Prediction of fetal growth restriction and complications: Individual Participant Data (IPD) meta-analysis with Decision Curve Analysis

International Prediction of Complications in Pregnancy: Fetal Growth restriction (IPPIC-FGR)

1. Summary of research

Background

One in 10 babies are born small. Of these, a third are at risk of perinatal complications associated with restricted growth. Fetal growth restriction is a major contributor to stillbirths and neonatal deaths. Half of the stillborn fetuses are considered to be growth restricted. Close monitoring of growth-restricted babies with serial ultrasound reduces perinatal deaths. But in two-thirds of babies born after 32 weeks, a diagnosis of fetal growth restriction is missed antenatally. A policy of universal ultrasound in all pregnant women to detect fetal growth restriction has not improved outcomes.

Any effort to prevent adverse offspring outcomes needs to: identify pregnancies that are at risk of delivering a growth restricted baby with severe complications to plan management such as delivery; assess the actual severity of smallness to determine the timing and frequency of surveillance – therefore two prediction models are required. Prediction of serious complications such as stillbirth and perinatal death need large sample sizes. Furthermore, robust external validation of models requires access to multiple large external datasets.

Our HTA funded IPPIC (International Prediction of Pregnancy Complications) Collaborative Network to predict pre-eclampsia consists of individual participant data (IPD) of about 3 million pregnancies. This large global repository, which includes 15 UK and 66 international datasets, has information on the relevant clinical, biochemical and ultrasound predictors of fetal growth and associated serious neonatal complications.

Aims and objectives

To develop and validate prediction models for accurate identification of women at risk of fetal growth restriction and at high risk of perinatal complications.

We will undertake an IPD meta-analysis to develop and validate (both internally and externally) the prediction models using maternal clinical characteristics, biochemical markers and ultrasound findings. They will allow individual outcomes to be predicted, including
(i) the risk of delivering a growth restricted baby with severe complications (birth weight < 10th centile adjusted for gestational age with stillbirth or neonatal death at any time or delivery before 32 weeks) (model 1)
(ii) a baby's birth weight at various potential gestational ages at delivery (with the flexibility to convert into centiles using existing fetal growth standards) to determine the likelihood of a baby being small, and the severity of its smallness (model 2).

Methods

In addition to the existing data in our IPPIC repository, we will update our search, add additional data from any new studies, and undertake the analysis. We will identify the relevant population from the IPD repository, develop/improve and validate the models using bootstrapping and internal-external cross-validation (IECV). Decision curves will be plotted for different models to decide which model offers the highest net-benefit and clinical utility.

Timelines for delivery

We anticipate the proposal to take 17 months for completion. We have already completed the cleaning, coding and standardisation of relevant data (until March 2017) across studies. The

required funding is to update the search, develop the protocol in detail, add and clean new datasets, consolidate the datasets, and undertake the analysis.

Applicability and expected impact

The models will be developed for application during routine antenatal visits at 12-week and 20-week (for ultrasound), and 28-weeks (for pregnancy check). Information on the risks of delivering a growth-restricted baby could reduce perinatal mortality and severe morbidity such as HIE (Hypoxic Ischaemic Encephalopathy). An accurate prediction tool has the potential to save costs to the NHS by £100 million / year.

2. Background

The problem

Perinatal mortality rates in the UK are higher than in many European countries. The fall in stillbirth rate is low; neonatal death rate is stagnant.[1] Until now, most stillbirths were classified as 'unexplained,' and by implication, unavoidable.[2] It is now recognised that failed fetal growth, defined as fetal growth restriction or intra uterine growth restriction, preceded the majority of these stillbirths - half of the 3000 babies who were stillborn every year in the UK were considered to be growth restricted.[3] This knowledge has shifted our classification of the proportion of stillbirths as 'unexplained' from 70% to 15%.[4] There is a need to identify women at risk of fetal growth restriction to minimise perinatal deaths.

Fetal growth restriction

The term 'fetal growth restriction' is often used interchangeably (and erroneously), with 'small for gestational age' (SGA), where the birth weight of the fetus is less than the 10th centile. Of the 70,000 babies who are born small each year in England and Wales,[5] most (70%) are constitutionally small, without major complications. But one in three small babies is growth restricted, with arrest or shift in rates of growth trajectory, which increases their risk of immediate and long-term complications.[3,6]

In 20-30% of growth-restricted fetuses, the condition is diagnosed early (<32 weeks), and is usually associated with hypertensive disorders of pregnancy and severe placental pathology.[7] These infants are often delivered early, with additional prematurity related complications. The majority (70-80%) of cases of fetal growth restriction are of late onset (>32 weeks). The diagnosis is unknown in three-quarters of these babies with late-onset growth restriction.[8-10] The odds of stillbirth (OR 7.1-10.0) and neonatal death (OR 3.4-9.4) are significantly higher in growth-restricted than normal weight fetuses at every week beyond the expected date of delivery in these babies.[11] The long-term complications include neurodevelopmental problems, poor growth and increased susceptibility to adult-onset diseases in infancy and adolescence, including obesity, metabolic syndrome, type 2 diabetes and cardiovascular disease in growth-restricted babies.[12]

Priority area for the NHS

Reduction in stillbirths and neonatal deaths[13] is a priority for the NHS.[3] The Secretary of State, supported by the Royal College of Obstetricians and Gynaecologists (RCOG), has declared a national ambition to halve the rates of stillbirths, neonatal deaths and intrapartum brain injuries by 2025, with a 20% reduction by 2020.[14] Reduction in fetal and neonatal mortality and morbidity is part of the NHS England Business Plan (2015-16),[15] and is a key indicator in the NHS Outcomes Framework.[16] Accurate prediction of fetal growth restriction is crucial to achieving this objective. A recent James Lind Alliance Research Priority Setting Initiative for stillbirth identified "the use of routine tests and monitoring procedures to improve the detection of growth restricted fetus to help prevent stillbirth" as number 5 of its top 10 goals.[17] The recent MBBRACE-UK (Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries in the UK) perinatal confidential

enquiry also highlighted the slow progress and the need for improving the diagnosis of fetal growth restriction and reducing stillbirths. [1]

What is needed?

Early and accurate prediction of fetal growth restriction is needed to identify women who need close monitoring in pregnancy, and to plan the setting and timing of delivery to minimise adverse perinatal outcomes. This requires well-developed, externally validated prediction models that are integrated for use within existing antenatal care. The models should include clinically relevant predictors, and predict outcomes that are critical to the management of women at risk of fetal growth restriction.

3. Current evidence and need for the study

3.1 Existing guidelines on identifying women at risk of fetal growth restriction

There is wide variation between guidelines on recommendations to identify women at risk of small for gestational age or fetal growth restriction. Current UK national guidelines (RCOG) advocate regular ultrasound for fetal growth in women with at least one 'major' risk factor in clinical history,[18] and to undertake further uterine artery Doppler when mothers have three or more 'minor' risk factors. The categorisation of risk factors as 'major' or 'minor' and their combination was arbitrarily determined. The guideline developers acknowledged the lack of robust information on the relationships between risk factors in an individual woman to predict the outcome. The Society of Obstetricians and Gynaecologists of Canada (SOGC) advocates clinical risk factors (not specified) based screening and recommends consideration of other investigations such as biochemical markers and uterine artery Doppler.[19] The ACOG (American College of Obstetricians and Gynecologists) does not recommend uterine artery Doppler or biochemical markers and cite lack of evidence on outcomes.[20] The Australians (RANZCOG) suggest a combination of biomarkers, Doppler ultrasound, and 'major' maternal clinical risk factors, with recommendations for further research to evaluate the predictive ability and cost-effectiveness of the tests.[21] None of the guidelines differentiates between early and late onset growth restriction, or with small for gestational age fetus.

3.2 Potential predictors of fetal growth restriction

Numerous systematic reviews and primary studies have identified the following tests to have the potential to predict fetal growth restriction or small for gestational age fetus (Table 1).

Table1: Markers for the prediction of fetal growth restriction

Type of marker	Predictors
Maternal clinical characteristics	Maternal characteristics: age, body mass index, smoking, alcohol intake or substance misuse, exercise, diet Medical history: chronic hypertension, diabetes, renal disease, heritable thrombophilia, autoimmune disease, cardiac disease Obstetric history: parity, previous SGA, previous stillbirth, previous pre-eclampsia, pregnancy interval Current pregnancy: mode of conception, weight gain, early pregnancy bleeding
Biochemical	PlGF, PAPP-A, sFlt-1, AFP, HCG, urine dipstick, 24 hour urine protein
Ultrasound	Uterine artery Doppler (resistance index, pulsatility index, unilateral or bilateral notching), abdominal circumference, fetal cerebral-placental ratio, estimated fetal weight, fetal echogenic bowel, nuchal translucency

PlGF- Placental Growth Factor, AFP- Alpha feto protein, HCG- Human Chorionic Gonadotrophin, PAPP-A Pregnancy associated plasma protein A, SGA Small for gestational age

3.3 Gaps in existing research

Prediction of fetal growth restriction

To-date, 337 studies have reported on the accuracy of individual tests to predict either fetal growth restriction or small for gestational age fetus.[22] Twenty-eight models have been published to predict the risk of small for gestational age fetus[23] - none of them is in clinical use (limitations of the models are provided in Section 5.2). Research to predict, screen or detect fetal growth restriction has been fraught with difficulties involving design, population, tests and outcomes. Firstly, the terms 'prediction' and 'screening', which have separate objectives, are often used interchangeably.[24] In the former, the outcome of interest (fetal growth restriction) has not yet occurred, while in the latter, the focus is on accurately detecting established fetal growth restriction. Some of the models to predict growth restriction in the fetus use tests as late as 36 weeks of pregnancy, which are more relevant for diagnosis than prediction.[23] Secondly, the population studied is often heterogeneous, or only limited to specific subgroups such as nulliparous women.[25] Thirdly, the predictors have often been dichotomised, thereby reducing their power. Fourthly, studies often predict small for gestational age fetus rather than growth-restricted infants. Fetal growth restriction is variously defined using either ultrasound characteristics (expected fetal weight, fetal abdominal circumference, Doppler blood flows) or by using birth weight.[26] Furthermore, both expected fetal weight and birth weight have been reported in centiles that were either adjusted for various maternal characteristics (customised) or for only gestational age (population-based).[27,28] The centile cut-offs to define growth restriction are varied (<10th, <5th, < 3rd). Very few studies have additionally included perinatal mortality and morbidity outcomes.[29]

Universal screening with ultrasound

The Cochrane review of randomised trials on universal screening with ultrasound in all pregnant women vs. current strategy of selective screening in high-risk women for fetal growth restriction has not shown any reductions in perinatal mortality and morbidity.[30] The latter strategy only detects 20% of small babies, while with the former strategy, for every small for gestational age fetus with complications, two more small babies that are otherwise normal are picked up.[25] The National Institute for Health and Care Excellence's (NICE) antenatal care guideline concluded that 'the methods by which small for gestational age fetus can be identified antenatally are poorly developed or are not tested by rigorous methodology.'[31] Universal ultrasound screening of all women for detection of fetal growth restriction will significantly strain finite resources. Furthermore, implementation of such a strategy in low-risk women in France did not lower the rates of complications in small for gestational age fetuses, but resulted in iatrogenic prematurity in screen-positive pregnancies.[32]

3.4 Why is an IPD meta-analysis needed?

We propose an IPD meta-analysis to develop and validate models to predict risks of delivering a growth-restricted baby with serious complications, and also predict the extent of its smallness. We consider IPD meta-analysis to be the optimal design to synthesise the existing evidence for the following reasons.

- Aggregate meta-analyses are restricted by the heterogeneity in the characteristics of the population, the timing of predictor measurement, choice of cut-offs to categorise continuous predictors, and the type/definition of outcome in published studies. IPD meta-analysis can reduce heterogeneity in patient selection through strict inclusion and exclusion criteria (i.e. removal and addition of particular patients in the dataset). While a primary study will only be able to evaluate the model's performance in a specific group of women (unselected or low risk or nulliparous), our large sample size allows us to assess the differential accuracy of the model for different subgroups.

- Studies included in aggregate meta-analysis often report on only one (or a few) predictors (test), despite available information on more than one predictor (test) in the raw data. They also consider different sets of included predictors, which make their synthesis problematic. Information on the predictive performance of multiple predictors in individual studies is often

missing, or only provided for the overall population (e.g. total predicted versus total observed with fetal growth restriction), and not for specific subgroups across the spectrum of predicted risks.

- The definition of fetal growth restriction is varied in primary studies. To robustly predict the clinically important, but relatively rare outcome growth restriction with serious perinatal complications (stillbirth, neonatal death, extreme preterm delivery), and early-onset growth restriction, large sample sizes are needed for the primary studies. In our IPD meta-analysis, we will be able to access the above outcomes and provide a more standardised definition across studies.

- Due to numerous problems of primary studies investigating predictor-outcome associations, especially publication bias and selective reporting, aggregate meta-analyses show inconsistent and even contradictory predictor-outcome associations.[33] In IPD meta-analysis, the association between future outcome and patient-level characteristics and study level characteristics (setting, timing, study design) can be assessed more reliably. For e.g., by using a more consistent set of adjustment factors and modelling biomarkers on their continuous scale (rather than categorisation).[34] Furthermore, there are now novel methods for imputation of systematically missing predictors across studies[35-39] for predictors which are entirely missing in some studies. These methods allow values to be imputed for missing predictors by borrowing information from other studies, under a missing at random assumption and accounting for between-study heterogeneity. This allows more predictors, and indeed more studies, to be included in the IPD than aggregate meta-analysis.

- Before application of a prediction model in clinical practice, there is a need to evaluate its performance in the population(s) in which it is intended for use. This requires external validation of the model in a dataset different to that in which it was developed. Lack of external validation is one of the key reasons for prediction models not being adopted in clinical practice. IPD meta-analysis offers an accepted way to overcome this current lack of validation.[40] We will maximise the data for model development and validation by using an 'internal-external cross validation' approach, which makes better use of multiple studies by rotating them toward model development and validation. Predictive performance (e.g. in terms of calibration and discrimination) can then be checked in each study, and summarised itself in a meta-analysis.[41] The net-benefit (clinical utility) of the model can also be checked in each study, and a summary decision curve provided, to examine the benefit versus harms of using the model in practice.[42]

- Problems with aggregate data arise with differential treatment (management) effects such as close monitoring in the third trimester with ultrasound, and use of aspirin by patient characteristics. Our IPD includes studies with this information, and will facilitate a more reliable meta-analysis, with details on intense monitoring and treatment with aspirin available at the individual-level. This will allow the external validation performance of a model to be evaluated across different groups of individuals defined by their treatment, and by considering the inclusion of treatment as a predictor in the model.

- Single primary studies usually have a small and local dataset, which leads to overfitting and optimism when deriving a new prediction model, such that the model does not perform as well when applied in new individuals. Combining IPD across studies will greatly increase the power to precisely estimate the predictor effects and minimise the potential for overfitting and optimism. Given the size of the IPD already available, it is cost-effective and more efficient to perform an IPD meta-analysis and make full use of the existing evidence, rather than setting up a new cohort study that would require recruitment and follow-up over a long period to attain the necessary sample size.

In summary, by accessing and synthesising the raw individual participant data, our IPD meta-analysis will address many issues facing an aggregate data meta-analysis or a new primary study (Table 2). In particular, a robust prediction model for fetal growth restriction will be developed and

validated by standardising the outcome definition, having a large sample size, using multiple candidate predictors in combination, applying novel statistical methods for imputation, and evaluating predictive performance overall and in relevant subgroups.

Table 2. Comparison of aggregate data and IPD meta-analysis approach for evidence synthesis on the prediction of fetal growth restriction

	Structured question components	Aggregate meta-analysis	IPD meta-analysis (IPPIC-FGR)
Population			
Pregnant women (nulliparous, low risk, or unselected)	Takes into account the different baseline risks across various groups of women in the included studies	No	Yes
Identification of any subgroups in which markers appear to perform best	The association between outcome and patient-level characteristics or between patient and study level characteristics (setting, study design) can be assessed in this group of women, without the ecological fallacy problem	No, as subgroup results and interactions rarely available	Yes
	Evaluates the differential performance of the prediction model according to the subgroups based on population (unselected vs. selected), timing of test (first vs. any trimester)	No, as predictive performance usually only provided for all individuals (if at all)	Yes
Tests (predictors)			
New and existing biochemical markers, ultrasound markers and combinations of markers and risk models	Adjust for multiple predictors such as clinical history, biochemical and ultrasound markers	Limited, as different adjustment factors used in each study	Yes
	The continuous predictors can be maintained as continuous values instead of dichotomous measures (using cut-points), thereby maximising the prognostic information of the tests	No, as often dichotomised (at different cut offs)	Yes
	Takes into account the effect of management (e.g. aspirin) that influences the outcome	No	Yes
Outcome			
Fetal growth restriction	Predictive performance of the model for fetal growth restriction defined using various centiles, and with complications (stillbirth, neonatal death at any time and delivery before 32 weeks)	No	Yes
Birth weight	Predicted birth weight for various gestational ages of delivery	No	Yes
Clinical applicability			
If findings suggest it is appropriate, model should	Produces a single, integrated prediction model that can be implemented in practice after validation in multiple datasets	No	Yes

be developed to explore its clinical use	Involves key global researchers in fetal growth restriction prediction and databases, with potential to improve implementation of the model	Limited	Yes

3.5 Developing and identifying a risk prediction tool with good predictive performance

A good prediction model is one that is accurate, validated in populations and datasets external to those used to develop the model, widely applicable in practice, acceptable to patients, and ultimately improves clinical outcomes by helping clinicians and patients make more informed decisions. Our prediction models will attempt to achieve this in the following ways: use rigorous statistical methods to develop the model and assess accuracy; undertake a formal internal and external validation within the IPD datasets; use unambiguous definitions of fetal growth restriction as an outcome; standardise definitions of predictors based on reproducible measurements using methods available in clinical practice; adjust and/or evaluate performance according to current clinical management; involve patient groups in model development and implementation; and produce personalised risk information that enables mothers and clinicians to make informed decisions on management aspects like frequent monitoring, appropriate referral to tertiary centres, and early delivery.

4. Aim

To accurately identify fetuses at risk of growth restriction and perinatal complications using prediction models.

Objectives

We will develop, internally and externally validate, and (if necessary) update separate models in pregnant women to determine (i) the overall risk of delivering a growth restricted fetus (birth weight less than 10th centile adjusted for gestational age) with serious perinatal complications (stillbirth or neonatal death at any time or delivery before 32 weeks) (ii) the birth weight if delivered at various gestational ages (with flexibility to convert into centiles using existing fetal growth standards) to assess the extent of smallness, using data that are already available in our large NIHR HTA funded IPD repository.

Primary

1. To establish whether existing models are suitable for the target population or if new models are needed through external validation, and where possible, recalibrate existing prediction models based on
   - Clinical characteristics only
   - Clinical and biochemical markers
   - Clinical and ultrasound markers
   - Clinical, ultrasound and biochemical markers
2. Using IPD meta-analysis, to develop and externally validate (using internal-external cross-validation) new multivariable prediction models for (i) fetal growth restriction with serious perinatal complications (model 1) and (ii) birth weight at various potential gestational ages at delivery (model 2).

Secondary

3. To compare the predictive performance of models according to (i) population (selected; unselected) (ii) trimester of testing (first <14 weeks; second ~20 weeks; third ~28 weeks) (iii) choice of predictors (clinical only; clinical and ultrasound; clinical and biochemical; clinical, ultrasound and biochemical and (iv) onset of fetal growth restriction (early <32 weeks; late >32 weeks).

4. To assess if the performance of the models are generalisable for various definitions of fetal growth restriction such as i) ultrasound parameters determined by Delphi consensus;[43] ii) and birth weight <10th centile adjusted for gestational age with associated neonatal morbidity.[44]
5. To assess the association between various birth weight centiles (<10th, <5th, < 3rd centiles) calculated using i) customised and ii) population-based standards, and perinatal mortality and morbidity.
6. To examine the clinical utility of the models using Decision Curve Analysis (DCA), which identifies whether there is an overall net-benefit for using models within clinical practice, i.e. to guide treatment decisions based on thresholds of predicted outcome risk.[42]

5. Work leading to the proposal

This proposal builds on the Programme of work supported by NIHR on improving perinatal outcomes such as: our large IPD IPPIC (International Prediction of Pregnancy Complications) funded by NIHR HTA to predict pre-eclampsia; KM's MRC Fellowship on individual predictors of fetal growth restriction and economic evaluation; GS's HTA funded study on the accuracy of universal late pregnancy ultrasound in diagnosing small for gestational age fetus in nulliparous women and association with adverse perinatal outcome with a health economic and a Value of Information analysis (Fig).

5.1 International Prediction of Pregnancy Complications (IPPIC) Network

Our HTA-funded International Prediction of Pregnancy Complication (IPPIC) Network of global researchers has access to the largest global IPD repository (Table 3). The IPPIC repository, supported by WHO (World Health Organization), was first established to develop IPD meta-analysis for prediction of pre-eclampsia (HTA No. 14/158/02).[45] Our collaborators have successfully obtained funding for further IPD meta-analyses using data in the repository to predict stillbirth (Sands Charity, UK) and spontaneous preterm birth (NHMRC, Australia). The repository holds IPD of over three million pregnancies (15 UK, 66 international datasets) with relevant predictor and outcome variables required to fulfil the objectives of IPPIC-FGR. Considerable resources have already been invested in this effort, which has led to most of the data needed for this proposal already mapped, cleaned, recoded and harmonised. We have categorised the datasets in the repository to be low and high quality using the PROBAST tool, thereby ensuring that only high quality studies or datasets are included in the analysis.

Access to clean IPPIC dataset will allow us to start validation of existing models within a short time frame, followed by as necessary, development and external validation of new prognostic models. We can, if needed, develop and externally validate new prognostic models. The proposal to access the IPPIC data has been approved by the data sharing committee and is supported by members of the IPPIC Network. A breakdown of the available analysable data in the IPPIC repository and the databases with the relevant predictors and outcomes is shown in Table 3.

Table 3: Studies and databases in the IPPIC repository

Primary Outcome	Predictors	No. of studies	No. of pregnancies
Fetal growth restriction complicated by stillbirth or neonatal death at any time or delivery before 32 weeks	Clinical only	42	1,239,396
	Clinical and biochemical	24	425,495
	Clinical and ultrasound	21	121,004
	Clinical, biochemical and ultrasound	17	108,499
Birth weight at various gestational ages	Clinical only	50	2,905,685
	Clinical and biochemical	27	427,757
	Clinical and ultrasound	27	123,390
	Clinical, biochemical and ultrasound	19	110,025

Study/database	Predictors				No of pregnancies	Birth weight and Gestational age of delivery ^	Stillbirth	Neonatal death	Neonatal morbidity
	Clinical	Clinical +Ultrasound	Clinical+ Biochemical	All					
AMND	*				141299	*	*	*	*
BORN	*		*		286718	*	*	*	*
Danish Birth Cohort	*				88020	*	*	*	*
Le Carpentie	*	*			245	*	*	*	*
MOBA	*				114744	*	*	*	*
NICHDHR 1998	*		*		3252	*	*	*	*
NICHDLR 1993	*				3171	*	*	*	*
Odibo 2011	*	*	*	*	1200	*	*	*	*
POP 2017	*	*	*	*	4212	*	*	*	*
Rumbold 2006	*				1877	*	*	*	*
SCOPE 2014	*	*	*	*	5628	*	*	*	*
TEST 2016	*	*	*	*	557	*	*	*	*
WHO 2015	*		*		7317	*	*	*	*
Baschat 201	*	*	*	*	9793	*		*	*
Llurba	*	*			11668	*		*	*
PARIS 2007	*				37341	*		*	*
Verlohren	*	*	*	*	566	*		*	*
Arenas 2003	*	*			319	*	*		*
BiB	*	*	*	*	13443	*	*		*
STORK 2010	*	*	*	*	823	*	*		*
Antsaklis 2000	*	*	*	*	3328	*	*	*	
Audibert 201	*	*	*	*	893	*	*	*	
Chei	*				406286	*	*	*	
Generation R	*	*	*	*	8824	*	*	*	
Ghana	*		*		1016	*	*	*	
Placental Health Study 2017	*	*	*	*	856	*	*	*	
Zhang 2001	*				1639	*	*	*	
Goffinet 200	*				3317	*		*	
STORK 2015	*	*	*	*	54677	*		*	
Vinter 2011	*				304	*		*	
Allen	*	*	*	*	1045	*	*		
ALSPAC	*		*		15444	*	*		
Andersen 2016	*	*	*	*	2161	*	*		
Conserva 2012	*				53	*	*		
Galindo	*	*	*	*	253	*	*		
Indonesian cohort	*				2281	*	*		

POUCH 2001	*		*		3019	*	*
Prefumo 200	*	*			273	*	*
Staff	*	*	*	*	240	*	*
Van Kuijk 2014	*		*		230	*	*
van Oostwaard 2012	*				425	*	*
van Oostwaard 2014	*				639	*	*
Chappell 1999	*	*	*	*	316	*	
Macleod 200	*	*			222	*	
Mbah 2012	*				1663167	*	
Ohkuchi 200	*	*			288	*	
Rang 2008	*	*			42	*	
Vatten 2007	*		*		736	*	
Velauthar	*	*	*	*	1210	*	
Vollebregt 2010	*	*			308	*	

^ *Information available to determine preterm birth before 32 weeks*

5.2 Systematic review of prediction models for fetal growth restriction

Our systematic review (Medline until 2012)[23] identified eighteen models for predicting fetal growth restriction or small for gestational age or low birth weight infants. Our updated search (Medline 2012 – 2018) done for this proposal identified ten additional prediction models for fetal growth restriction, and none were externally validated. Overall, 75% involved low-risk women, 7% unselected and 18% only high-risk groups. Only a third (10/28, 36%) were internally validated, and only two (2/28, 7%) were externally validated. Calibration measures were reported in 14% (4/28), and a prediction formula, rule or score was reported in sixteen (16/28, 57%) models. Only one model defined fetal growth restriction as a combination of small for gestational age fetus and perinatal complications.

5.3 Predictors of fetal growth restriction

Over the last two decades, the co-applicants undertook many studies to minimise offspring complications by predicting and managing small for gestational age, and growth-restricted fetuses.[25] However, until now, there is a lack of a co-ordinated effort to optimise evidence synthesis in this area, with suboptimal use of resources that have been invested in this area. In their roles as guideline developers (KM, JK) for the identification and management of small for gestational age fetuses with complications (RCOG UK, SOGC Canada),[18,19] the applicants have systematically reviewed the accuracy of predictors for fetal growth restriction.[22,46-48] The systematic review on biomarkers for small for gestational age fetus (81 studies, 382,005 women)[46] evaluated maternal serum alpha fetoprotein (AFP), human chorionic gonadotrophin (HCG), unconjugated estriol, pregnancy associated plasma protein A (PAPP-A), serum inhibin A and triple test (serum AFP, HCG and unconjugated estriol). Our findings highlighted the need for the development of prediction models incorporating clinical characteristics and uterine artery Doppler to increase the accuracy of risk assessment. Through our survey of 40 researchers, clinicians and members of the clinical speciality groups, we identified the most important predictors of fetal growth restriction.

5.4 Fetal growth restriction and complications

Fetal growth restriction is defined either in terms of fetal or birth weight parameters. Applicants (WG, BT, AP, SJG) published an international consensus statement using fetal growth and uterine artery Doppler measurement parameters to define early and late onset fetal growth restriction.[43]

Members (GS) previously used a birth weight of less than 10th centile to define small for gestational age fetus when evaluating the accuracy of third-trimester ultrasound as a screening test to detect infants at risk.[25] Fetal growth restriction is also defined as small for gestational age fetus with associated perinatal complications. The associated complications are varied (Table 4).

Our systematic review (59 studies, 2,600,383 individuals) showed poor association when arbitrary birth weight cut-offs were used (<2.5 Kg or as <10th centile) with childhood morbidity (OR 0.98, 95% CI 0.87, 1.1 and OR 1.49, 95% CI 1.02, 2.19 respectively).[49] We also did not observe any association with adult morbidity outcomes for these cut-offs. Similarly, a definition of small for gestational age (<10th centile) using either customised or population-based fetal standards, or population-based birth weight standards was not able to discriminate between infants with perinatal morbidity (ROC 0.56, 0.56, 0.55 respectively).[44]

These findings question the rationale of using these arbitrary cut offs to define small for gestational age as an outcome in any prediction model, leading to our choice of primary outcomes of birth weight, a continuous measure, at various gestational ages, and small baby with additional serious perinatal complications. As a secondary objective, we will assess the performance of the model for growth-restricted babies with neonatal morbidity, whose components will be informed from our ongoing core-outcome set (COSGROVE) development work for fetal growth restriction.[50]

5.5 Prediction model development and validation: IPD meta-analyses and primary studies

Our team has previously developed a detailed protocol for developing and validating prediction models for pre-eclampsia using IPD (HTA No. 14/158/02). The mapping of the predictor and outcome variables that was done for this project has given us a detailed insight into the quality of the datasets. We are confident that we have sufficient information to develop and validate models for fetal growth restriction. After completion of IPPIC-pre-eclampsia in December 2018, we are able to continue to accrue new datasets and clean and code them through our funding secured to predict stillbirth (May 2017- December 2018), and for preterm birth (January 2019-December 2022) using the IPD in the repository. Our experience in development of an integrated mobile and web App for the HTA funded models (HTA No. 09/22/163) to predict complications in women with early onset pre-eclampsia has enabled us to identify the requirements to develop similarly for IPPIC-FGR.[51]

6. Research Plan

Our IPD meta-analysis approach will follow existing guidelines for prediction model development and validation,[52-54] and our output will comply with the PRISMA statement and adhere to IPD meta-analysis and TRIPOD reporting guidelines.[55,56] We have the raw individual data from studies and datasets in the IPPIC data repository. These have already been standardised, cleaned, recoded and harmonised ready for use in our analysis. We will summarise the available evidence by data synthesis, while preserving the clustering of patients within the studies. We will address the structured questions presented in Table 4 in our project.

Table 4: Structured questions for IPD meta-analysis on prediction of birth weight and fetal growth restriction with complications

Question Components	
Population	Pregnant women
Predictors	Maternal clinical characteristics, biochemical markers, ultrasound markers (details in Table 1)
Outcomes	**Primary outcomes**

	Fetal growth restriction with severe complications (birth weight less than 10th centile adjusted for gestational age at delivery, complicated by stillbirth or neonatal death at any time or delivery before 32 weeks); Birth weight at various gestational ages **Secondary outcomes:** Early onset (<32 weeks) and late onset (≥32 weeks) fetal growth restriction Ultrasound based diagnosis for early (EFW<3rd centile, AC<3rd centile, absent EDF in umbilical artery Doppler) and late fetal growth restriction (EFW <3rd centile, AC <3rd centile) Neonatal morbidity: cord blood pH <7 at birth, hypoxic ischemic encephalopathy, respiratory distress syndrome, septicaemia, admission to neonatal unit, Apgar score <7 at 1' and 5'
Study design	IPD meta-analysis of observational studies and cohorts nested within randomised trials

EFW expected fetal weight; AC Abdominal circumference; EDF End Diastolic Flow

6.1 Primary outcome(s)

Our primary outcomes are (i) fetal growth restriction with severe complications (birth weight <10th centile adjusted for gestational age with stillbirth or neonatal death at any time or delivery before 32 weeks) (ii) birth weight for deliveries at various gestational ages to reflect the extent of the restricted growth.

Rationale for the choice of outcome(s)

Model 1. We chose fetal growth restriction with severe complications for the following reasons: the definition excludes small but healthy babies; the components of the composite include severe complications of mortality or extreme prematurity (both iatrogenic and spontaneous preterm births before 32 weeks are reflective of the severity of the condition). Any prediction model will need to take into consideration the effects of treatment paradox, where delivery could have prevented stillbirth or neonatal death that may have otherwise occurred.[57] We have addressed this by including delivery before 32 weeks as a component of the outcome.

Model 2: Until now, prediction models have used arbitrary cut-offs to define fetal growth restriction or small for gestational age fetus using only birth weight < 10th or < 3rd centile. Dichotomisation of the outcome limits the power and usefulness of a prediction model. Besides, the prognosis for a fetus with a predicted birth weight on the 3rd centile at 26 weeks is far worse than that predicted to be on the 9th centile at 37 weeks, despite both being labelled as small with less than 10th centile birth weight. A baby diagnosed to be small using a particular fetal growth standard (e.g., GROW, INTERGROWTH 21st, WHO) may not be categorised so with another standard, thereby limiting the generalisability of the model use. To address this, we will use birth weight as our outcome to be predicted at various potential gestational ages at delivery for the following reasons: it is a continuous measure not limited by arbitrary cut-offs; the predicted birth weight can still be converted into predicted centiles using any fetal growth standard in use; it provides information on both severity of the restricted growth, and the expected timing of onset to plan appropriate management. For e.g., a baby with a predicted birth weight on the 5th centile at 28 weeks gestation will require frequent monitoring starting from 26 weeks.

6.2. IPD meta-analysis

Our IPD meta-analysis will update the literature search, obtain IPD of the new studies, assess the quality of studies and datasets, complete data cleaning, coding and standardisation, and synthesise data.

6.2.1. Updating literature searches

Our current search (completed in March 2017) has identified the predictors for fetal growth restriction. We will update these for relevant reviews, and primary studies on prediction models for fetal growth restriction or birth weight, as new research evidence may have appeared since completion of our work. The following databases will be searched: MEDLINE, EMBASE, BIOSIS, LILACS, Pascal, Science Citation Index, Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment Database (HTA). Research reported in the grey literature will be sought by searching a range of relevant databases including the Inside Conferences, Systems for Information in Grey Literature (SIGLE), Dissertation Abstracts and Clinical Trials.gov. Internet searches will also be carried out using specialist search gateways (such as JISC: https://www.jisc.ac.uk/), general search engines (such as Google: http://www.google.co.uk/) and meta-search engines (such as Copernic: http://www.copernic.com/). Language restrictions will not be applied to the electronic searches. Identification of prognostic studies has been hindered by the lack of standard descriptors and indexing terms. We will overcome this by using search strategies with high sensitivity in identifying prognostic studies in Medline, such as exp epidemiologic studies OR incidence.sh OR prognos:.tw OR predict:.tw OR course:.tw along with terms specific to FGR.[58] This work will be completed during the pre-grant with our available resources.

6.2.2. Strengthening the IPPIC Collaboration

We will strengthen the existing IPPIC network by adding new researchers of studies identified in Section 6.2.1. The IPPIC network currently consists of 73 collaborators from 21 countries (15 UK datasets). The Network holds the largest repository of standardised pregnancy data (>3 million pregnancies) for key predictors of outcomes including pre-eclampsia, stillbirth, neonatal death, birth weight, neonatal morbidity, and gestational age at delivery. We will ensure that UK data are used in the external validation of the prediction models, which is similar to our current approach for developing IPPIC-pre-eclampsia prediction models.[59] The proposed project will bring together key researchers, guideline makers and clinicians involved in the efforts to identify early women at risk of fetal growth restriction. A buy-in from these individuals and groups is essential to ensure that the developed prediction models are applied in clinical practice.

All authors have signed a memorandum of understanding covering the provision of their data, and that any publication of the IPD meta-analysis will be in the name of the collaborative group. We maintain regular contact with collaborators via face-to-face meetings or teleconference. A Data Access Committee (DAC) that assesses any requests to access the data by evaluating the scientific quality of the research question, expertise of the research group, suitability of the IPPIC dataset to answer the research question, availability of required variables, and IT security for the shared IPPIC dataset has approved access to the IPPIC repository for this proposal.

6.2.3. Data collection, standardisation and quality assessment

i. Standardisation of data

We will add any data from new studies identified from our updated search, recode and harmonise these in line with the clean formatted IPPIC datasets, with rigorous range and consistency checks. If any new predictor variables have been considered to be important in existing datasets within the IPPIC repository, or in new studies, we will undertake the same process as above. Missing data, obvious errors, and inconsistencies between variables or outlying values will be queried and rectified with input from the original authors.

We will transform all new original data sets to Stata format, and generate data sets consisting of the essential variables (essential datasets). We will then verify the script corresponding to the datasets, and transform and harmonise the variables (core datasets). We have already customised

the web interfaces for the original, essential and core data sets, according to the levels of access needed by the collaborators and statisticians.[60]

ii. Quality assessment of the included studies and datasets

We will follow the methods used in our previous IPD meta-analyses to assess the quality of newly identified studies (and their IPD) using the PROBAST tool.[61] The tool assesses participant selection (adequate description of data sources, details on recruitment), predictors (appropriately defined, assessed blinded to outcome, assessed in the same way for all participants), and outcomes (appropriately defined and determined in a similar way for all participants, predictors excluded from the outcome definition, outcome determined without knowledge of predictor information and appropriate interval between assessment of predictor and outcome determination). The tool also evaluated the applicability of the studies or datasets. For each IPD, we will also assess the integrity of the data collection, ascertainment of the outcome and quality of the datasets using the PROBAST tool.

6.2.4. Data synthesis

Summarising the prognostic effect of individual predictors of fetal growth restriction

For each predictor and outcome separately, we will perform a two-stage IPD meta-analysis to obtain the summary effect of each predictor, firstly unadjusted and then adjusted for a pre-defined set of predictors identified from our systematic review and survey. For the birth weight outcome, the summary effect will be (adjusted) mean difference; for the binary fetal growth restriction with complications outcome, the summary effect will be an (adjusted) odds ratio. In the second stage of the meta-analysis, a random effects model will be used to account for unexplained heterogeneity between studies. Confidence intervals will be derived using the Hartung-Knapp approach (to account for uncertainty in variance estimates), and heterogeneity will be summarised by I-squared, tau-squared and 95% prediction intervals. One-stage models will be fitted as a sensitivity analysis, but as the data are large, one-stage and two-stage models should be similar.[62]

Assessing the performance of existing prediction models

We will use our IPD to evaluate the performance of the existing (and relevant) published models identified from our systematic review.[23] The external validation cohort will be from our large IPD database of the IPPIC Collaborative Network. For the external validation, we will quantify the predictive performance of each existing model in terms of calibration (using plots, E/O, calibration slopes), discrimination (using C statistic), overall fit (using R^2 and Brier score) and net-benefit.[63] We will assess the extent to which they need to be improved or tailored for the target population to predict birth weight at various gestational ages, and the risk of delivering a growth restricted infant with serious complications. For example, recalibration techniques will be considered.[53] This can include re-estimation of the intercept and adjustment of individual predictor weights (regression coefficients). Using these data, the probability of fetal growth restriction with complications, and the birth weight, for each individual patient in our validation cohort will be calculated. We expect the sample sizes to be adequate; as for such validation only one parameter (the linear predictor of the original model) is evaluated.[64] With at least 1000 babies classed as fetal growth restriction with serious complications, our available data far exceeds that of existing models. Missing data will be multiply imputed within studies and, as appropriate, systematically missing predictors imputed across studies using latest methods and guidelines.[54,65]

Improving the performance of prediction models

Using the relevant population from the IPD repository, we will develop, internally and externally validate new prediction models. A set of candidate predictors will be identified *a priori*, based on prior evidence and clinical judgment, with up to 50 predictor parameters. A suitable multivariable modelling framework will be chosen, such as logistic regression for the binary fetal growth restriction outcome, and linear regression for the continuous birth weight outcome. Intricate

modelling decisions will be pre-defined, such as the handling of continuous predictors (i.e. fractional polynomial modelling), selection of predictors (e.g. backwards selection), identification of non-linear trends, methods for dealing with missing data (i.e. multiple imputation), [37,66] dealing with heterogeneity and clustering in an IPD meta-analysis model. In model 2, gestational age at delivery will be included as a predictor, so that predictions of birth weight in new individuals can be made conditional on an assumed gestational age at delivery. Interactions between gestational age and other predictors will be examined. The final model will allow birth weight to be predicted for new individuals across a range of assumed delivery ages (rather than for just one gestational age). Shrinkage (penalisation) techniques (e.g. uniform shrinkage, the Lasso) to adjust for optimism in regression coefficients will also be incorporated to produce the final model, although – given the huge sample size – these are not likely to have much impact.

All the IPD will be used to develop the model, as data for model development are precious and splitting the data is not recommended.[67] However, our developed models will still be validated both internally and externally. Internal validation will be conducted using non-parametric bootstrapping, to examine overfitting and produce optimism-adjusted models and performance measures. Then, external validation (generalisability) of the model will be examined using internal-external cross-validation (IECV), as follows. Let there be IPD available from K studies. First, study 1 is excluded and the risk prediction model is developed using the remaining data (studies 2 to K). Study 1 can then be used to externally validate the model. This is then repeated, excluding study 2 rather than study 1 and fitting the model using studies 1 and 3 to K. Study 2 is then used to externally validate the model. This process is continually repeated, each time omitting a different study, until the model has been fitted excluding each study once. This process therefore provides K values for each validation statistic of interest, one for each excluded study. Random effects meta-analysis will then be used to summarise the performance across studies, to obtain summary measures of the model performance.

We will ensure that each cycle of the IECV approach retains sufficient sample size for model development. In this manner, each cycle will retain the majority of the available IPD for model development, and so the final models produced in each cycle are likely to be very similar to each other. A consistent model development strategy will be used in each cycle of the IECV approach. A wide range of validation statistics will be considered, focusing primarily on discrimination and calibration. Performance measures will include the C statistic, calibration slope, calibration in the large, and the Brier score for the fetal growth restriction with complications outcome, and the calibration slope and R^2 statistic for the birth weight outcome. Calibration plots will also be given to aid clinical interpretation. Optimism-adjusted performance statistics will be obtained (e.g. using bootstrapping), which adjust the apparent predictive performance to reduce bias due to overfitting, and produce estimates more likely to be observed when the models are applied to new individuals.

Examining heterogeneity and potential subgroup effects

The external validation performance will be investigated not just on average (i.e. the average across all IPD studies), but also in terms of the heterogeneity in performance across studies, settings, and clinically relevant subgroups (e.g. defined by treatment and populations). We will produce forest plots and meta-analyses for each validation statistic,[68] to summarise average performance and heterogeneity, with the latter reported in terms of I-squared, tau-squared, and a 95% prediction interval for the expected performance in a single population.[69,70] An ideal model will have little heterogeneity and consistently good performance. The sensitivity of external validation performance of the developed models will be checked in relation to the risk of bias. If required, these validated prediction models will be updated (or recalibrated) to different subgroups of women.

6.3. Sample size considerations for the new models

The effective sample size for the development and validation of prediction models is driven by the total number of events (for logistic regression of a binary outcome) or the total number of subjects (for linear regression of a continuous outcome). However, the number of subjects/events must be

large relative to the number of candidate predictors to be included in the model, to reduce the potential for overfitting and optimism.

Our first model to be developed will have the binary outcome of fetal growth restriction with serious perinatal complications. For logistic regression, sample size is highly debated: some authors recommend at least 10 outcome events per candidate predictor,[71] or at least 15,[52] or even at least 20.[72] We expect 10% of all pregnancies to be small, and 10% of small babies to suffer severe adverse perinatal outcome. Therefore, in our IPD meta-analysis, we expect the number of outcome events in our database (for a model with clinical, biochemical and ultrasound predictors) to be 108,499* 0.1*0.1 = 1084. Hence, even if we consider up to 50 predictor parameters, we have 1084/50 = 22 events per predictor parameter, thus far exceeding the aforementioned minimum requirements for sample size. For validation, during the internal-external cross-validation and bootstrapping approach, there will be well over 100 events in each sample, which is often recommended as a minimum for validation.[73]

The second model to be developed will have the continuous outcome of birth weight. Austin and Steyerberg recently recommended at least 2 subjects per candidate predictor for linear regression.[74] However, we have 110,025 subjects in our database for models incorporating clinical, biochemical and ultrasound predictors, and still many thousands when considering how many provide key predictors of interest. Thus, even if we considered up to 50 predictor parameters, our sample size is enormous. The sample sizes for both the primary outcomes are even higher (~ 3 million) when the models include only clinical, clinical and biochemical or clinical and ultrasound parameters (Table 3).

6.4. Decision curve analysis (DCA)

Decision curve analysis is a method for evaluating and comparing prediction models (in addition to the traditional validation measures of calibration and discrimination) in terms of their clinical utility i.e. whether one model offers greater net benefit than another when used to inform clinical decision making based on a threshold of predicted risk.[42] The net benefit of the model is plotted against different risk thresholds to produce a decision curve.[75] To obtain the curve, the prediction model is evaluated at different probability thresholds where the threshold is taken as a point above which a patient would be treated, and below which a patient would not be treated. The curve can then be compared to the treat all and treat no-one strategies to see the range of probabilities at which the model may be useful. Decision curves can also be plotted for different models on the same graph for comparison, and to help decide which model offers the most benefit. The model with the highest curve (over a range of thresholds) is considered to have the greatest net benefit.

6.5. Health economic evaluation and Decision Analytic Modelling

A. Economic evaluation using the NICE model for monitoring fetal growth

We will conduct a cost-effectiveness analysis comparing the costs and outcomes of the various IPPIC FGR prediction models for fetal growth restriction (FGR) with perinatal complications (model 1) and small for gestational age (SGA) fetus (model 2). We will build upon previous work, using the economic model structure and care pathways for monitoring fetal growth which was published as part of the NICE clinical guideline on antenatal care (No. 62).[76] The economic model will be in the form of a decision tree from an NHS and personal social services perspective.

The economic model will compare the existing strategies in the NICE 2008 Antenatal Care guideline (strategies 1-3), with the strategies for monitoring fetal growth using the IPPIC FGR prediction models (strategies 4 and 5):

1. No measurement or monitoring of fetal growth
2. Measure and monitor growth of all fetuses by ultrasound
3. Measure and monitor growth of all fetuses by symphysis-fundal height (SFH) measurement and ultrasound
4. Measure and monitor fetal growth by ultrasound according to risk predicted by **IPPIC FGR model 1** for fetal growth restriction with perinatal complication

5. Measure and monitor fetal growth by ultrasound according to risk predicted by **IPPIC FGR model 2** for small for gestational age (SGA) fetus

The strategies will be updated if new information emerges from the literature reviews. The clinical effectiveness data inputs for the decision tree will be informed by the IPD from the IPPIC Collaborative network,[77] existing economic models,[78]ongoing cost effectiveness work on routine ultrasound for fetal growth in the third trimester[79] and the updated literature reviews. This information will include but not limited to: prevalence and incidence rates; accuracy of ultrasound scanning (true positives, true negatives, false positives and false negatives); the mode of delivery (birth); the number of neonatal deaths (including stillbirths); and number of neonatal admissions. We will be guided by our clinical experts (assumptions), and where necessary we will use information from the economic model on test treatment strategies for prevention of fetal growth restriction.[78]

Resource use and costs will include: routine ultrasound scans and any extra scans; monitoring appointments; any tests/treatments (to predict fetal growth restriction or SGA); costs of labour differentiated by type of delivery (normal or a caesarean section); costs of neonatal intensive care admissions; and any post-mortem or litigation costs. Unit costs will be attached to each resource item to generate an overall cost per patient and will be obtained from routine sources such as NHS reference costs[80] and Unit Costs of Health and Social Care.[81] Costs will be updated to 2016/17 prices using the latest inflation indices.[81]

The primary outcome will be quality-adjusted life years (QALYs) and a secondary outcome will be the number of cases where fetal growth restriction severe complication is avoided. Results will be expressed as cost per QALY quality-adjusted life year gained and cost per fetal growth restriction severe complication was avoided. As cost data is skewed we will use non-parametric bootstrapping to produce 95% confidence intervals around the mean cost estimate.[82] Any future costs and outcomes will be discounted at 3.5% per annum. We shall use sensitivity analyses to explore the robustness of these results and to consider the broader issue of the generalisability.

B. Resource Impact Assessment

We will assess the potential impact of resource use and cost savings to the NHS of different strategies to identify and monitor babies with fetal growth restriction and complications; and small for gestational age babies. Using the templates provided by NICE for the clinical guideline on antenatal care (number 62)[76] and the inputs obtained from both the IPD from the IPPIC Collaborative network [77] and the updated literature reviews we will estimate the budget impact to the NHS.

7. Potential difficulties and solutions

From our previous experiences in undertaking large IPD meta-analyses,[83] we are aware of the potential hurdles in this project. A major challenge in any IPD meta-analysis is to persuade the primary study authors to provide their raw data, and we have already achieved it.[34] Authors who have developed prediction models for fetal growth restriction have also been invited to join the IPPIC Network. The other difficulty is obtaining timely responses (sometimes multiple times) from collaborators on clarification of the data and study details, and additional relevant data that may not have been published. We have ensured that sufficient time is planned into the proposal to accommodate any delayed response.

Our project involves multiple datasets acquired from studies with varied study design. The IPD meta-analysis will take this into account with adequate steps to standardise the data, and ensure that the findings are applicable to the relevant population. Multiple imputation will be used to handle missing data, including systematically missing variables across studies using new methods.[38] A significant amount of time and resources are spent in cleaning, coding and standardising the data. We have already completed this process. We have established the publication and data access committees to ensure that all collaborators are adequately acknowledged in any outputs to allay concerns of primary investigators in sharing the data.

8. Dissemination, outputs and anticipated impact

Dissemination of research findings is a key responsibility of the researcher. Apart from it being an ethical obligation, dissemination of the results to the following groups is necessary to facilitate rapid translation of relevant findings into clinical care where appropriate.

Funder: The findings will be provided as a detailed report to NIHR and other relevant agencies that fund the work. Any outputs as scientific publications, presentations and websites will highlight the support provided by the NIHR.

Scientific papers and presentations: Every effort will be made to ensure that the studies are conducted and reported with the highest standard necessary for publication in high impact journals. The findings will be disseminated to peers and experts through presentations in relevant specialty conferences and network meetings and through peer reviewed and practitioner publications according to the NIHR open access policy.

Patient and Public: A regular newsletter will be sent to the collaborators updating and highlighting the work. We will also liaise closely with Katie's Team and Sands charity and other interested groups regarding the dissemination of the findings of the analysis

Websites: The details of the project and findings will be provided through the institutional websites of the collaborators. Additionally, websites dedicated to collaborative endeavours such as GONET and CoLab will be a vehicle for dissemination.

Mainstream and Social media: QMUL has an active press department to facilitate the research findings to the public by staging press releases that are relevant, factual and informative. The increasing integration of social media in our day-to-day lives will be exploited to effectively disseminate the findings through applications such as Twitter and LinkedIn.

Professional Societies: Through existing links, the findings will be disseminated to the Association of Medical Royal Colleges (AoMRC), Royal College of Obstetricians and Gynaecologists (RCOG), Royal College of Midwifery and Nursing (RCM), Royal College of General Practitioners (RCGP) and British Maternal Fetal Medicine Society (BMFMS).

Expected impact of predicting fetal growth restriction

Clinical

In women who are predicted to be at risk of delivering a growth-restricted fetus, serial testing with ultrasound (for fetal growth, umbilical artery blood flow Doppler, and amniotic fluid volume) can detect the condition early and assess its severity. Stillbirth rates are halved when the diagnosis of fetal growth restriction is known (9.7/1000) vs. not known (19.8/1000).[5] A diagnosis of fetal growth restriction made relatively early in pregnancy (26-32 weeks) with subsequent close monitoring reduced perinatal death rates (8% vs. 11%); seven in ten of these infants (345/490, 70%) survived without severe morbidity.[84] At term (≥37 weeks), the diagnosis can prevent stillbirths and severe morbidity including HIE (Hypoxic Ischaemic Encephalopathy) by not continuing the pregnancy beyond 41 weeks, and by offering early delivery instead.[85] Induction of labour at term reduces perinatal complications, and lowers the rates of caesarean sections compared with expectant management.[86]

National and international guidelines

Our effort to predict fetal growth restriction resonates with the priorities of the Department of Health and RCOG to reduce stillbirths and neonatal deaths. In addition to contributing to the large IPPIC database, the project provides a platform for leading global researchers in the field of fetal growth restriction to develop and adopt unified prediction models. By including clinical academics involved

in the development of the forthcoming RCOG Green Top national guideline on small for gestational age fetus, and by obtaining the support of the RCOG fetal medicine clinical study group (CSG), we expect the model to be incorporated within national and international recommendations.

Cost

Improved prediction of growth restriction has the potential to save costs for the NHS.[87] In a maternity unit caring for 1000 women, the estimated cost savings include: prevention of stillbirths (£5000/y), reduced neonatal intensive care admission (£20,000/y), HIE (£25,000/y), and litigation costs (£70,000/y). About 60% of NHS Litigation Authority cases involving stillborn fetuses had a missed antenatal diagnosis of fetal growth restriction.[88] The total potential savings from early identification of fetuses at risk of growth restriction is £120,000/y for every 1000 births – equivalent to £100 million per year for all births in the UK.[87,89]

9. Contribution to Collective Research Effort

Fetal growth restriction continues to be a research priority area. This is in part from ongoing efforts to reduce stillbirth and adverse perinatal outcomes at term, and from works to reduce prematurity related complications. Since fetal growth restriction is commonly attributed to placental insufficiency, research on this topic is also complementary to efforts on predicting (HTA No.14/158/02) and preventing pre-eclampsia, which is associated with placental diseases.

The NICE guidelines and RCOG have called for predictive tests or strategies to identify women at risk of small baby, particularly for growth-restricted infant with complications.[18,31] Studies on the accuracy of tests in the third trimester are focussed on the detection of small for gestational age fetus, with limited sample sizes to detect fetal growth restriction complicated by stillbirths and neonatal deaths (POP study).[25] Our proposal will focus in this area, and develop and assess the performance of a prediction model for growth restriction with complications, which has not been assessed to-date due to lack of sample size.

In recent years there have been considerable advances in the evaluation of individual tests, including novel biomarkers and ultrasound techniques to identify women at risk of having a small baby with or without complications. Despite these advances, early identification of mothers who are at risk of true fetal growth restriction with severe offspring complications still poses considerable challenges. In addition to the large sample sizes needed for such a study, a lack of agreement between various teams on the fetal or birth weight standards used to define growth-restricted or small for gestational age fetus has limited progress. In the UK, some maternity units use the GAP (Growth Assessment Protocol) tool with customised fetal growth standards, while others use population-based charts. An ongoing cluster randomised trial (DESiGN trial) of the two strategies is expected to identify optimal method of monitoring a small baby. By predicting the actual birth weight for various gestational ages, our model provides the flexibility to convert the outcome into birth weight centiles using the growth standard chart that is found to be most effective.

GS's HTA funded study (No. 15/105/01) on universal late pregnancy ultrasound in nulliparous women will provide health economic data which is relevant for the current proposal, as it will establish the economic case for future large scale trials of screening and intervention and the health economic case for implementing screening. The NIHR EME commissioned ongoing primary study (SPREE) on the prediction of pre-eclampsia also reports on birth weight, and fetal growth restriction as secondary outcomes. We will work closely with this team to ensure that the research output benefits from incorporation of the study findings within our IPD framework. If this is not feasible within the lifetime project, we will ensure that steps are taken to enable such a plan in the future, by standardising the relevant databases.

Identification of subsets of women who are at high risk of fetal growth restriction, will allow us to evaluate whether targeted management of these high groups, will improve the outcome. We will liaise with the TRUFFLE Collaborative group, who are partners in this proposal, on optimal management of early onset fetal growth restriction.

Our economic evaluation of fetal growth restriction will be built of existing and ongoing work in this area: NICE economic model for monitoring fetal growth published as part of NICE Antenatal care guideline 2008; decision analytic model on prediction and prevention of fetal growth restriction, and ongoing HTA project 'Value of undertaking a study to determine the clinical and cost effectiveness of late pregnancy ultrasound to prevent adverse perinatal outcome in nulliparous women' (15/105/01). The cost implications of using the IPPIC FGR models will be built on NICE work on Resource Impact Assessment.

Many networks such as the Co-Lab, SCOPE, STORK and PRE-EMPT have already joined forces in the standardisation of the definitions, collection of relevant datasets and outcomes. Our proposal is collaborative and complementary to the above efforts and involves the leading researchers active in the above endeavours. Continued growth and refinement of the central IPPIC repository will enable future work to be seamlessly continued, by incorporating emerging data on new biomarkers, and updating the developed models beyond the lifetime of the project.

10. Project timetable

We estimate the project duration to be 17 months. Since the repository is currently live with funding for data management, during the pre-grant phase we plan to deposit the data from the ten new studies that are identified. We will devote the first 3 months of the grant to further update the search, clean, code and standardise the new data, develop the protocol, establish collaborators workshop and finalise the standardisation of datasets; next 3 months for external validation of the models using our IPPIC database; 6 months for model development and IECV, and decision curve analysis; 2 months for economic evaluation and 3 months for write-up of the HTA report.

	Quarter/Year	Q4 - 2018		Q1 - 2019			Q2 - 2019			Q3 - 2019			Q4 - 2019			Q1 - 2020			Q2 - 2020			Q3 - 2020		
	Project Month	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	Calendar Month	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep
1	Pre grant work (add relevant new data to the IPPIC repository, 10 new studies)																							
2	HTA Grant																							
3	Protocol development																							
4	Update literature search																							
5	Establish collaborator workshops																							
6	Standardise new datasets																							
7	External validation of published models																							
8	Model development, IECV and DCA																							
9	Health economic evaluation																							
10	Write up and report production																							

11. Project management

Queen Mary University of London (QMUL) will be the sponsor and host organisation, with ST as the Chief Investigator. Subcontracts will be put in place between QMUL and Institutions of the co-applicants on data sharing, responsibilities and the expected contributions of each party. We already have agreements in place to access the datasets, and we will get the collaborators to reconfirm the use of their data for this particular project. The IPPIC Data Access Committee have reviewed the proposal for this project and given their support and approval to access the IPPIC data.

The CI is responsible for conduct of the project and decision-making. All staff will share the same duty of care to prevent unauthorised disclosure of personal information and act according to the Data Protection Act 1998 & Good clinical governance. A project management group (PMG) will manage the work with monthly meetings. An Independent Project Steering Committee will provide overall supervision and ensure adherence to Research Governance framework and GCP Guidelines. The Project Steering Committee will meet three times and will include at least one PPI member.

12. Ethics

The current project is an evidence synthesis project involving meta-analysis of anonymised datasets. No further ethical considerations or approvals are needed for this project.

13. Patient and Public Involvement

The Katherine Twining Network's Katie's Team has been actively involved in the preparation of this application and is part of the research team (co-applicant) undertaking the project. The group includes mothers, pregnant women, carers and family members with an interest in improving the quality of research within women's health. Katie's Team members contributed to the fine-tuning of the primary outcomes of this proposal by providing feedback on what they would consider to be an important outcome.

A PPI member will provide input though participation in Steering Committee meetings. Development of the prediction model will take into account the input from service users on the acceptability of the various predictors being evaluated. Furthermore, they will provide input on the value of various risk scores to the women. Their involvement will ensure that the prediction strategy takes into account the opinion of mothers, and ensure acceptability in clinical practice. Katie's team will contribute to study reports and help in the dissemination of the findings. We will also liaise closely with Sands and other interested groups and obtain input.

References

1. Draper ES, Kurinczuk JJ, Kenyon S (Eds.) on behalf of MBRRACE-UK. MBRRACE-UK 2017 Perinatal Confidential Enquiry: Term, singleton, intrapartum stillbirth and intrapartum-related neonatal death. The Infant Mortality and Morbidity Studies, Department of Health Sciences, University of Leicester: Leicester, 2017.
2. Flenady V, Middleton P, Smith GC, et al. Stillbirths: the way forward in high-income countries. *The Lancet* 2011; **377**(9778): 1703-17.
3. NHS England: Saving Babies' Lives: A care bundle for reducing stillbirth https://www.england.nhs.uk/wp-content/uploads/2016/03/saving-babies-lives-car-bundl.pdf.
4. Froen JF, Gardosi JO, Thurmann A, Francis A, Stray-Pedersen B. Restricted fetal growth in sudden intrauterine unexplained death. *Acta obstetricia et gynecologica Scandinavica* 2004; **83**(9): 801-7.
5. Gardosi J, Madurasinghe V, Williams M, Malik A, Francis A. Maternal and fetal risk factors for stillbirth: population based study. *BMJ* 2013; **346**: f108.
6. Vayssiere C, Sentilhes L, Ego A, et al. Fetal growth restriction and intra-uterine growth restriction: guidelines for clinical practice from the French College of Gynaecologists and Obstetricians. *European journal of obstetrics, gynecology, and reproductive biology* 2015; **193**: 10-8.
7. Dall'Asta A, Brunelli V, Prefumo F, Frusca T, Lees CC. Early onset fetal growth restriction. *Matern Health Neonatol Perinatol* 2017; **3**: 2.
8. Hepburn M, Rosenberg K. An audit of the detection and management of small-for-gestational age babies. *British journal of obstetrics and gynaecology* 1986; **93**(3): 212-6.
9. Figueras F, Gardosi J. Intrauterine growth restriction: new concepts in antenatal surveillance, diagnosis, and management. *American journal of obstetrics and gynecology* 2011; **204**(4): 288-300.
10. S MK, Gardosi J. Perinatal mortality and fetal growth restriction. *Best practice & research Clinical obstetrics & gynaecology* 2004; **18**(3): 397-410.
11. Divon MY, Haglund B, Nisell H, Otterblad PO, Westgren M. Fetal and neonatal mortality in the postterm pregnancy: The impact of gestational age and fetal growth restriction. *American journal of obstetrics and gynecology* 1998; **178**(4): 726-31.
12. Barker DJ, Osmond C, Golding J, Kuh D, Wadsworth ME. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. BMJ : British Medical Journal. 1989;298(6673):564-567.
13. Lindqvist PG, Molin J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome? *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2005; **25**(3): 258-64.
14. New ambition to halve rate of stillbirths and infant deaths. Department of Health and Social Care and The Rt Hon Jeremy Hunt MP. https://www.gov.uk/government/news/new-ambition-to-halve-rate-of-stillbirths-and-infant-deaths.
15. NHS England: Our Business Plan for 2016/17. https://www.england.nhs.uk/wp-content/uploads/2016/03/bus-plan-16.pdf.
16. Clinical Indicators Team, Health and Social Care Information Centre. NHS Outcomes Framework: May 2016 Quarterly Publication V1.0. Copyright © 2016, Health and Social Care Information Centre. All rights reserved. file:///C:/Users/hmx108/Downloads/nhs-out-fram-ind-may-16-comm2.pdf.
17. James Lind Alliance, Priority Setting Patnerships. Stillbirth Priority 5 from the Stillbirth PSP. JLA question ID 0029/5. 2 February 2015. http://www.jla.nihr.ac.uk/priority-setting-partnerships/stillbirth/top-10-priorities/priority-5-from-the-stillbirth-psp.htm.

18. Royal College of Obstetricians and Gynaecologists, The Investigation and Management of the Small–for–Gestational–Age Fetus - RCOG Green-top Guideline No. 31 : 2nd Edition | February 2013 | Minor revisions – January 2014. https://www.rcog.org.uk/globalassets/documents/guidelines/gtg_31.pdf.
19. Lausman A, Kingdom J, Gagnon R, et al. Intrauterine Growth Restriction: Screening, Diagnosis, and Management. *Journal of Obstetrics and Gynaecology Canada* 2013; **35**(8): 741-8.
20. American College of O, Gynecologists. ACOG Practice bulletin no. 134: fetal growth restriction. *Obstet Gynecol* 2013; **121**(5): 1122-33.
21. The Royal Australian and New Zealand College of Obstetricians and Gynaecologists. Screening in Early Pregnancy for Adverse Perinatal Outcomes. Prenatal Screening for Adverse Pregnancy Outcomes. C-Obs 61 (July 2015). https://www.ranzcog.edu.au/RANZCOG_SITE/media/RANZCOG-MEDIA/Women's%20Health/Statement%20and%20guidelines/Clinical-Obstetrics/Screening-in-Early-Pregnancy-for-Adverse-Perinatal-Outcomes-(C-Obs-61)-New-July-2015.pdf?ext=.pdf.
22. RK M. Prediction and prevention of fetal growth restriction and compromise of fetal wellbeing. Systematic reviews and Meta-analyses with model based economic evaluation. A thesis submitted to the University of Birmingham for the degree of Doctor of philosophy. College of Medical and Dental Sciences The University of Birmingham September 2010 Volume I http://etheses.bham.ac.uk/1319/1/Morris11PhD.pdf. 2010.
23. Kleinrouweler CE, Cheong-See Mrcog FM, Collins GS, et al. Prognostic models in obstetrics: available, but far from applicable. *American journal of obstetrics and gynecology* 2015.
24. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine* 2015; **162**(1): W1-73.
25. Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *The Lancet* 2015; **386**(10008): 2089-97.
26. World Health Organisation (WHO), Children's environmental health indicators. Intrauterine growth retardation in newborn children http://www.who.int/ceh/indicators/iugrnewborn.pdf.
27. Zhang X, Platt RW, Cnattingius S, Joseph KS, Kramer MS. The use of customised versus population-based birthweight standards in predicting perinatal mortality. *BJOG : an international journal of obstetrics and gynaecology* 2007; **114**(4): 474-7.
28. McCowan LM, Harding JE, Stewart AW. Customized birthweight centiles predict SGA pregnancies with perinatal morbidity. *BJOG : an international journal of obstetrics and gynaecology* 2005; **112**(8): 1026-33.
29. Seed PT, Chappell LC, Black MA, et al. Prediction of preeclampsia and delivery of small for gestational age babies based on a combination of clinical risk factors in high-risk women. *Hypertension in pregnancy* 2011; **30**(1): 58-73.
30. Bricker L, Medley N, Pratt JJ. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *The Cochrane database of systematic reviews* 2015; (6): CD001451.
31. National Institute for Health and Care Excellence (NICE), Antenatal care for uncomplicated pregnancies. Clinical guideline [CG62]. Published March 2008. ISBN: 978-1-4731-0891-2. nice.org.uk/guidance/cg62.
32. Monier I, Blondel B, Ego A, Kaminiski M, Goffinet F, Zeitlin J. Poor effectiveness of antenatal detection of fetal growth restriction and consequences for obstetric management

and neonatal outcomes: a French national study. *BJOG : an international journal of obstetrics and gynaecology* 2015; **122**(4): 518-27.
33. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS medicine* 2013; **10**(2): e1001380.
34. Abo-Zaid G, Sauerbrei W, Riley RD. Individual participant data meta-analysis of prognostic factor studies: state of the art? *BMC medical research methodology* 2012; **12**: 56.
35. Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. *Statistics in medicine* 2016; **35**(17): 2938-54.
36. Resche-Rigon M, White IR. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. *Stat Methods Med Res* 2016.
37. Resche-Rigon M, White IR, Bartlett JW, Peters SA, Thompson SG, Group P-IS. Multiple imputation for handling systematically missing confounders in meta-analysis of individual participant data. *Statistics in medicine* 2013; **32**(28): 4890-905.
38. Vincent Audigier, Ian R. White, Shahab Jolani, et al. Multiple imputation for multilevel data with continuous and binary variables. https://arxiv.org/pdf/1702.00971.pdf. February 6, 2017.
39. Jolani S, Debray TP, Koffijberg H, van Buuren S, Moons KG. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. *Statistics in medicine* 2015; **34**(11): 1841-63.
40. Ahmed I, Debray TP, Moons KG, Riley RD. Developing and validating risk prediction models in an individual participant data meta-analysis. *BMC medical research methodology* 2014; **14**: 3.
41. Pennells L, Kaptoge S, White IR, Thompson SG, Wood AM, Emerging Risk Factors C. Assessing risk prediction models using individual participant data from multiple studies. *American journal of epidemiology* 2014; **179**(5): 621-32.
42. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; **352**: i6.
43. Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2016; **48**(3): 333-9.
44. Sovio U, Smith GCS. The effect of customization and use of a fetal growth standard on the association between birthweight percentile and adverse perinatal outcome. *American journal of obstetrics and gynecology* 2018; **218**(2S): S738-S44.
45. Allotey J, Snell KIE, Chan C, et al. External validation, update and development of prediction models for pre-eclampsia using an Individual Participant Data (IPD) meta-analysis: the International Prediction of Pregnancy Complication Network (IPPIC pre-eclampsia) protocol. *Diagnostic and Prognostic Research* 2017; **1**(1).
46. Morris RK, Cnossen JS, Langejans M, et al. Serum screening with Down's syndrome markers to predict pre-eclampsia and small for gestational age: systematic review and meta-analysis. *BMC pregnancy and childbirth* 2008; **8**: 33.
47. Morris RK, Khan KS, Coomarasamy A, Robson SC, Kleijnen J. The value of predicting restriction of fetal growth and compromise of its wellbeing: Systematic quantitative overviews (meta-analysis) of test accuracy literature. *BMC pregnancy and childbirth* 2007; **7**: 3.
48. Morris RK, Bilagi A, Devani P, Kilby MD. Association of serum PAPP-A levels in first trimester with small for gestational age and adverse pregnancy outcomes: systematic review and meta-analysis. *Prenat Diagn* 2017; **37**(3): 253-65.
49. Malin GL, Morris RK, Riley RD, Teune MJ, Khan KS. When is birthweight at term (>/=37 weeks' gestation) abnormally low? A systematic review and meta-analysis of the prognostic

and predictive ability of current birthweight standards for childhood and adult outcomes. *BJOG : an international journal of obstetrics and gynaecology* 2015; **122**(5): 634-42.
50. Core Outcome Set for GROwth restriction: deVeloping Endpoints (COSGROVE) COMET Initiative http://www.comet-initiative.org/studies/details/689.
51. PREP Apps. Prediction of Risks of Complications in Early Onset Pre-eclampsia http://www.prepapp.qmul.ac.uk/.
52. Harrell FE. Regression modeling strategies, with applications to linear models, logistic regression, and survival analysis (2nd edition). New York: Springer; 2015.
53. Riley RD, Ensor J, Snell KI. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016; **353**.
54. Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Statistics in medicine* 2013; **32**(18): 3158-80.
55. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535.
56. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Annals of internal medicine* 2015; **162**(1): 55-63.
57. Cheong-See F, Allotey J, Marlin N. Prediction models in obstetrics: understanding the treatment paradox and potential solutions to the threat it poses. *BJOG : an international journal of obstetrics and gynaecology* 2016; **123**.
58. Systematic Reviews. CRDs guidance for undertaking systematic reviews in healthcare. Centre for Reviews and Dissemination, University of York, 2009.
59. Screening programme for pre-eclampsia. ISRCTN83611527 https://doi.org/10.1186/ISRCTN83611527.
60. Accuracy of clinical characteristics, biochemical and ultrasound markers in the prediction of preeclampsia: an Individual Patient Data (IPD) Meta-analysis https://kamolo.org.ar/ippicipd/.
61. PROBAST: a risk of bias tool for prediction modelling studies. The Cochrane Colloquium Vienna. http://2015.colloquium.cochrane.org/abstracts/probast-risk-bias-tool-prediction-modelling-studies. Accessed 27 Mar 2017.
62. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Statistics in medicine* 2017; **36**(5): 855-75.
63. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009; **338**: b605.
64. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *Journal of clinical epidemiology* 2005; **58**(5): 475-83.
65. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine* 2011; **30**(4): 377-99.
66. Burgess S, White IR, Resche-Rigon M, Wood AM. Combining multiple imputation and meta-analysis with individual participant data. *Statistics in medicine* 2013; **32**(26): 4499-514.
67. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models. *Journal of clinical epidemiology* 2001; **54**(8): 774-81.
68. Pennells L, Kaptoge S, White IR, Thompson SG. Wood AM and Emerging Risk Factors C. Assessing risk prediction models using individual participant data from multiple studies. *American journal of epidemiology* 2014; **179**.
69. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**.

70. van Klaveren D, Steyerberg EW, Perel P, Vergouwe Y. Assessing discriminative ability of risk models in clustered data. *BMC medical research methodology* 2014; **14**: 5.
71. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology* 1996; **49**(12): 1373-9.
72. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of clinical epidemiology* 2016; **76**: 175-82.
73. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Statistics in medicine* 2016; **35**(2): 214-26.
74. Austin PC, Steyerberg EW. The number of subjects per variable required in linear regression analyses. *Journal of clinical epidemiology* 2015; **68**(6): 627-36.
75. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; **26**(6): 565-74.
76. National Collaborating Centre for Women's and Children's Health (UK). Antenatal Care: Routine Care for the Healthy Pregnant Woman. London: RCOG Press; 2008 Mar. (NICE Clinical Guidelines, No. 62.).
77. HTA funded IPPIC-FGR study.
78. Morris RK. Prediction and prevention of fetal growth restriction and compromise of fetal wellbeing. Systematic reviews and meta-analyses with model based economic evaluation. University of Birmingham Thesis. 2010.
79. HTA project Value of undertaking a study to determine the clinical and cost effectiveness of late pregnancy ultrasound to prevent adverse perinatal outcome in nulliparous women (15/105/01)
80. Department of Health. NHS reference costs 2015 to 2016. 2016; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.
81. Curtis, L. and A. Burns, Unit Costs of Health and Social Care 2016, Kent: Personal Social Services Research Unit.
82. Manly, B.F., Randomization, bootstrap and Monte Carlo methods in biology. Vol. 70. 2006: CRC Press.
83. International Weight Management in Pregnancy Collaborative G. Effect of diet and physical activity based interventions in pregnancy on gestational weight gain and pregnancy outcomes: meta-analysis of individual participant data from randomised trials. *BMJ* 2017; **358**: j3119.
84. Lees C, Marlow N, Arabin B, et al. Perinatal morbidity and mortality in early-onset fetal growth restriction: cohort outcomes of the trial of randomized umbilical and fetal flow in Europe (TRUFFLE). *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* 2013; **42**(4): 400-8.
85. Gardosi J, Giddings S, Buller S, Southam M, Williams M. Preventing stillbirths through improved antenatal recognition of pregnancies at risk due to fetal growth restriction. *Public Health* 2014; **128**(8): 698-702.
86. 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* 2012; **344**: e2838.
87. Gardosi J, Williams M. Serial ultrasound scanning in pregnancies at risk of intrauterine growth restriction. Cost Benefit Analysis. Perinatal Institute www.perinatal.org.uk. http://www.perinatal.org.uk/FetalGrowth/GAP/Resources/RCOG_serial_scans_cost_benefit_analysis.pdf 18 August 2014.
88. Symon A. NHS Litigation Authority report on claims involving stillbirths. *British Journal of Midwifery* 2010; **18**(9): 598-9.

89. Mistry H, Heazell AE, Vincent O, Roberts T. A structured review and exploration of the healthcare costs associated with stillbirth and a subsequent pregnancy in England and Wales. *BMC pregnancy and childbirth* 2013; **13**: 236.