should not be prescribed for this indication outside of research studies with explicit participants' consent.

Study reference	**Subtil 2003a, Subtil 2003b**
Study Characteristics	Design RCT, but only a single-arm relevant to the review Objective To assess the effectiveness of a pre-eclampsia prevention strategy based on routine uterine artery Doppler flow velocity waveform examination during the second trimester of pregnancy, followed by a prescription for 100 mg aspirin in the case of abnormal Doppler findings Dates Enrolment 20 December 1993 to 24 June 1998 Country France and Belgium Setting 12 maternity centres
Population Characteristics	Patient recruitment Women were invited to participate in this study during a prenatal visit to a participating centre. The invitation was followed by an explanation, consent and random allocation. This trial was part of a larger study evaluating routine prescription of low dose aspirin in nulliparous women (Subtil 2003b). The parallel trial presented here was ran in 12 of 29 centres that had Doppler imaging facilities. Eligibility criteria Women were required to be nulliparous (no previous delivery before ≥22 weeks), between 14 and 20^{+6} weeks, with no history of hypertension and no clear indications for or contraindications to the prescription of aspirin or another anticoagulant during the pregnancy. Multiple pregnancy was not an exclusion criterium. Allocation methods Women within the main RCT were randomised into 3 groups: Half underwent a uteroplacental artery Doppler at the same time as the second-trimester anatomical ultrasound (22–24 weeks), with low dose aspirin (100 mg) prescribed only if the findings were abnormal. This treatment was to begin the day after the uteroplacental artery Doppler and continue through 36 weeks. The other half was further divided randomly into two groups (i.e. daily treatment with low dose aspirin [100 mg] or placebo until the end of the 34th week). Only data for the women with abnormal Doppler results and prescribed aspirin are presented here. A randomisation list was computer-generated, balanced in blocks of 16 and stratified by centre. Each patient was randomly allocated to a group immediately after inclusion by connection to an always available server. Blinding It is assumed that it was not possible to blind women to the intervention (ultrasound and aspirin versus no aspirin). Data collection A seven-page case report form contained the data, recorded prospectively, for each patient. Definition of stillbirth [or SGA <3rd centile] Antenatal death among pregnancies with births ≥22 weeks and excluding terminations. SGA<3rd centile defined as birthweight ≤3rd centile for gestational age according to standard curves used in France Prevalence of stillbirth [SGA<3rd centile] NR Sample size N screened/invited = 3,118 N eligible = NR

Study reference	**Subtil 2003a, Subtil 2003b**
	N enrolled = 2,491 (randomised): 1,253 (Doppler group), 239 (Abnormal Doppler results and aspirin prescription) N excluded (with reason) =NR N lost to follow-up = 9 (from all women allocated to the Doppler group) N completed = 1,244 (from all women allocated to the Doppler group) N excluded from analysis = NR N included in analysis = 239 in the abnormal Doppler group, Maternal characteristics *Reported for women where Doppler examination had been performed (N=1,175)* • Age: 24.7 years (SD 4.4) • Gestational age at inclusion =16.8 weeks (SD 2.1) • Parity/gravida: all nulliparous • BMI ≥25 kg/m^2, n=192 (17.8%) • Comorbidities: NR • Previous pregnancy complications: NR • Multiple pregnancies: n=18 (1.5%) *Reported for women randomised to the Doppler group (N=1,253)* • Smoking: 339 (27.1%)
Intervention	Intervention group (N=239) Patients allocated to the Doppler group were given an appointment date between 22 and 24 weeks, for a uteroplacental artery Doppler, which in most cases coincided with the fifth-month ultrasound). The examinations were performed by physicians experienced in current obstetric ultrasound techniques, with training meetings for these physicians organised at the beginning of the study. The uteroplacental artery Doppler was considered abnormal if a notch was observed on either side or if the mean resistance index (S–D/S) of the two uterine arteries (right and left) was equal to or greater than 0.61. The aspirin treatment boxes prescribed in the case of abnormal Doppler findings were manufactured specifically for the study. Aspirin was prescribed through the end of 36 weeks.
Outcomes Measured	The principal outcome was pre-eclampsia, the other outcomes were pregnancy-related hypertension, very low or low birthweight for gestational age (birthweight ≤3rd or ≤10th centile of the standard curves used in France), HELLP syndrome, placental abruption or a caesarean delivery because of fetal indication (uncompensated maternal hypertension, suspected IUGR, meconium-stained amniotic fluid or placental abruption).
Effectiveness of the Intervention	Antenatal death (among births ≥22 weeks and excluding terminations): 0/239 Birthweight ≤3rd centile: • 11/239 (4.6%) in women with abnormal Doppler • 12/947 (1.3%) in women with normal Doppler • RR 3.6 (95% CI 1.6 to 8.1) (lower risk of incidence among women with normal Doppler)
Authors' Conclusions	A prevention strategy based on a routine uteroplacental artery Doppler in nulliparous patients followed by the prescription of aspirin did not reduce the frequency of either pre-eclampsia or other complications of pregnancy. These results do not justify a recommendation for a routine Doppler, since the aspirin prescription in the case of abnormal findings did not reduce the risk of these complications.

Study reference	**von Dadelszen 2011**
Study Characteristics	Design Retrospective case-control Objective To report the use of Sildenafil citrate as an innovative therapy in the management of severe early-onset IUGR, and compare the outcomes of Sildenafil-treated pregnancies with similar pregnancies that remained Sildenafil-naïve Dates 2004 to 2009 Country

Study reference	**von Dadelszen 2011**
	Canada Setting Unclear, likely University of British Columbia Women's Hospital
Population Characteristics	Patient recruitment and eligibility criteria Women were offered Sildenafil if their pregnancy was complicated by severe early-onset IUGR (ultrasound estimation of the fetal AC of <5th percentile) with an estimated probability of 'intact' survival of <50% excluding known aneuploidy, anomaly, syndrome or congenital infection, or if there was a plan to terminate the pregnancy. Contemporaneous Sildenafil-naive controls were identified within the Diagnostic Ambulatory Program Ultrasound Database at British Columbia Women's Hospital. Matching criteria were as follows: (i) maternal age (±5 years); (ii) gestational age at eligibility (±14 days); (iii) parity (0, 1, ≥2); and (iv) eligibility to be offered Sildenafil. Allocation methods Case-control, appears to have been conducted as an open-label study. Blinding Outcomes assessors (sonographers) were not blind to the treatment. Data collection Management was similar between the two groups, and included increased fetal (umbilical artery and ductus venosus Doppler indices, fetal biometry; amniotic fluid index, deepest vertical amniotic fluid pocket and nonstress tests) and maternal [measurement of blood pressure, proteinuria, pulse oximetry, complete blood count, creatinine, uric acid, aspartate transaminase, bilirubin and albumin] surveillance. Fetal surveillance occurred at least as frequently as every 6 to 8 days for outpatients and at least twice weekly for inpatients. Maternal tests were repeated at least every 14 days in outpatients and every week in inpatients. Women in the Sildenafil-treated group were also monitored for adverse side-effects, such as flushing, light-headedness and visual disturbance. Fetuses and infants were followed until 28 days of life or primary hospital discharge, whichever was later. Women were followed until primary hospital discharge. Definition of stillbirth [or SGA <3rd centile] NR Prevalence of stillbirth [SGA<3rd centile] NR Sample size N screened/invited =NR N eligible =NR N enrolled =NR N excluded (with reason) =NR N lost to follow-up =NR N completed =NR N excluded from analysis =NR N included in analysis =NR Maternal characteristics

Characteristic	Sildenafil naïve N=17	Sildenafil-treated N=10
Maternal age, median (IQR)	33 (28 to 36.5) years	34 (25 to 40.5) years
GA at delivery, median (IQR)	181 (166 to 208) days	190 (179 to 230) days
Nulliparous, n (%)	8 (47)	5 (50)
Smoking status	NR	NR
Comorbidities	NR	NR
Previous pregnancy complications	NR	NR

Intervention	Sildenafil citrate (N=10) 25 mg *per os*, thrice daily until delivery Sildenafil-naïve (N=17)

Study reference	von Dadelszen 2011
	No sildenafil treatment; no details regarding other treatment reported
Outcomes Measured	The primary outcome for this analysis was the proportion of women in each group for whom fetal AC growth velocity increased post-eligibility. Secondary outcomes for this study were live birth, neonatal survival to hospital discharge, intact survival, combined non-CNS severe morbidity and adverse maternal side-effects of medication.
Effectiveness of the Intervention	Sildenafil-treated: 1 stillbirth occurred within 48 hours of starting Sildenafil (reversed end-diastolic flow on day of prescription), 1 stillbirth occurred during in utero transfer to USA (NICU occupancy), 1 permissive stillbirth (estimated fetal weight of <500 g) Sildenafil-naïve: 5 stillbirths as a result of late termination, 6 permissive stillbirths (estimated fetal weight of <500 g) No statistical analysis comparing frequency of stillbirths was provided
Authors' Conclusions	The data suggests that Sildenafil treatment may offer a new opportunity to improve perinatal outcomes for women whose pregnancies are complicated by severe early-onset IUGR, thought they are not sufficiently robust to guide decision-making about the use of sildenafil in pregnancies complicated by early-onset IUGR, until a randomised controlled trial can be conducted.

Question 4

Study reference	Boers 2010 (DIGITAT)
Study Characteristics	Design RCT Objective To compare the effect of induction of labour with a policy of expectant monitoring for intrauterine growth restriction near term Dates November 2004 to November 2008 (recruitment) Country The Netherlands Setting Fifty-two maternity hospitals
Population Characteristics	Patient recruitment and eligibility criteria The trial included pregnant women between 36+0 and 41+0 weeks' gestation who had a singleton fetus in cephalic presentation, suspected IUGR and who were under specialised obstetric care. Fetuses were included regardless of Doppler flow velocity. IUGR was defined as AC<10th centile, EFW<10th centile, growth curve flattening in the third trimester, or all three factors. It appears that population charts were used to estimate fetal size (and therefore IUGR), as the authors note that customised growth centile charts are rarely applied in the Netherlands and were not used in the study. Exclusion criteria were previous CS, diabetes mellitus or gestational diabetes requiring insulin therapy, renal failure, HIV, PROM, severe pre-eclampsia, HELLP syndrome, or a fetus with aneuploidy or congenital abnormalities suspected on ultrasound. Fetuses with decreased or absent movements, and those with abnormal heart rate tracings, were also excluded. Allocation methods Women were randomly allocated to either induction or expectant monitoring in a 1:1 ratio using varied sized block randomisation with stratification for centre and parity (nulliparous or parous women). Blinding Participants, obstetricians or outcomes assessors were not blinded Data collection

Study reference	Boers 2010 (DIGITAT)

Outcome data was collected prospectively by obstetricians, nurses and midwives associated with the Dutch Obstetric Consortium.

Definition of stillbirth [or SGA<3rd centile]
Stillbirth was not specifically defined. Birthweight<3rd centile was calculated according to Dutch fetal growth charts relating weight to gestational age

Prevalence of stillbirth [SGA<3rd centile]
NR

Sample size
N screened/invited = NR
N eligible = 1116
N enrolled = 650
N excluded (with reason) = none
N lost to follow-up (reported as "unknown method of delivery"= 1 (expectant monitoring), 1 (induction of labour)
N completed = NR
N excluded from analysis = 0
N included in analysis = 321 (induction of labour), 329 (expectant monitoring)

Demographics

Characteristic	Induction of labour (n=321)	Expectant monitoring (n=329)
Maternal age, median years (IQR)	27 (23 to 31)	27 (23 to 31)
Smoking during pregnancy, n (%)	138 (46.9) (n=294)	127 (40.8) (n=311)
BMI at study entry, kg/m^2 median (IQR)	22 (20 to 25) (n=275)	22 (20 to 26) (n=295)
Gravidity	NR	NR
Nulliparous, n (%)	182 (56.7)	201 (61.1)
Gestational age, days median (IQR)	263 (258 to 269)	263 (258 to 270)
Gestational hypertension, n (%)	9 (2.8)	19 (5.8)
Pre-eclampsia, n (%)	18 (5.6)	27 (8.2)
Prior pregnancy complications	NR	NR

Details of Planned Delivery Method

Induction of labour, N=321
Participants allocated to the induction of labour group were induced within 48 hours of randomisation. If the Bishop score at randomisation was greater than 6, labour was induced with amniotomy and, if necessary, augmented with oxytocin. Otherwise cervical ripening was performed with intracervical or intravaginal prostaglandin (E1 or E2 analogue, repeated once after six hours) or a Foley balloon catheter filled with 30 mL sodium chloride.

Time between randomisation and onset of labour was 0.9 days (IQR 0.7–1.7)

Details of Comparator

Expectant monitoring N=329
Participants allocated to the expectant monitoring group were monitored until the onset of spontaneous labour with daily fetal movement counts and twice weekly heart rate tracings, ultrasound examination, maternal blood pressure measurement, assessment of proteinuria, laboratory tests of liver and kidney function, and full blood count. Women were monitored as either an outpatient or an inpatient, according to local protocol. In the expectant monitoring group, induction of labour or planned caesarean section was performed for obstetrical indications—such as suboptimal fetal heart rate tracings, prolonged rupture of membranes, or postmaturity between T+7 and T+14 days – at the obstetrician's discretion.

Time between randomisation and onset of labour was 10.4 days (IQR 5.6–16.0)

Outcomes Measured

The primary outcome was a composite measure of adverse neonatal outcome, defined as death before hospital discharge, 5 min Apgar score <7, umbilical artery pH <7.05, or admission to NICU.

Secondary outcomes were delivery by caesarean section, instrumental vaginal delivery, length of stay in

Study reference	**Boers 2010 (DIGITAT)**
	NICU or neonatal ward, length of stay in the maternal hospital, and maternal morbidity.

Effectiveness of Planned Delivery

Outcomes	Induction of labour (n=321)	Expectant monitoring (n=329)	Difference in %
Stillbirth	**0**	**0**	**NA**
Birthweight<3rd centile	**40 (12.5)**	**100 (30.6)**	**−18.1 (−24.3 to −12.0), p<0.001**
Neonatal deaths	0	0	NA
Fetal deaths	0	0	NA
Composite adverse neonatal outcome	17 (5.3)	20 (6.1)	−0.8 (−4.3 to 2.8)
Onset of labour, n (%)			
Spontaneous	12 (3.7)	151 (46.0)	−42.3 (−48.1 to −36.5)
Planned Caesarean section	2 (0.6)	11 (3.3)	−2.7 (−4.9 to −0.6)
Induction	306 (95.6)	166 (50.6)	45.0 (39.2 to 50.9)
Apgar score after 5 minutes <7	7 (2.2)	2 (0.6)	1. 6 (−0.2 to 3.4)
Admission to intensive care	9 (2.8)	13 (4.0)	−1.2 (−4.0 to 1.6)
Length of stay in the neonatal intensive care unit (days)	9 (6 to 14)	13 (6 to 22)	P=0.2

Authors' Conclusions

A significantly higher number of babies with a birthweight<3rd centile in the expectant monitoring group suggests that a substantial number did not continue to grow along their own expected growth curves. This should be a compelling reason for induction and merits further investigation.

Induction was not associated with any increase in operative and instrumental delivery rates, thus, it is rational to choose induction in patients with intrauterine growth restriction near term to prevent possible neonatal morbidity and stillbirth.

Study reference	**Rabinovich 2018**
Study Characteristics	Design Retrospective cohort study Objective To examine the fetal and neonatal morbidity and mortality rates associated with induction of labor versus expectant management in a unique cohort of fetuses with isolated growth restriction between $34^{0/7}$ and $38^{6/7}$ weeks; and (2) determine the optimal gestational age in which delivery of such fetuses will be associated with the lowest rate of fetal, neonatal and maternal complications Dates NR Country Israel Setting Soroka University Medical Center (a tertiary medical facility)
Population Characteristics	Patient recruitment Eligible patients with isolated FGR between $34^{0/7}$ and $38^{6/7}$ weeks of gestation were identified from the hospital's electronic database. Exclusion criteria included multiple pregnancies, fetal chromosomal and congenital abnormalities, poly or oligohydramnion, placental abruption, chorioamnionitis, and preeclampsia. Allocation methods NA – retrospective study

Study reference	**Rabinovich 2018**

Blinding
NA – retrospective study

Data collection
Data from the hospital's electronic database was extracted on maternal demographics, obstetrical history, labour and delivery events, immediate neonatal outcome and detailed information of neonatal hospitalisation.

Definition of stillbirth [or SGA<3rd centile]
NR

Prevalence of stillbirth [SGA<3rd centile]
NR

Sample size
N screened/invited = 273,940
N eligible = 2,232
N excluded (with reason) = NR
N randomised = NA (retrospective study)
N lost to follow-up = NA (retrospective study)
N completed = NA (retrospective study)
N excluded from analysis = NR
N included in analysis = 1,428 (labour induction), 804 (expectant management)

Demographics

Characteristic	Labour induction (N=1,428)		Expectant management (n=804)	P value
	Elective CS (n=348)	Other induction (n=1,080)		
Maternal age, mean years (SD)	30.5 (6.07)	26.6 (5.31)	27.138 (6.03)	<0.001
Smoking during pregnancy, n (%)	NR	NR	NR	NA
BMI ≥30 kg/m^2, n (%)	NR	NR	NR	NA
Gravidity, median	3	2	3	<0.001
Parity, median	3	2	2	<0.001
Fertility treatments, %	6	4.5	3.9	<0.001
Gestational age in weeks, mean (SD)	36.5 (1.12)	36.9 (1.04)	36.8 (1.22)	<0.001
Comorbidities	NR	NR	NR	NA
Prior pregnancy complications	NR	NR	NR	NA

Details of Planned Delivery Method	Induction of labour was achieved either by cervical balloon catheters, prostaglandin E2, oxytocin or artificial rupture of membranes
Details of Comparator	The decision of expectant management was based on the daily assessment of fetal heart tracing, biophysical profile, sequential Doppler studies and repeated assessment of fetal growth. When abnormal results were encountered, the attending physicians decided whether to deliver the fetus, chose the route of delivery, and how urgent the intervention was needed. Delivery by a planned CS was preferred in pregnancies with abnormal fetal lie and presentation, EFW <3rd centile and when the attending physician estimated that the fetus would eventually require an emergency CS
Outcomes Measured	Not pre-specified, but assumed to be fetal and neonatal mortality. A composite neonatal outcome, defined as neonatal death, birth asphyxia, neonatal sepsis and prematurity complication, was pre-specified

Effectiveness of Planned Delivery

	Labour induction (N=1,428)		Expectant management (N=804)	P value
	Elective CS (n=348)	Other induction (n=1,080)		
Stillbirth, %	0.3	0.6	1.5	0.042

Study reference **Rabinovich 2018**

Apgar at 1 minute <5, %	2.6	1.6	5.8	<0.001
Apgar at 5 minutes <7, %	0.9	0.3	1.9	0.002
Fetal distress, %	0	2.5	6.5	<0.001
Neonatal death, %	0.9	0.2	2	<0.001
Neonatal hospitalisation length, days	5	3	3	<0.001
Neonatal hypoglycaemia, %	4.5	5.5	7	0.212
Neonatal sepsis, %	0.3	0.3	0.6	ns
Prematurity complications, %	3	1	2	0.032
Low birth weight <2500g, %	82.5	78.6	80.1	0.283
Very low birth weight <1500g, %	4.6	0.5	3.2	<0.001

	Perinatal mortality		Crude OR	Adjusted OR	P value
	Yes (n=40)	No (n=2,192)			
Induction of labour	30%	64.4%	0.23 (0.12 to 0.46)	0.29 (0.14 to 0.59)	0.001

	Late pre-term (34 0/7 to 36 6/7)			Early term (37 0/7 to 38 6/7)		
	Induction (N=477)	Expectant management (N=292)	P value	Induction (N=951)	Expectant management (N=512)	P value
Stillbirth, %	0.8	3.1	0.001	0.3	0.6	0.428
Initial Apgar score <5, %	2.8	9.4	<0.001	1.4	3.8	0.003
Subsequent Apgar score <7, %	0.6	3.6	0.006	0.3	1	0.174
Fetal distress, %	1.9	9.6	<0.001	1.9	4.7	0.002
Neonatal death, %	0.8	4.5	0.039	0.1	0.6	0.126
Neonatal hospitalisation length, days	6	6	0.532	3	2	0.001
Neonatal hypoglycaemia, %	6.7	7.8	0.617	4.5	6.7	0.088
Neonatal sepsis, %	0.2	1.2	0.123	0.3	0.2	1

Study reference	**Rabinovich 2018**						
	Prematurity complications, %	3	5.3	0.136	0.7	0.2	0.434
	Low birth weight <2500g, %	93.9	95.5	0.335	72.3	71.3	0.668
	Very low birth weight <1500g, %	4.2	7.2	0.072	0.1	1	0.022
Authors' Conclusions	The results are in concordance with previous publications indicating no substantial advantage to induction of labour over expectant management at early term and specifying 37th to 38th week as the optimum time for induction of labour in pregnancies complicated by FGR. Statistically significant lower rates of stillbirth, neonatal death, fetal distress and low 1 and 5 min Apgar scores were shown with labour induction compared with expectant management at the late preterm gestational period and there was no particular perinatal or neonatal advantage to induction by a CS.						

Study reference	**Walker 2016**
Study Characteristics	Design RCT Objective To test the hypothesis that induction of labour at 39 weeks of gestation would reduce the rate of Caesarean delivery among nulliparous women of advanced maternal age Dates Participants were recruited between August 2012 and March 2015 Country UK Setting Thirty-eight NHS hospitals and one Primary Care Trust
Population Characteristics	Patient recruitment Eligible women were nulliparous, were to be ≥35 years of age on their expected due date, and had a singleton live fetus in a cephalic presentation. Women were ineligible to participate if their pregnancy was complicated by a known fetal congenital abnormality that would lead to neonatal death or if they had any contraindications to labour (e.g., evidence of fetal compromise), vaginal delivery (e.g., placenta previa), or expectant management (e.g., gestational diabetes). Women who had undergone a myomectomy, who had not undergone US examination (for estimation of gestational age) before 22 weeks of gestation, or who had undergone in vitro fertilisation with the use of donor eggs were also excluded. Allocation methods Women were randomised at 36 weeks 0 days to 39 weeks 6 days of gestation, in a 1:1 ratio according to a computer-generated code with the use of random permuted blocks of randomly varying size generated by the Nottingham Clinical Trials Unit. The randomisation was stratified into three categories according to trial centre and maternal age (35 to 37 years of age, 38 to 39 years of age, and 40 years of age or older). Blinding There was no blinding of the treatment Data collection Post-randomisation outcome data was collected at delivery by a designated individual within a local centre from hospital notes, immediately following discharge. Definition of stillbirth [or SGA<3rd centile] A baby delivered with no signs of life after 24 wk of gestation Prevalence of stillbirth [SGA<3rd centile]

Study reference	**Walker 2016**
	NR Sample size N screened/invited = 6,455 N eligible = 4,542 N excluded (with reason) = 1,913 (did not meet inclusion criteria), 3,923 (declined participation) N randomised = 619 (total), 305 (labour induction), 314 (expectant management) N lost to follow-up = 0 N completed = 265 (adhered to labour induction), 297 (adhered to expectant management) N excluded from analysis = 2 in labour induction: 1 withdrew consent and 1 had no available outcome data N included in analysis = 304 (labour induction), 314 (expectant management)
Details of Planned Delivery Method	Induction of labour, N=304 Induction of labour was to be performed between 39 weeks 0 days and 39 weeks 6 days of gestation, local policies for induction of labour were followed (most participating units used prostaglandin ripening followed, if necessary, by amniotomy and oxytocin infusion).
Details of Comparator	Expectant management, N=314 Women randomised to the expectant management group could undergo induction between 41 weeks 0 days and 42 weeks 0 days of gestation (i.e., 7 to 14 days after the due date), with the exact time determined by their preference and the physician's usual practice. No additional monitoring before 42 weeks 0 days of gestation was offered unless it was the physician's usual practice. If the woman declined to undergo induction at 42 weeks of gestation, she could undergo scanning to determine fetal growth and amniotic-fluid volume and daily or every other day cardiotocographic monitoring according to the physician's usual practice. If induction was indicated in this group for any reason, staff were encouraged to use the same induction protocol as would have been used for the labour induction arm.
Outcomes Measured	The primary outcome was Caesarean delivery. The secondary maternal outcomes were the method of delivery other than Caesarean section, the onset of labour, the indication for induction of labour, the method of labour induction, the indication for Caesarean section, intrapartum complications, and postpartum complications (e.g., systemic infection or the need for a blood transfusion).

Demographics

Characteristic	Labour induction (N=305)	Expectant management (n=314)
Maternal age at expected date of delivery, mean years (SD), range	37 (2.2), 35 to 45	37 (2.2), 35 to 44
Smoking during pregnancy, n (%)	9 (3)	5 (2)
BMI ≥30 kg/m^2, n (%)	85 (28)	83 (26)
Gravidity	NR	NR
Nulliparous, n (%)	All (per inclusion criteria)	All (per inclusion criteria)
Gestational age, days median (IQR)	NR	NR
History of hypertension, n (%)	4 (1)	3 (1)
History of any disease, n (%)	48 (16)	50 (16)
History of renal disease, n (%)	0	1 (<1)
History of other condition, n (%)	46 (15)	46 (15)
Prior pregnancy complications	NR	NR

Study reference	**Walker 2016**
	The secondary neonatal outcomes were live birth or stillbirth, birth weight, admission to NICU, birth trauma, and two composite outcomes for serious neonatal complications (direct trauma and hypoxia). Other secondary outcomes included the mothers' expectations and experience of childbirth, as measured with the use of the Childbirth Experience Questionnaire, which was sent to the women one month after the birth The trial was not designed or powered to assess the effects of labour induction on stillbirth.
Effectiveness of Planned Delivery	**No stillbirths** occurred in either trial arm. SGA<3rd centile was not reported. There was no difference between the rate of Caesarean sections or maternal and neonatal outcomes between the groups.

Outcome	Induction Group (N=304)	Expectant Management Group (N=314)	Relative Risk (95% CI)	P Value
Stillbirth	0	0	-	-
Birth weight <2500g	4	6	0.68 (0.19 to 2.4)	0.56
Apgar score at 5 min, n				
<4	0	1		
4 to 7	11	11	1.04 (0.40 to 2.69)	0.94
Admission to NICU for >4 days, n	6	7	0.88 (0.26 to 3.06)	0.85
Complication, n				
Composite outcome				
Hypoxia	2	2	1.03 (0.14 to 7.50)	0.98
Hypotonia ≥2 hr	1	0		
Required intervention, n				
Tube feeding >4 days	0	2		
Intubation and ventilation >24 hours	1	2	0.51 (0.45 to 5.82)	0.59
Cooling	1	2	0.52 (0.47 to 5.68)	0.59
Oxygen	9	7	1.32 (0.58 to 2.99)	0.50
CPAP	4	4	1.02 (0.22 to 4.86)	0.97

Authors' Conclusions	The trial did not address the effect of labour induction at 39 weeks on preventing stillbirth. However, the results provide support for performing a larger trial of labour induction on stillbirth rates in women ≥35 years of age.

Appraisal for quality and risk of bias

Quality assessments of included studies are reported below.

Table 24. Quality assessment of studies relevant to question 1

Question	Akolekar 2016a	Akolekar 2016b, Aupont 2016	Bakalis 2015	Chaiworapongsa 2017	Chaiworapongsa 2013a (cohort study)	Chaiworapongsa 2013b (case-control study)	Dugoff 2004	Dugoff 2005	Dugoff 2008	Familiari 2016
PARTICIPANT SELECTION										
Was a consecutive or random sample of pregnancies enrolled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a case-control design avoided?	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Did the study avoid inappropriate exclusions?	Yes	Yes	Yes	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes
Could the selection of pregnancies have introduced bias?	No	No	No	Yes	Unclear	Yes	Unclear	No	No	Yes
Is there concern that the included pregnancies do not match the review question?	Low	Low	Low	Low	High	High	Low	Low	Low	Low
INDEX TESTS										
Were the index test results interpreted without knowledge of the reference standard?	No	No	No	Yes	No	No	Unclear	Unclear	Yes	Yes
If a threshold was used, was it pre-specified?	No	No	No	Yes	Unclear	Unclear	Yes	Yes	No	No
Could the conduct or	Yes	Yes	Yes	No	Yes	Unclear	No	No	Yes	Yes

Question	Akolekar 2016a	Akolekar 2016b, Aupont 2016	Bakalis 2015	Chaiworapongsa 2017	Chaiworapongsa 2013a (cohort study)	Chaiworapongsa 2013b (case-control study)	Dugoff 2004	Dugoff 2005	Dugoff 2008	Familiari 2016
interpretation of the index test have introduced bias?										
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
REFERENCE STANDARD										
Is the reference standard likely to correctly classify the test condition?	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	Unclear	No	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear	Unclear	High	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	No
Is there concern that the target condition as defined by the reference standard does not match the review question?	Low	Low	Low	Low	High	High	Low	Low	Low	Low
PARTICIPANT FLOW										
Did all participants receive a reference standard?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes

Question	Akolekar 2016a	Akolekar 2016b, Aupont 2016	Bakalis 2015	Chaiworapongsa 2017	Chaiworapongsa 2013a (cohort study)	Chaiworapongsa 2013b (case-control study)	Dugoff 2004	Dugoff 2005	Dugoff 2008	Familiari 2016
Did participants receive the same reference standard?	Yes	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Unclear	Yes
Were all pregnancies included in the analysis?	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes
Could the participant flow have introduced bias?	No	No	No	Yes	No	Yes	No	No	Yes	No

Question	Hemming 2011	Marttala 2010	Mastrodima 2016	Odibo 2012	Poon 2013	Singh 2012	Smith 2014
PARTICIPANT SELECTION							
Was a consecutive or random sample of pregnancies enrolled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a case-control design avoided?	Yes	Yes	Yes	Yes	Yes	Yes	No
Did the study avoid inappropriate exclusions?	Yes	Yes	Yes	Yes	Yes	No	No
Could the selection of pregnancies have introduced bias?	Yes	Yes	No	Yes	Yes	Yes	Yes
Is there concern that the included pregnancies do not match the review question?	Low	Low	Low	Low	Low	High	Low

Question	Hemming 2011	Marttala 2010	Mastrodima 2016	Odibo 2012	Poon 2013	Singh 2012	Smith 2014
INDEX TESTS							
Were the index test results interpreted without knowledge of the reference standard?	Unclear	Unclear	No	Unclear	No	No	No
If a threshold was used, was it pre-specified?	Yes	Yes	No	Yes	Yes	Yes	Yes
Could the conduct or interpretation of the index test have introduced bias?	No	No	Yes	No	No	No	No
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Low	Low	Low	Low	Low	Low	High
REFERENCE STANDARD							
Is the reference standard likely to correctly classify the test condition?	Yes	Unclear	Unclear	Unclear	Yes	Yes	Yes
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	Unclear	Unclear	No	Yes	Yes	Yes
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear	Unclear	Unclear	Yes	No	No	Unclear
Is there concern that the target condition as defined by the reference standard does not match the	Low	High	Low	High	Low	High	Low

Question	Hemming 2011	Marttala 2010	Mastrodima 2016	Odibo 2012	Poon 2013	Singh 2012	Smith 2014
review question?							
PARTICIPANT FLOW							
Did all participants receive a reference standard?	Unclear	Unclear	Yes	Yes	No	Unclear	Unclear
Did participants receive the same reference standard?	Unclear	Unclear	Yes	Unclear	No	Unclear	Unclear
Were all pregnancies included in the analysis?	Yes	Yes	Yes	Yes	Yes	No	No
Could the participant flow have introduced bias?	No	No	No	Unclear	Yes	Yes	Yes

Question	Smith 2007a	Smith 2007b	Sutan 2010	Tancrede 2015	Trudell 2015	Trudell 2017	Valino 2016a	Valino 2016b	Yerlikaya 2016
PARTICIPANT SELECTION									
Was a consecutive or random sample of pregnancies enrolled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a case-control design avoided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Did the study avoid inappropriate exclusions?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Could the selection of pregnancies have introduced bias?	No	Yes	Yes	No	Yes	Yes	No	No	No
Is there concern that the included pregnancies do not match the review question?	Low	Low	Low	Low	Low	Low	Low	Low	Low

Question	Smith 2007a	Smith 2007b	Sutan 2010	Tancrede 2015	Trudell 2015	Trudell 2017	Valino 2016a	Valino 2016b	Yerlikaya 2016
INDEX TESTS									
Were the index test results interpreted without knowledge of the reference standard?	No	Unclear	No	Yes	No	No	No	No	No
If a threshold was used, was it pre-specified?	Yes	Yes	No	Yes	Yes	No	No	No	No
Could the conduct or interpretation of the index test have introduced bias?	No	No	Yes	No	No	Yes	Yes	Yes	Yes
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Low	Low	Low	Low	Low	Low	Low	Low	Low
REFERENCE STANDARD									
Is the reference standard likely to correctly classify the test condition?	Yes	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Unclear
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	Unclear	No	Yes	Yes	Unclear	No	No	Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear	Unclear	Unclear	No	No	Unclear	Yes	Yes	Unclear
Is there concern that the target condition as defined by the reference standard	Low	Low	High	High	High	High	Unclear	Unclear	Low

Question	Smith 2007a	Smith 2007b	Sutan 2010	Tancrede 2015	Trudell 2015	Trudell 2017	Valino 2016a	Valino 2016b	Yerlikaya 2016
does not match the review question?									
PARTICIPANT FLOW									
Did all participants receive a reference standard?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Did participants receive the same reference standard?	Yes	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Yes
Were all pregnancies included in the analysis?	Yes	Unclear	Unclear	Yes	No	No	No	No	Yes
Could the participant flow have introduced bias?	No	Unclear	Unclear	No	No	Yes	Yes	Yes	No

SLRs

Question	Conde-Agudelo 2015
Was an 'a priori' design provided?	Yes
Was there duplicate study selection and data extraction?	Yes
Was a comprehensive literature search performed?	Yes
Was the status of publication (i.e. grey literature) used as an inclusion criterion?	Yes
Was a list of studies (included and excluded) provided?	No
Were the characteristics of the included studies provided?	Yes
Was the scientific quality of the included studies assessed and documented?	No
Was the scientific quality of the included studies used appropriately in formulating conclusions?	No
Were the methods used to combine the findings of studies appropriate?	Yes
Was the likelihood of publication bias assessed?	Yes
Was the conflict of interest included?	No

Table 25. Quality assessment of studies relevant to question 2

Question	Lees 2015	Spaggiari 2013	Tveit 2009
EXTERNAL VALIDITY			
Have actual probability values been reported for the main outcomes except where the probability value is less than 0.001?	No	Yes	Yes
Were the subjects asked to participate in the study representative of the population of interest for this review?	Yes	Yes	Yes
BIAS			
Was an attempt made to blind study subjects to the intervention they have received?	No	NA	NA
Was an attempt made to blind those measuring the main outcomes of the intervention?	No	No	No
If any of the results of the study were based on “data dredging”, was this made clear?	No dredging	No dredging	No dredging
Were the statistical tests used to assess the main outcomes appropriate?	NA	Yes	Yes
Was compliance with the intervention/s reliable?	Yes	Yes	Yes
Were the main outcome measures used accurate (valid and reliable)?	No	Yes	No
Were the main outcome measures defined using UK definitions?	No	No	No
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Unclear	Yes
Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Yes	Yes	Yes
Were study subjects randomised to intervention groups?	Yes	No	No
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	Yes	NA	NA
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Yes	No	Yes
Were losses of patients to follow-up taken into account?	Yes	Yes	Yes
POWER			
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	Unclear	Unclear	Unclear

Table 26. Quality assessment of studies relevant to question 3

Question	Ayala 2012	Haddad 2016	Kingdom 2011	Sharp 2018 (STRIDER)	Subtil 2003a, Subtil 2003b	Von Dadelszen
EXTERNAL VALIDITY						

Question	Ayala 2012	Haddad 2016	Kingdom 2011	Sharp 2018 (STRIDER)	Subtil 2003a, Subtil 2003b	Von Dadelszen
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Yes	Yes	Yes	Yes	NA	Yes
Were the subjects asked to participate in the study representative of the population of interest for this review?	No	No	Unclear	Unclear	Yes	Unclear
BIAS						
Was an attempt made to blind study subjects to the intervention they have received?	Yes	No	No	Yes	No	No
Was an attempt made to blind those measuring the main outcomes of the intervention?	Yes	Yes	No	Yes	No	No
If any of the results of the study were based on “data dredging”, was this made clear?	Dredging	No dredging	No dredging	No dredging	No dredging	No dredging
Were the statistical tests used to assess the main outcomes appropriate?	Yes	Yes	NA	Yes	NA	NA
Was compliance with the intervention/s reliable?	Yes	Unclear	Yes	Yes	Unclear	Unclear
Were the main outcome measures used accurate (valid and reliable)?	Unclear	Yes	Unclear	Unclear	Yes	Unclear
Were the main outcome measures defined using UK definitions?	No	No	No	No	No	No
Were the patients in different intervention groups (trials and cohort	Yes	Yes	Yes	Yes	NA	No

Question	Ayala 2012	Haddad 2016	Kingdom 2011	Sharp 2018 (STRIDER)	Subtil 2003a, Subtil 2003b	Von Dadelszen
studies) or were the cases and controls (case-control studies) recruited from the same population?						
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes	Yes	Yes	No	No
Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Yes	Yes	Yes	Yes	Yes	Yes
Were study subjects randomised to intervention groups?	Yes	Yes	Yes	Yes	No	No
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	Yes	Yes	Yes	Unclear	NA	NA
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Unclear	Yes	Yes	Yes	NA	NA
Were losses of patients to follow-up taken into account?	Yes	Yes	Yes	Yes	Yes	Unclear
POWER						
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference	Unclear	No	No	No	NA	No

Question	Ayala 2012	Haddad 2016	Kingdom 2011	Sharp 2018 (STRIDER)	Subtil 2003a, Subtil 2003b	Von Dadelszen
being due to chance is less than 5%?						

Table 27. Quality assessment of studies relevant to question 4

Question	DIGITAT[15]	Rabinovich 2018[16]	Walker 2016[17]
EXTERNAL VALIDITY			
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Yes	Yes	NA
Were the subjects asked to participate in the study representative of the population of interest for this review?	Yes	Unclear	Unclear
BIAS			
Was an attempt made to blind study subjects to the intervention they have received?	No	NA	No
Was an attempt made to blind those measuring the main outcomes of the intervention?	No	NA	No
If any of the results of the study were based on "data dredging", was this made clear?	Dredging	No dredging	No dredging
Were the statistical tests used to assess the main outcomes appropriate?	Yes	Yes	NA
Was compliance with the intervention/s reliable?	Yes	NA	Yes
Were the main outcome measures used accurate (valid and reliable)?	Unclear	Unclear	Yes
Were the main outcome measures defined using UK definitions?	Unclear (stillbirth), no (SGA)	Unclear	Yes
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	No	Yes
Were study subjects in different intervention groups	Yes	Yes	Yes

Question	DIGITAT[15]	Rabinovich 2018[16]	Walker 2016[17]
(trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?			
Were study subjects randomised to intervention groups?	Yes	No	Yes
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	Yes	NA	Yes
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Unclear	No	NA
Were losses of patients to follow-up taken into account?	Yes	Yes	Yes
POWER			
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	Unclear	Unclear	No

Appendix 4 – Guidance for quality assessment of included studies

Table 28. QUADAS-2 guidance for the quality assessment of studies included for Question 1[62]

Question		
PARTICIPANT SELECTION		
Was a consecutive or random sample of pregnancies enrolled?	A study should ideally enrol all consecutive, or a random sample of, eligible patients – otherwise there is potential for bias. Studies that make inappropriate exclusions, e.g. excluding "difficult to diagnose" patients, may result in overoptimistic estimates of diagnostic accuracy	**Yes** if all pregnancies (or a random sample of patients) within the study period were included **No** if patients were selected in a different way, e.g. by referral or convenience sample **Unclear** if all screened pregnancies are enrolled but it is not specified if the screening test is routinely administered at the study site
Was a case-control design avoided?	Studies enrolling patients with known disease and a control group without the condition may exaggerate diagnostic accuracy	**Yes** if the study was a prospective or retrospective cohort study **No** if cases (diagnosed post-delivery) were matched to controls
Did the study avoid inappropriate	Exclusion of patients with "red flags" for the target condition, who	**Yes** if all pregnancies were included, or if exclusions were appropriate

Question		
exclusions?	may be easier to diagnose, may lead to underestimation of diagnostic accuracy	and unlikely to lead to bias **No** if any group within the screening population was systematically excluded
Could the selection of pregnancies have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the included pregnancies do not match the review question?	There may be concerns regarding applicability if patients included in the study differ, compared to those targeted by the review question, in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols	**Low** if patients overall are low-risk pregnancies representative of the screening population **High** if patients overall are not representative of the screening population, such as pregnancies with at least one moderate risk factor as specified in UK guidelines
INDEX TESTS		
Were the index test results interpreted without knowledge of the reference standard?	This item is similar to "blinding" in intervention studies. Interpretation of index test results may be influenced by knowledge of the reference standard	**Yes** if screening results were interpreted before the diagnosis was confirmed (i.e. before birth) **No** if screening results were only examined after the diagnosis was confirmed
If a threshold was used, was it pre-specified?	Selecting the test threshold to optimise sensitivity and/or specificity may lead to overoptimistic estimates of test performance, which is likely to be poorer in an independent sample of patients in whom the same threshold is used	**Yes** if the criteria used to diagnose stillbirth or SGA were explicitly stated, well-defined, and specified before the study **No** if criteria were not stated, were insufficiently well-defined, or were specified retrospectively
Could the conduct or interpretation of the index test have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the index test, its conduct, or interpretation differ from the review question?	Variations in test technology, execution, or interpretation may affect estimates of its diagnostic accuracy. If index tests methods vary from those specified in the review question there may be concerns regarding applicability	**Low** if the screening test is similar to tests or screening tests administered as part of UK clinical practice **High** if any aspect of the index test, including its conduct or interpretation, is substantially different from clinical practice in a UK setting (as outlined in the CG 62 NICE guidance[35])
REFERENCE STANDARD		
Is the reference standard likely to correctly classify the test condition?	Estimates of test accuracy are based on the assumption that the reference standard is 100% sensitive and specific. Disagreements between the reference standard and index test are assumed to result from incorrect classification by the index test	**Yes** if stillbirth was confirmed postnatally or, if liveborn, SGA<3rd centile was confirmed postnatally **No** if no postnatal diagnosis or if postnatal diagnosis was performed inconsistently, or if the methods used are likely to be unreliable
Were the reference standard results interpreted without knowledge of the results of the index test?	Potential for bias is related to the potential influence of prior knowledge on the interpretation of the reference standard	**Yes** if the final diagnosis of stillbirth or liveborn SGA<3rd centile were made by a blinded investigator **No** if the screening results were known by the investigator making the final diagnosis **Unclear** if it is not clear whether the investigator was aware of the test result when making the final diagnosis

Question		
Could the reference standard, its conduct, or its interpretation have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain
Is there concern that the target condition as defined by the reference standard does not match the review question?	The reference standard may be free of bias but the target condition that it defines may differ from the target condition specified in the review question. For example, when defining urinary tract infection, the reference standard is generally based on specimen culture but the threshold above which a result is considered positive may vary	**Low** if the definition of stillbirth was the standard UK definition: *Death of a baby before or during birth after 24 weeks of gestation. The baby must not at any time breathe or show any other signs of life* **High** if the reference standard defined stillbirth in any other way
PARTICIPANT FLOW		
Was there an appropriate interval between the index test(s) and the reference standard?	Ideally results of the index test and reference standard are collected on the same patients at the same time. If there is a delay or if treatment is started between index test and reference standard, misclassification may occur due to recovery or deterioration of the condition. The length of interval leading to a high risk of bias will vary between conditions. A delay of a few days may not be a problem for chronic conditions, while for acute infectious diseases a short delay may be important	Question removed – not relevant to the review
Did all participants receive a reference standard? Did participants receive the same reference standard?	Verification bias occurs when not all of the study group receive confirmation of the diagnosis by the same reference standard. If the results of the index test influence the decision on whether to perform the reference standard or which reference standard is used, estimated diagnostic accuracy may be biased	**Yes, Yes** if all screened patients had postnatal confirmation of their diagnosis, and all were diagnosed postnatally in the same manner (similarly trained staff, similar timing of diagnosis) **No, Yes** if not all patients were diagnosed postnatally, but those who were got diagnosed in the same manner **No** if patients received different reference standards **Unclear** if staff diagnosing and recording stillbirth may have been from different centres, or may not have received the same training
Were all pregnancies included in the analysis?	All patients who were recruited into the study should be included in the analysis. There is a potential for bias if the number of patients enrolled differs from the number of patients included in the 2x2 table of results, for example because patients lost to follow-up differ systematically from those who remain	**Yes** if all screened patients were included in the final analysis **No** if any screened patients were not included in the final analysis
Could the participant flow have introduced bias?	If all signalling questions for a domain are answered "yes" then risk of bias can be judged "low". If any signalling question is answered "no" this flags the potential for bias	Answered based on the previous questions in this domain

Table 29. Downs and Black guidance used to assess the quality of studies included for Question 2 to 4[63]

Question	Guideline criteria for stillbirth studies
EXTERNAL VALIDITY	

Question	Guideline criteria for stillbirth studies
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	Answer should relate to the outcome measures of interest to this review
Modified question: Were the subjects asked to participate in the study representative of the population of interest for this review? Original question: Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	**Yes** only when pregnant women were identified as being at risk of stillbirth through screening and the population they were identified from was representative of the expected screening population, that is, women who would be part of the normal antenatal care pathway (NICE CG 62)1 **No** if participants were at risk of stillbirth as determined by risk factors or tests or they were determined to be at risk of severe FGR or pre-eclampsia, where the risk factors would also put them at risk of stillbirth OR if the population from which participants were initially selected from was not representative of the expected screening population
Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?	Question removed – studies with unrepresentative treatment settings to be excluded from the review
BIAS	
Was an attempt made to blind study subjects to the intervention they have received?	**Yes** if patients were blind to treatment allocation, and if methods of blinding were appropriate, such as use of matching placebos **No** if any patients were aware of treatment allocation **NA** if observational study
Was an attempt made to blind those measuring the main outcomes of the intervention?	**Yes** if outcome assessors were blind to treatment allocation, and if methods of blinding were appropriate, such as data analysis taking place at a separate site using blinded datasets **No** if any outcome assessors (i.e. the healthcare professionals making the diagnosis of stillbirth or SGA) were aware of treatment allocation
If any of the results of the study were based on “data dredging”, was this made clear?	**No dredging** if all outcomes were pre-specified, and all outcomes listed in the methods section are fully reported **Dredging** if the authors report that any analyses were post hoc, or if a large number of results **Unclear** if the methods section or protocol do not specify a list of primary and secondary outcomes, and it is not clear whether outcomes were pre-specified
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	Question removed – all studies should include follow-up from intervention to birth
Were the statistical tests used to assess the main outcomes	**Yes** if treatment groups were compared appropriately using risk difference, risk ratios, odds ratios, unpaired t-tests or similar; for single-arm trials a paired t-test may be appropriate; other methods may

Question	Guideline criteria for stillbirth studies
appropriate?	also be appropriate if justified in the publication **No** if the statistical tests were not appropriate – to be determined on a case-by-case basis **Unclear** if the statistical methods were not specified **NA** if no statistical tests were performed
Was compliance with the intervention/s reliable?	**Yes** if compliance or adherence were reported and were above 80% **No** if compliance or adherence were below 80% **Unclear** if compliance or adherence were not reported
Were the main outcome measures used accurate (valid and reliable)?	Answer should relate to the outcome measures of interest to this review **Yes** when the definition of stillbirth or SGA was pre-specified and the criteria were fully reported (gestational age, weight, size). SGA should be defined using international standards. Cases should be routinely recorded in a reliable manner. Particularly for SGA, consider whether staff were adequately trained to make measurements **No** if the definition was not pre-specified, the criteria were unclear, outcomes were not routinely and consistently recorded, or staff were not adequately trained to make measurements
Question added: Were the main outcome measures defined using UK definitions?	For stillbirth, **Yes** if the study used the standard UK definition: "Death of a baby before or during birth after 24 weeks of gestation. The baby must not at any time breathe or show any other signs of life" For SGA, **Yes** if the definition was an infant born with a birth weight less than the 3rd centile based on local or national charts from the UK **No** if the study defined stillbirth or SGA in any other way
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	**Yes** if patients from all intervention groups were recruited from the same population **No** if different intervention groups were recruited from different populations, such as different geographical location, different baseline characteristics, or patients selected using a different screening test
Question added: Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	**Yes** if baseline characteristics were similar between treatment groups, particularly maternal age, BMI, pre-existing disorders, smoking or other narcotic use, and screening results **No** if there were significant differences between the groups in any of the characteristics listed above **Unclear** if relevant baseline characteristics were not reported or it is not clear if the differences between these are significant
Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	**Yes** if patients from all intervention groups were recruited over the same period of time **No** if patients from different intervention groups were recruited at different times, such as historical control groups
Were study subjects randomised to intervention groups?	**Yes** if randomisation was performed using computer-generated random numbers or random number tables

Question	Guideline criteria for stillbirth studies
	Inadequate if alternation, case record numbers, birth dates or week days were used to allocate patients to treatment arms **No** if no attempt was made at randomisation
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	**Yes** if the allocation sequence was protected before and until assignment, using methods such as: centralised or pharmacy-controlled randomisation, serially-numbered identical containers, on-site computer-based system with a randomisation sequence that is not readable until allocation, or other approaches with robust methods to prevent foreknowledge of the allocation sequence **No** if inadequate methods of randomisation were used, or if random number lists could have been viewed before allocation, such as open random number lists or serially numbered envelopes **NA** in non-randomised studies
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Answer should relate to the outcome measures of interest to this review **Yes** if analyses were adjusted for differences in key baseline characteristics (maternal age, BMI, pre-existing disorders, smoking or other narcotic use, and screening results), or if adjustment was not necessary **No** if adjustment was necessary but was not performed
Were losses of patients to follow-up taken into account?	Answer should relate to the outcome measures of interest to this review **Yes** if there were no imbalances in drop-outs between groups, or if there was an imbalance in drop-outs but this was discussed and accounted for in the statistical analyses; for RCTs, check whether an intention-to-treat (ITT) analysis was used and whether this was appropriate (generally appropriate for superiority studies, not appropriate for non-inferiority studies) **No** if drop-out rates were unbalanced and this was not explained or adjusted for or when ITT analysis was used incorrectly or inappropriately
POWER	
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	**Yes** if power calculations are reported and an adequate sample size was used **No** if power calculations are reported and an adequate sample size was not reached **Unclear** if power calculations are not reported (adequate sample sizes may be calculated for each outcome when a clinically important difference has been determined)

Table 30. The AMSTAR checklist guidance used for quality assessment of the Conde-Agudelo 2015 SLR[64]

1. Was an 'a priori' design provided?	The research question and inclusion criteria should be established before the conduct of the review.	*Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a "yes."*
2. Was there duplicate study selection and data extraction?	There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.	*Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the*

		other's work.
3. Was a comprehensive literature search performed?	At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found	*Note: If at least 2 sources + one supplementary strategy used, select "yes" (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).*
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.	*Note: If review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SINGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit*
5. Was a list of studies (included and excluded) provided?	A list of included and excluded studies should be provided.	*Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no."*
6. Were the characteristics of the included studies provided?	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.	*Note: Acceptable if not in table format as long as they are described as above.*
7. Was the scientific quality of the included studies assessed and documented?	A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.	*Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).*
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.	*Note: Might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.*
9. Were the methods used to combine the findings of studies appropriate?	For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).	*Note: Indicate "yes" if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.*
10. Was the likelihood of publication bias assessed?	An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).	*Note: If no test values or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.*
11. Was the conflict of interest	Potential sources of support should be clearly acknowledged in both the	*Note: To get a "yes," must indicate source of funding or*

included?	systematic review and the included studies.	*support for the systematic review AND for each of the included studies.*

Appendix 5 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table 31.

Table 31. UK NSC reporting checklist for evidence summaries

	Section	Item	Page no.
1.	TITLE AND SUMMARIES		
1.1	Title sheet	Identify the review as a UK NSC evidence summary.	Title page
1.2	Plain English summary	Plain English description of the executive summary.	5
1.3	Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review.	6
2.	INTRODUCTION AND APPROACH		
2.1	Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search. Method – briefly outline the rapid review methods used.	10-17
2.2	Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided *a priori*.	18-23
2.3	Appraisal for quality/risk of bias tool	Details of tool/checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR.	24, 186-193
3.	SEARCH STRATEGY AND STUDY SELECTION (FOR EACH KEY QUESTION)		

3.1	Databases/ sources searched	Give details of all databases searched (including platform/interface and coverage dates) and date of final search.	76
3.2	Search strategy and results	Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used. Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.	76-81
3.3	Study selection	State the process for selecting studies – inclusion and exclusion criteria, number of studies screened by title/abstract and full text, number of reviewers, any cross checking carried out.	18, 81
4.	STUDY LEVEL REPORTING OF RESULTS (FOR EACH KEY QUESTION)		
4.1	Study level reporting, results and risk of bias assessment	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up period, outcomes reported, statistical analyses etc.). Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available. For each study, present the results of any assessment of quality/risk of bias.	Study level reporting: 97-174 Quality assessment: 175-186
5.	QUESTION LEVEL SYNTHESIS		
5.1	Description of the evidence	For each question, give numbers of studies screened, assessed for eligibility, and included in the review, with summary reasons for exclusion.	26, 46, 53
5.2	Combining and presenting the findings	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of studies. Consideration of four components should inform the reviewer's judgement on whether the criterion is 'met', 'not met' or 'uncertain': quantity; quality; applicability and consistency.	27-44, 47-51, 54-63
5.3	Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/risk of bias issues for each question. Have the criteria addressed been 'met', 'not met' or 'uncertain'?	45, 52, 64
6.	REVIEW SUMMARY		
6.1	Conclusions and implications	Do findings indicate whether screening should be recommended?	65-67

	for policy	Is further work warranted? Are there gaps in the evidence highlighted by the review?	
6.2	Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.	68-69

Appendix 6 – Likelihood ratios and graphs of sensitivity versus (1−specificity)

A 'perfect' diagnostic test is one that is able to discriminate between test subjects who truly have and truly do not have the test condition (that is, 100% sensitivity and 100% specificity [FPR rate of 0%]). Therefore, the closer the pair of sensitivity/FPR is to the top left of the graph, the better the test performance. These graphs can also be used to consider each test result in relation to a given likelihood ratio (LR), which indicates how much that result will increase or decrease the probability of the patient having the condition that they are tested for, and is considered to be a powerful measure of diagnostic test accuracy.[92] If only sensitivity and specificity values are available, the receiver-operator curve (ROC) analysis allows for the visualisation of LR categories delineated by a positive LR of 1, 2, 5 and 10 and a negative LR− of 0.1, 0.2, 0.5 and 1. The general consensus is that tests with LR+ >10 and LR− <0.1 are considered to have an acceptable accuracy, and could be considered for use in screening for a condition in clinical practice.[92, 93]

Figure 6. Positive likelihood ratio categories

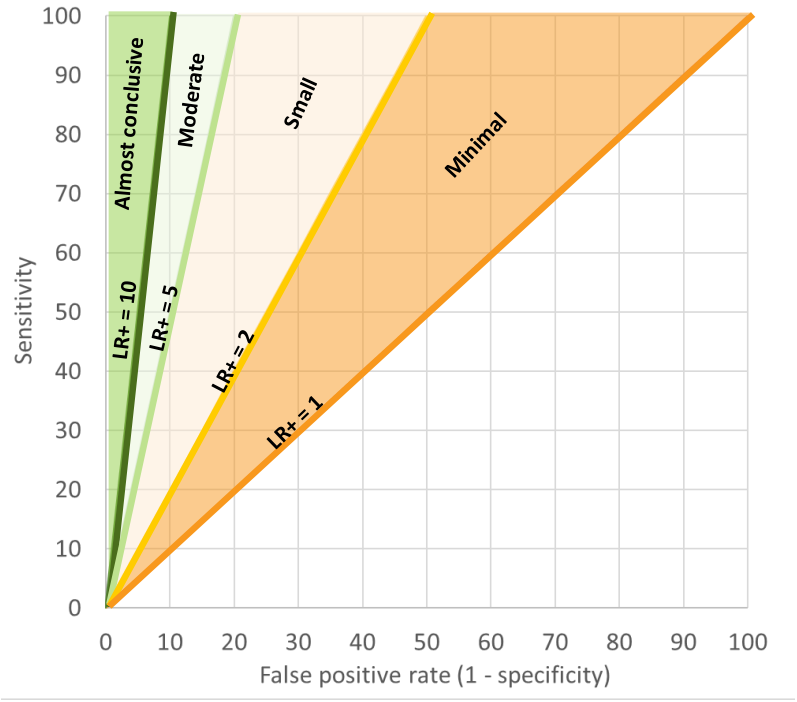
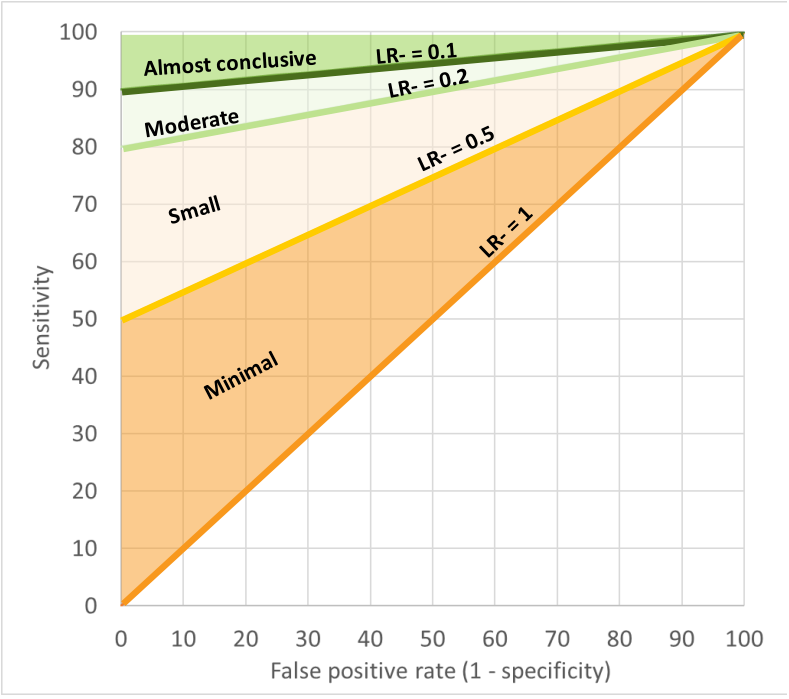


Figure 7. Negative likelihood ratio categories



Appendix 7 – Graphs of test accuracy for the prediction of stillbirths arising from impaired placentation

Figure 8. Measures of test accuracy for the prediction of stillbirths arising from impaired placentation at 5% FPR

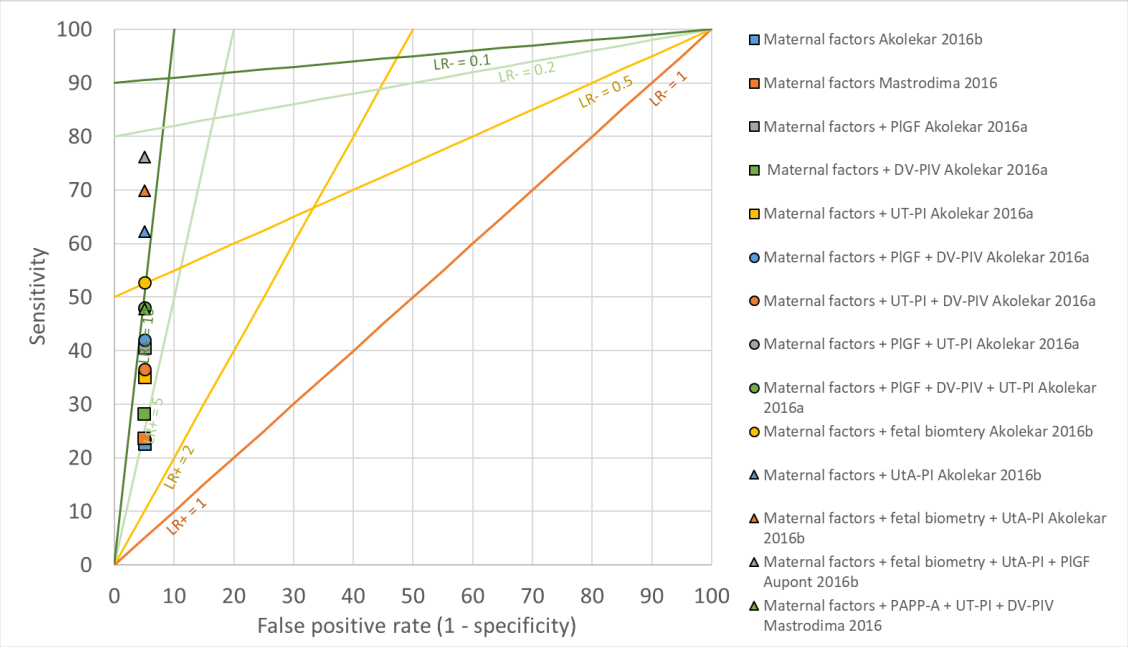
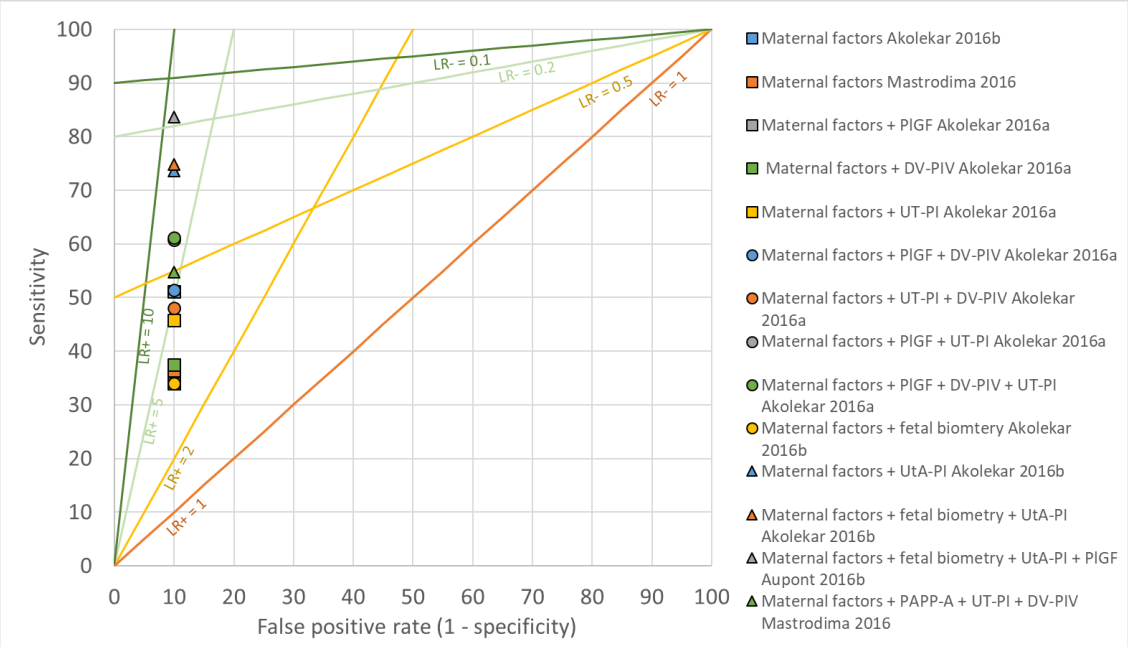


Figure 9. Measures of test accuracy for the prediction of stillbirths arising from impaired placentation at 10% FPR



Appendix 8 – Screening tests for the prediction of all-cause stillbirth

Table 32. Measures of test accuracy for screening tests for all-cause stillbirth at any gestational age

Test	Study	Women	Sens (%)	Spec (%)	PPV	NPV	LR+	LR−
Biochemical tests								
AFP <0.4 to 0.5 MoM	Conde-Agudelo 2015	111,607	6	94	NR	NR	1.0	1.0
AFP ≥1.7 to 1.8 MoM	Conde-Agudelo 2015	93,252	13	95	NR	NR	2.6	0.9
AFP >2.0 MoM	Conde-Agudelo 2015[a]	131,466	11	96	NR	NR	3.1	0.9
	Tancrede 2015	3,466	0	95.9	0	99.8	NR	NR
AFP ≥2.5 MoM	Conde-Agudelo 2015[b]	186,802	9	98	NR	NR	4.0	0.9
Inhibin A ≥2.0 MoM	Conde-Agudelo 2015	33,145	19	97	NR	NR	6.1	0.8
Inhibin A (threshold NR)	Dugoff 2008 (FPR 1%)	35,253	8	99	NR	NR	NR	NR
	Dugoff 2008 (FPR 5%)	35,253	17	95	NR	NR	NR	NR
	Dugoff 2008 (FPR 10%)	35,253	27	90	NR	NR	NR	NR
sVEGFR-1 >97.5th centile	Chaiworapongsa 2017	804	27	97	NR	NR	8.4	0.75
sEng >90th centile	Chaiworapongsa 2017	804	64	89	NR	NR	5.9	0.41
sEng >97.5th centile	Chaiworapongsa 2017	804	55	96	NR	NR	13.7	0.47
PlGF <10th centile	Chaiworapongsa 2017	804	55	87	NR	NR	4.1	0.52
PlGF <2.5th centile	Chaiworapongsa 2017	804	45	95	NR	NR	9.9	0.57
PAPP-A <0.25 to 0.30 MoM	Conde-Agudelo 2015[b]	21,158	15	95	NR	NR	3.3	0.9
PAPP-A <0.4 to 0.5 MoM	Conde-Aguelo 2015[b]	114,343	14	95	NR	NR	2.7	0.9
PAPP-A <10th centile	Dugoff 2004	33,395	15.79	89.91	0.44	99.73	NR	NR
PAPP-A <5th centile	Dugoff 2004	33,395	10.53	94.81	0.58	99.74	NR	NR

PAPP-A <1st centile	Dugoff 2004	33,395	3.16	98.94	0.84	99.72	NR	NR
PAPP-A	Marttala 2010	19,536	15	95.3	1.0	99.7	NR	NR
hCG <0.5 MoM	Conde-Agudelo 2015	13,180	4	94	NR	NR	0.7	1.0
hCG >2.0 MoM	Tancrede 2015	3,466	40	89.9	0.9	99.8	NR	NR
hCG ≥2.0 to 2.5 MoM	Conde-Agudelo 2015[c]	170,617	12	93	NR	NR	1.6	1.0
hCG >4.0 MoM	Conde-Agudelo 2015	2,406	70	75	NR	NR	2.8	0.4
Free β-human chorionic gonadotrophin ≤5th centile MoM	Conde-Agudelo 2015	56,814	12	93	NR	NR	1.8	0.9
Unconjugated oestriol ≤0.5 to 0.7 MoM	Conde-Agudelo 2015[d]	58,417	15	96	NR	NR	4.0	0.9
Ultrasound tests								
UtA Bilateral notch	Conde-Agudelo 2015	243	100	93	NR	NR	13.3	0.0[e]
UtA Any notch	Conde-Agudelo 2015	346	100	58	NR	NR	2.4	0.0[e]
UtA-PI (threshold NR)	Familiari 2016 (FPR 10%)	23,894	28	90	NR	NR	NR	NR
UtA-PI >95th percentile	Valino 2016a	30,261	16.3	94.5	NR	NR	NR	NR
UtA-PI >90th percentile	Conde-Agudelo 2015	65,819	40	90	NR	NR	4.0	0.7
UtA-PI >90th percentile	Poon 2013	65,819	39.9	NR	NR	NR	NR	NR
UtA-RI >0.58	Conde-Agudelo 2015	10,873	16	91	NR	NR	1.8	0.9
UtA-RI 90th percentile	Singh 2012	15,786	46.2	NR	0.46	95.73	NR	NR
UtA-RI >95th percentile	Singh 2012	15,786	35.4	NR	NR	NR	NR	NR
UtA-RI >99th percentile	Singh 2012	15,786	15.4	NR	NR	NR	NR	NR
Ductus venosus Doppler (Reversed A wave)	Conde-Agudelo 2015	33,452	8	97	NR	NR	3.0	0.9
Fetal nuchal translucency (Any increased NT measurement)	Conde-Agudelo 2015[b]	58,772	10	95	NR	NR	2.0	0.9
Fetal nuchal translucency (≥2 to 3 mm)	Conde-Agudelo 2015[b]	10,777	13	95	NR	NR	2.6	0.9
Fetal nuchal translucency (ΔNT	Conde-Agudelo 2015	47,995	9	95	NR	NR	1.8	1.0

>95th centile)								
Fetal growth measurements								
Femur length centile	Familiari 2016 (FPR 10%)	23,894	26	90	NR	NR	NR	NR
Isolated short femur <10th percentile	Conde-Agudelo 2015	73,613	2	99	NR	NR	5.3	1.0
Fetal growth standard	Hemming 2011 (FPR 10%)	540,849	43	90	2.11	99.68	4.29	NR
Suboptimal fetal growth	Conde-Agudelo 2015	743,975	32	75	NR	NR	1.3	0.9
Customised growth standard	Smith 2014[d]	196	39	NR	NR	NR	NR	NR
	Odibo 2012	59,016	55.5	90.4	3.8	99.6	NR	NR
Population growth norms	Smith 2014[d]	196	14	NR	NR	NR	NR	NR
	Odibo 2012	59,016	19.5	89.2	1.8	99.1	NR	NR
Customised growth chart with US biometric parameters	Odibo 2012	59,016	54.9	90.3	3.5	99.6	NR	NR
Ultrasound growth norm (Hadlock)	Smith 2014[d]	196	33	NR	NR	NR	NR	NR
Non-sex-specific growth standard	Trudell 2015	57,170	32	92.8	2.4	99.6	4.43	0.73
Sex-specific growth standard	Trudell 2015	57,170	64	92	4.3	99.8	7.96	0.39
Combinations of tests								
Maternal factors	Yerlikaya 2016 (FPR 5%)	113,415	18.4	95	NR	NR	NR	NR
Maternal factors	Yerlikaya 2016 (FPR 10%)	113,415	29	90	NR	NR	NR	NR
Maternal factors	Familiari 2016 (FPR 10%)	23,894	19	90	NR	NR	NR	NR
Maternal factors (clinical cut-point of 3)	Trudell 2017	64,173	53.1	65.4	NR	NR	1.54	NR
Maternal factors (maternal age, smoking during pregnancy and maternal height)	Sutan 2010	541,811	4.2	99.4	1.4	99.8	NR	NR
Maternal factors + PlGF + UT-PI + DV-PIV (11 to 13 weeks)	Akolekar 2016a (FPR 5%)	45,452	31.7	95	NR	NR	NR	NR
	Akolekar 2016a (FPR 10%)		41.9	90	NR	NR	NR	NR
Maternal factors (19 to 24 weeks)	Akolekar 2016b (FPR 5%)	70,003	19.0	95	NR	NR	NR	NR

	Akolekar 2016b (FPR 10%)		29.5	90	NR	NR	NR	NR
Maternal factors + fetal biometry (19 to 24 weeks)	Akolekar 2016b (FPR 5%)	70, 003	32.2	95	NR	NR	NR	NR
	Akolekar 2016b (FPR 10%)		42.5	90	NR	NR	NR	NR
Maternal factors + UtA-PI (19 to 24 weeks)	Akolekar 2016b (FPR 5%)	70, 003	41.8	95	NR	NR	NR	NR
	Akolekar 2016b FPR (10%)		52.6	90	NR	NR	NR	NR
Maternal factors + fetal biometry + UtA-PI (19 to 24 weeks)	Akolekar 2016b (FPR 5%)	70, 003	45.1	95	NR	NR	NR	NR
	Akolekar 2016b (FPR 10%)		54.7	90	NR	NR	NR	NR
Maternal factors + biometry + UtA-PI + PlGF (19 to 24 weeks)	Aupont, 2016 (FPR 5%)	70, 003	50.7	95	NR	NR	NR	NR
	Aupont, 2016 (FPR 10%)		57.6	90	NR	NR	NR	NR
Cerebroplacental ratio <5th centile (MCA-PI/UA-PI MoM)	Bakalis 2016	30,780	8.5	94.8	NR	NR	NR	NR
Angiogenic index-1 (PlGF/s VEGFR-1 <10th centile (24 to 28 weeks gestation)	Chaiworapongsa 2017	840	64	89	NR	NR	5.7	0.41
Angiogenic index-1 (PlGF/s VEGFR-1 <2.5th centile (24 to 28 weeks gestation)	Chaiworapongsa 2017	840	55	96	NR	NR	14.6	0.47
Angiogenic index-1 (PlGF/s VEGFR-1 <0.12 MoM (30 to 34 weeks gestation)	Chaiworapongsa 2013a	1269	80	94	5	100	14.2	0.2
Angiogenic index-1 (PlGF/s VEGFR-1 ≤0.046 (at 30 to 34 weeks gestation)	Chaiworapongsa 2013b	35	80	93	NR	NR	NR	NR
PlGF/sEng <10th centile	Chaiworapongsa 2017	840	64	87	NR	NR	5.1	0.42
PlGF/sEng <2.5th centile	Chaiworapongsa 2017	840	55	96	NR	NR	13.7	0.47
PlGF/sEng	Chaiworapongsa 2013a	1269	60	89	2	99	5.5	0.4
PlGF/sEng ≤11.7	Chaiworapongsa 2013b	35	80	93	NR	NR	NR	NR
≥2 abnormal markers (AFP ≥ 2.0, hCG ≥ 2.0, uE3 ≤ 0.5, or inhibin A ≥ 2.0 MoMs)	Dugoff 2005	33,145	13.4	98.63	1.23	99.74	4.23	0.89

versus single or no abnormal markers (<2)								
Maternal factors + femur length centile + uterine artery Doppler (UtA-PI)	Familiari 2016 (FPR 10%)	23,894	31	90	NR	NR	NR	NR
Maternal factors + PAPP-A + UT-PI + DV-PIV (at 11to13 weeks)	Mastrodima 2016 (FPR 5%)	76,897	32.5	95	NR	NR	NR	NR
	Mastrodima 2016 (FPR 10%)	76,897	39.9	90	NR	NR	NR	NR
UtA-PI + UA-PI + MCA-PI + MAP + PlGF + sFlt-1	Valino 2016b	8268	30.4	10	NR	NR	NR	NR
First-trimester Down screening (Risk ≥1:270-300)	Conde-Agudelo 2015[b]	34,013	10	96	NR	NR	2.8	0.9
Second-trimester Down screening (risk ≥1:190-270)	Conde-Agudelo 2015[b]	2,566	67	61	NR	NR	1.8	0.5
First- and second-trimester sequential integrated screening for Down syndrome (risk ≥1:200)	Conde-Agudelo 2015	38,120	36	81	NR	NR	1.9	0.8

Abbreviations: AFP, alpha-fetoprotein; Ang-1/Ang-2, angiopoietin 1 and 2; DV PIV, ductus venous pulsatility index for veins; FPR, false positive rate; hCG, human chorionic gonadotropin; MoM, multiples of the median; NPV, negative predictive value; NR, not reported; NT, nuchal translucency; PAPP-A, pregnancy-associated plasma protein A; PI, pulsatility index; PLGF, placenta growth factor; PPV, positive predictive value; RI, resistance index; Sens, sensitivity; Spec: specificity, UtA, uterine artery

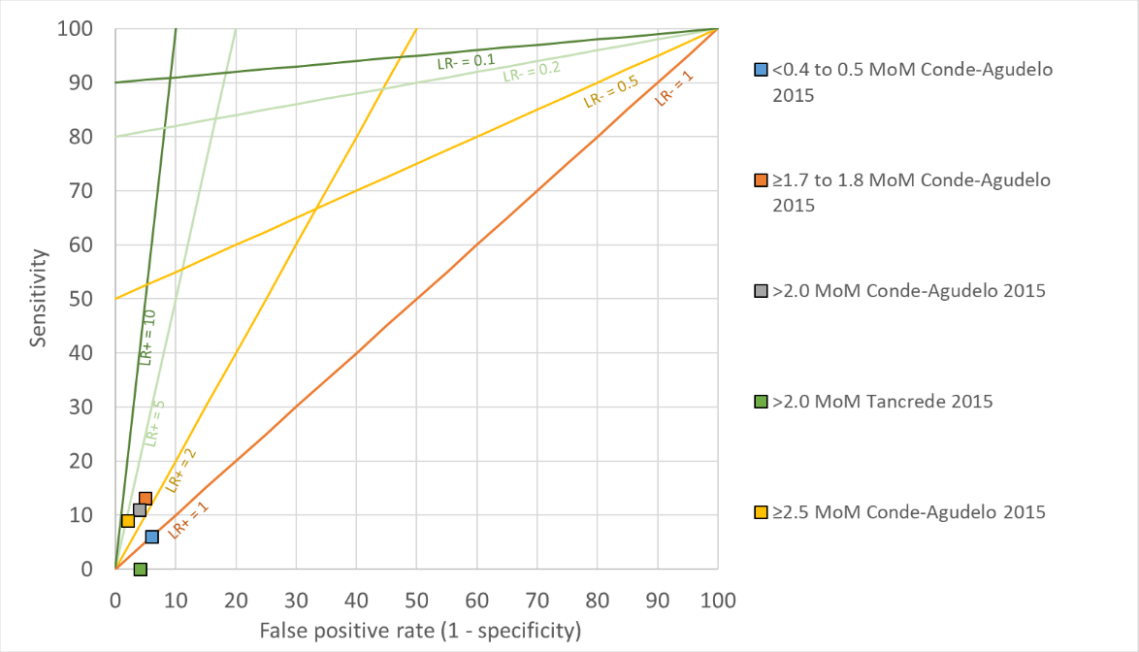
a Two studies conducted in countries where population and antenatal care pathway differ from the UK setting.

b One study conducted in countries where population and antenatal care pathway differ from the UK setting.

c Four studies conducted in countries where population and antenatal care pathway differ from the UK setting

d Sensitivity for identification of SGA as a predictor of stillbirth

Figure 10. AFP for predicting stillbirth



* Tancrede 2015: pre-term births only

Figure 11. Inhibin A for the prediction of all-cause stillbirth

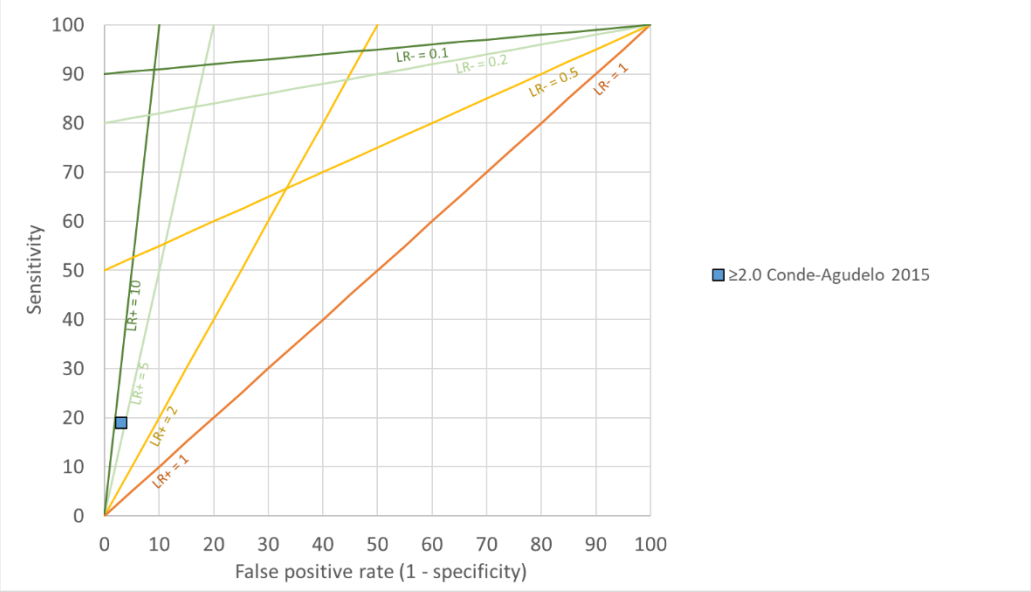


Figure 12. sVEGFR-1 for the prediction of all-cause stillbirth

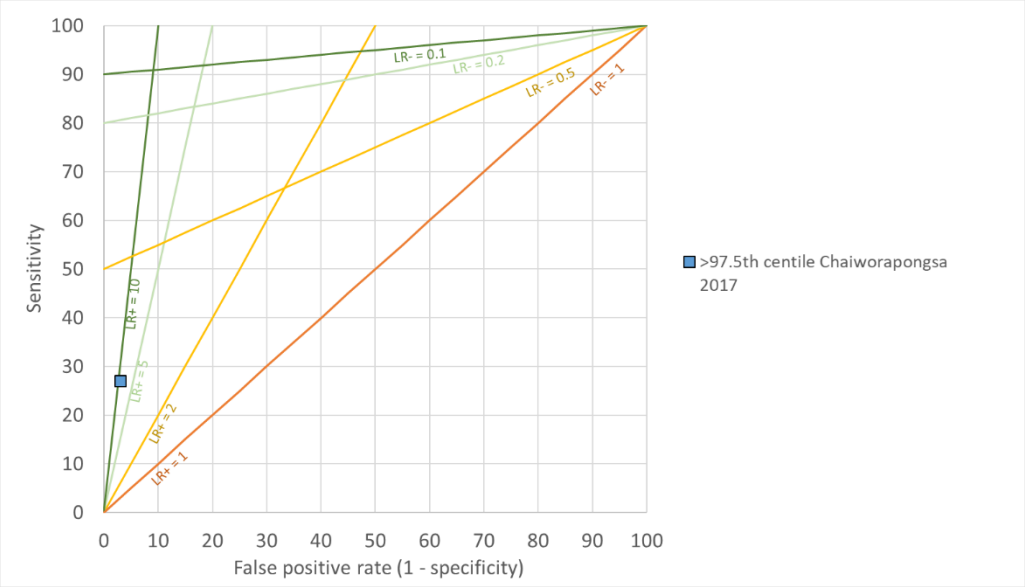


Figure 13. sEng for the prediction of all-cause stillbirth

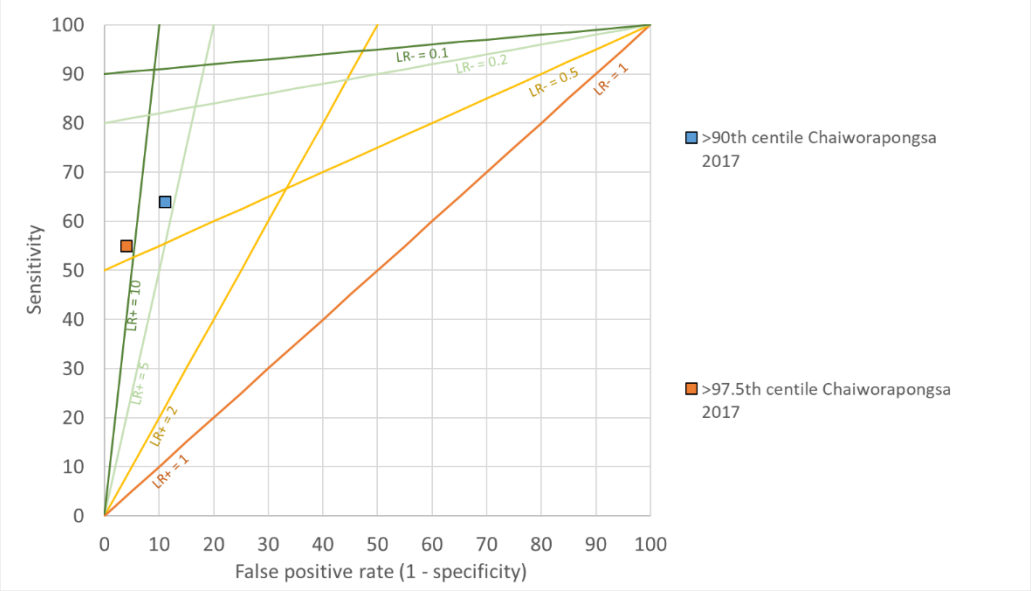


Figure 14. PlGF for the prediction of all-cause stillbirth

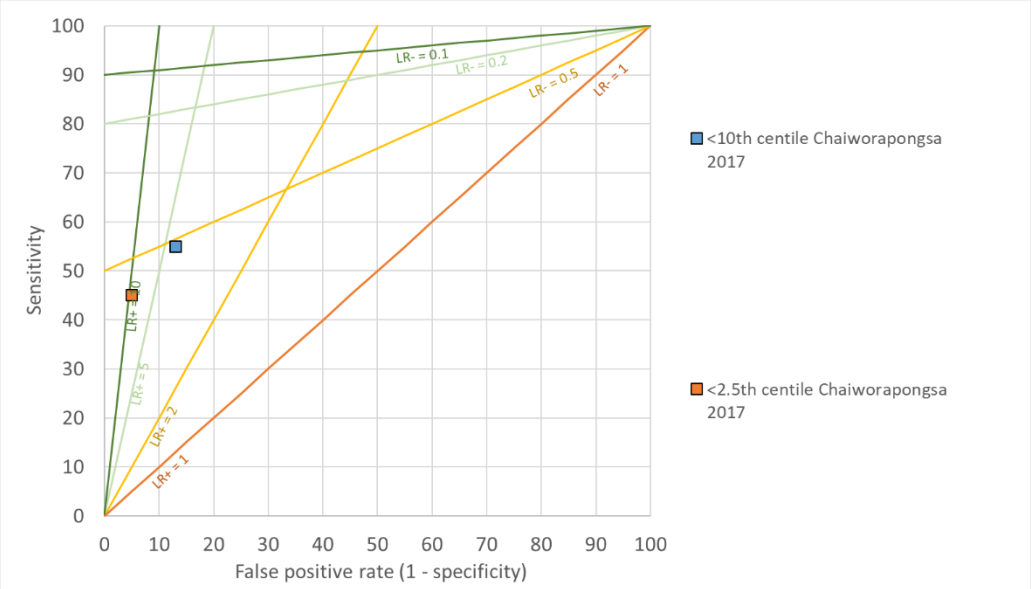


Figure 15. PAPP-A for the prediction of all-cause stillbirth

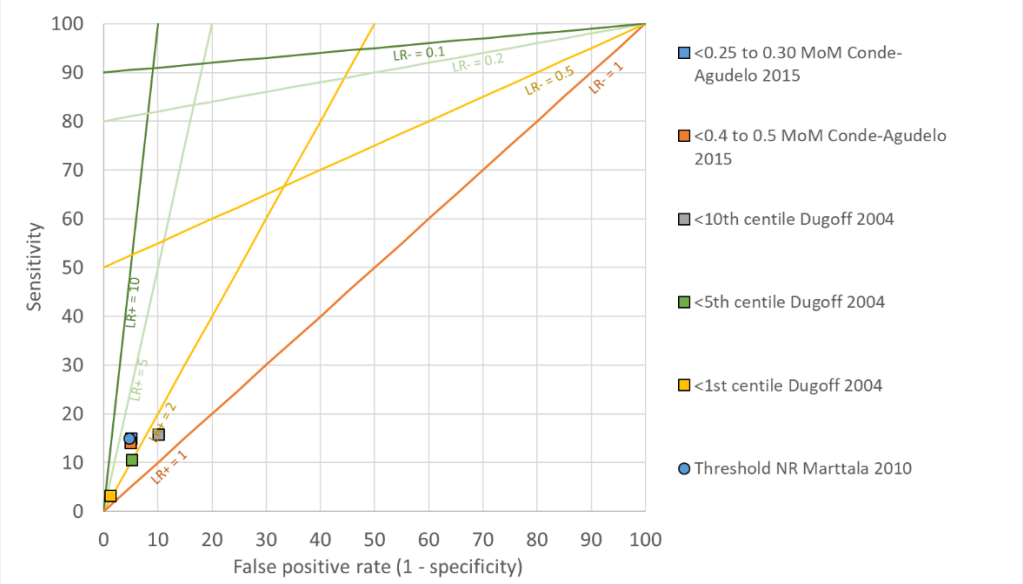


Figure 16. HCG for the prediction of all-cause stillbirth

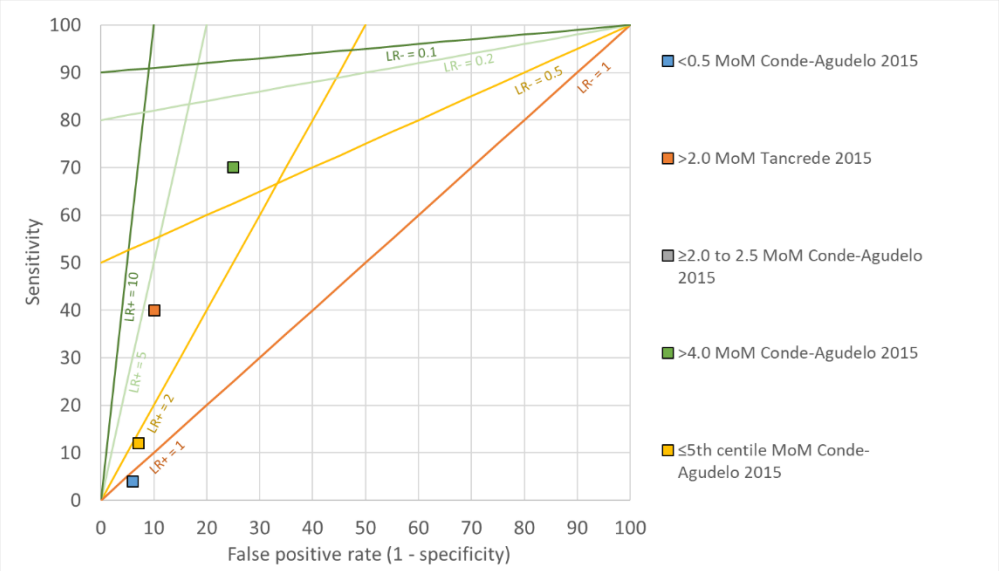


Figure 17. Unconjugated oestriol for the prediction of all-cause stillbirth

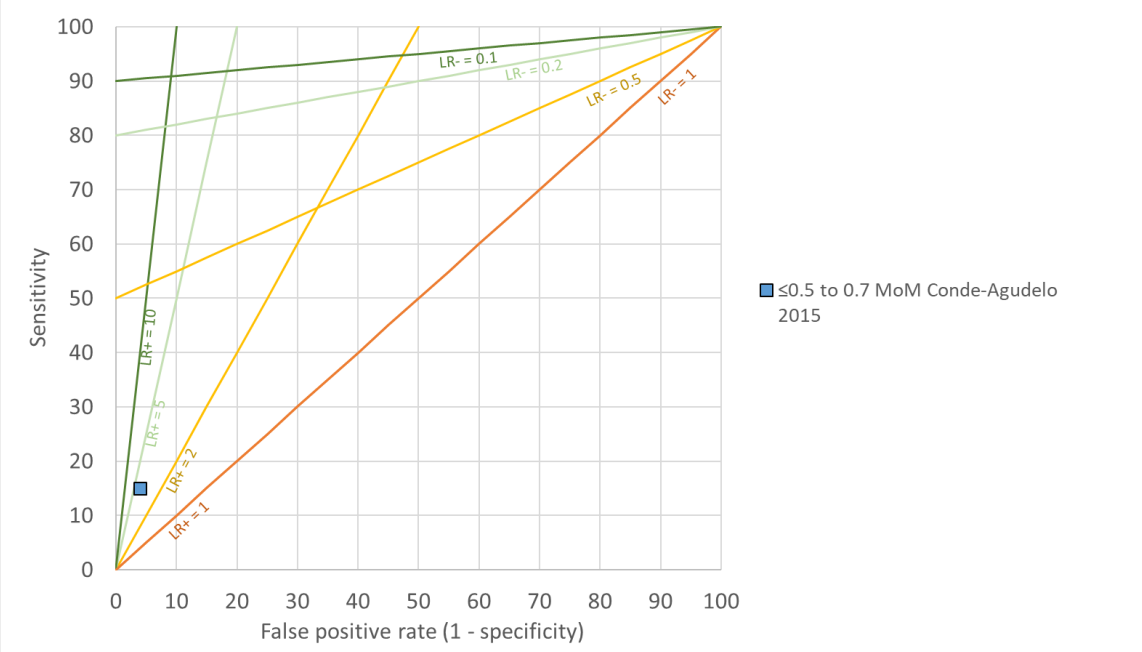


Figure 18. Ultrasound Doppler tests for the prediction of all-cause stillbirth

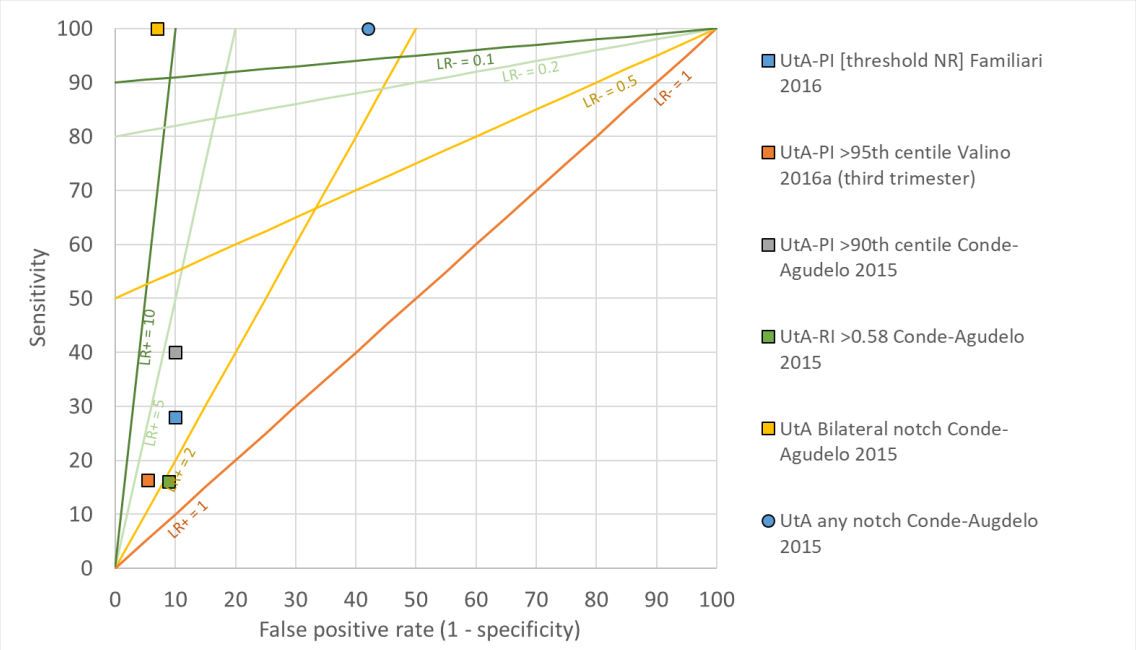


Figure 19. Nuchal translucency for the prediction of all-cause stillbirth

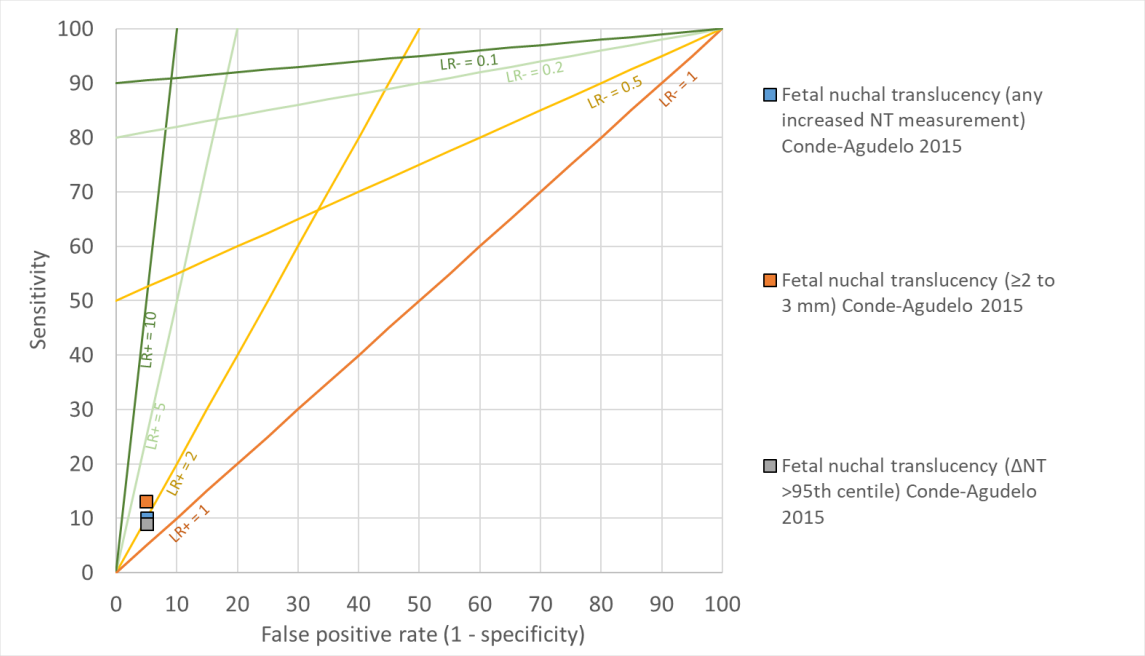


Figure 20. Fetal growth monitoring for the prediction of all-cause stillbirth

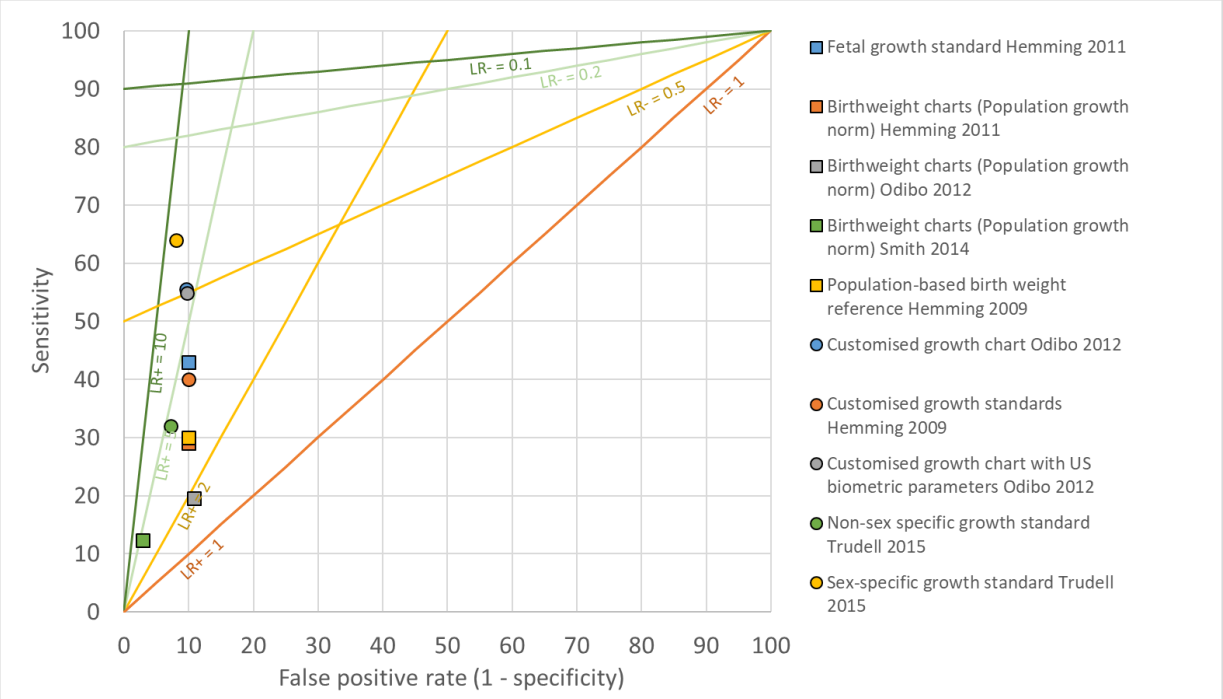


Figure 21. Combination tests in the first or second trimester for the prediction of all-cause stillbirth

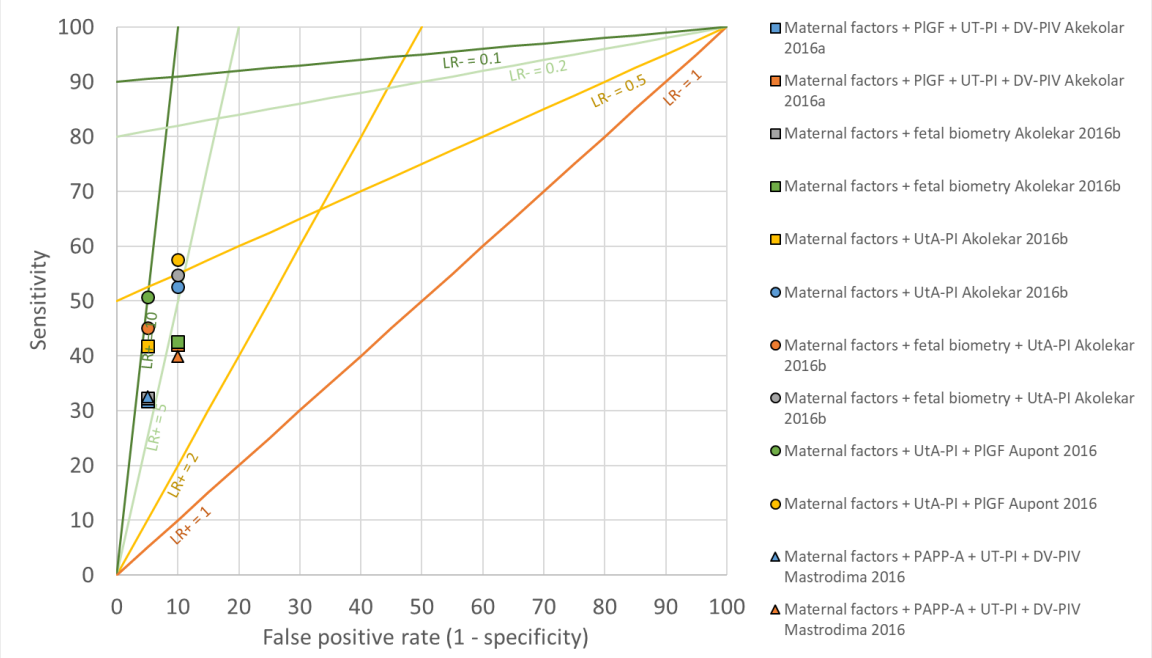


Figure 22. Combination tests in the first or second trimester for the prediction of all-cause stillbirth

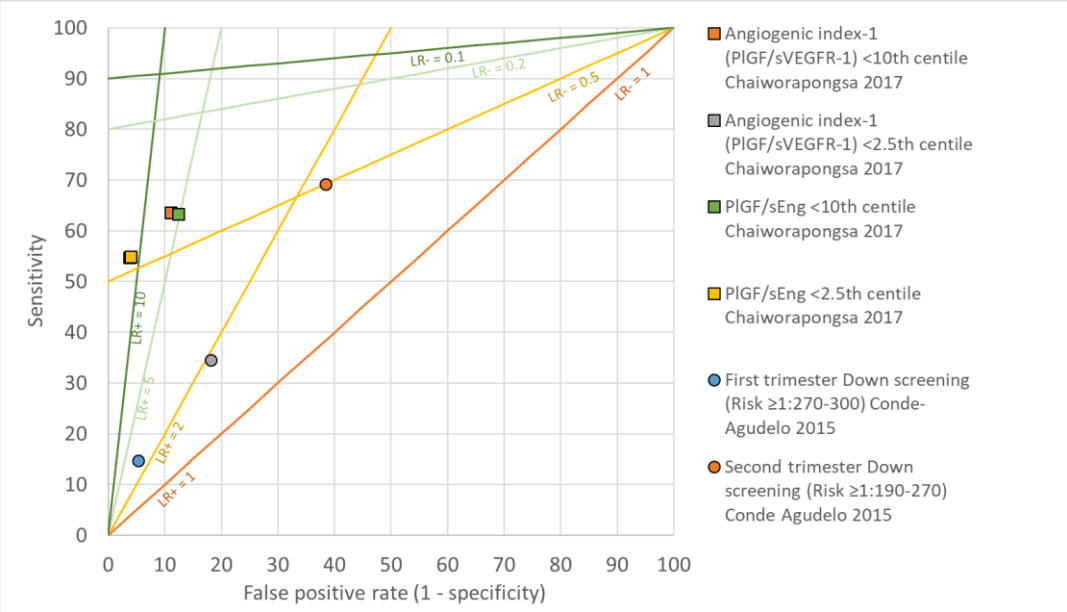
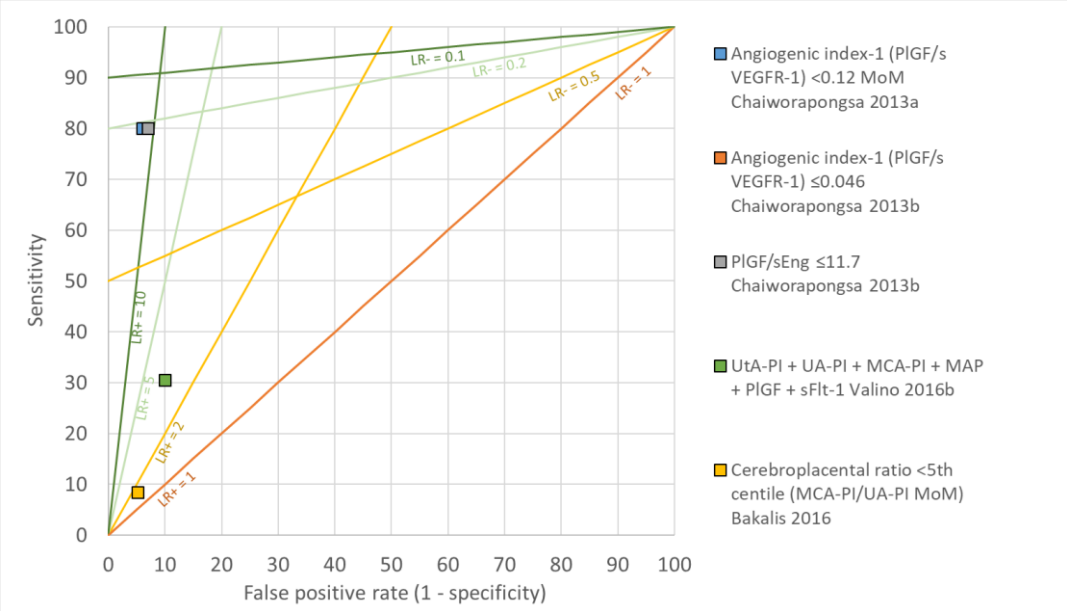


Figure 23. Combination tests in the third trimester for the prediction of all-cause stillbirth



As discussed in the results section, only one index test, as reported by the Conde-Agudelo 2015 SLR, achieved a positive LR >10 and a negative LR <0.1, thereby falling into the 'acceptable accurate'*' category. This test, using UtA bilateral notch measurements, reportedly achieved a LR+ of 13.3 and a LR− of 0 (100% sensitivity and 93% specificity). However, as this study had a very small sample size with only 10 cases of stillbirth, it is likely that the test accuracy has been overestimated. A meta-regression conducted in the Conde-Agudelo 2015 SLR found that studies with fewer than 25 stillbirths in the study sample often overestimated the accuracy of the test, additionally concluding that using UtA bilateral notch is unlikely to be a sufficiently accurate test for stillbirth due to placental insufficiency.[44]

Aa number of index tests approached this threshold of LR+ >10 and LR– <0.1. UtA notching (any notch), as reported by the Conde-Agudelo 2015 SLR achieved a LR– of 0, but had a LR+ of only 2.4.[44] However, this result was from a study with a small sample size of 346 women.[107] Chaiworapongsa 2013a (prospective cohort study) reported a sensitivity of 80% at 94% specificity for the prediction of all-cause stillbirth for an angiogenic index-1 <0.12, and Chaiworapongsa 2013b (case-control study) reported a sensitivity of 80% at 93% specificity for an angiogenic index-1 ≤0.046. The case-control study additionally reported a sensitivity of 80% with 93% specificity for the prediction of all-cause stillbirth for a PlGF/sEng ratio ≤11.7.[69] These tests thereby fall into the category of LR+ >5 and almost into the category of LR– of 0.2. However, the case-control study had a very small sample size of 35 births, with only five cases of stillbirth. Furthermore, both the prospective cohort and case-control studies were considered to have a high risk of bias for participant selection and for the interpretation of the index test. In particular, thresholds were not pre-specified and so it is likely that the reported threshold was chosen to optimise sensitivity and specificity of the test.

This review additionally identified studies examining screening for biomarkers in the third trimester of pregnancy (Figure 23).[68, 69, 76] It should be noted that this is not reflective of general UK clinical practice for low-risk pregnancies; tests for biochemical abnormalities are typically performed in the first or second trimester. High measures of test accuracy for the angiogenic-1 index and the PlGF/sEng ratio performed at 30 to 34 weeks of gestation were reported by the Chaiworapongsa 2013 cohort and case-control studies.[69] However, these studies were at high risk of bias, and so these results alone do not support that screening for biochemical markers in isolation or as a combination in the third trimester could have superior accuracy than screening in the first or second trimester.

The most accurate growth assessment method was the use of a sex-specific growth standard as reported by Trudell 2015,[84] achieving a LR+ of 7.96 and LR– of 0.39. In comparison, using a non-sex specific growth standard for the prediction of stillbirth produced a LR+ of 4.43 and a LR− of 0.73.