IRIS study - The IUGR RIsk Selection study

IRIS study - The IUGR RIsk Selection study

Project team

Ank de Jonge, project leader	PhD, Midwife researcher | Department of Midwifery Science, AVAG / EMGO+, VU University Medical Center, Amsterdam
Petra Jellema, project coordinator	PhD, Epidemiologist | Department of Midwifery Science, AVAG / EMGO+, VU University Medical Center, Amsterdam
Viki Verfaille, junior-researcher	MSc, Midwife researcher | Department of Midwifery Science, AVAG / EMGO+, VU University Medical Center, Amsterdam
Henriette van der Horst, supervisor	MD, PhD, Professor of General Practice | Department of General Practice, EMGO+, VU University Medical Center, Amsterdam
Arie Franx, supervisor	MD, PhD, Professor of Obstetrics | Department of Obstetrics, UMCU, Utrecht

Steering committee

Sabina Ledda, advisor	Ultrasonographer, chair of BEN ('Beroepsvereniging Echoscopisten Nederland')
Maya Kruijt, advisor	Representative of Dutch consortium for studies in women's health and reproductivity | Consortium Trialbureau, AMC, Amsterdam
Eva Pajkrt, advisor	MD, PhD, perinatologist | Department of Obstetrics and Gynaecology, Academic Medical Centre, Amsterdam, The Netherlands
Judith Bosmans, advisor	PhD, Assistant Professor in Health Technology Assessment | Department of Health Sciences of the VU University and the Department of Epidemiology and Biostatistics of the VU University Medical Center, Amsterdam
Marlou de Kroon, advisor	MD, PhD, Child healthcare physician and chair of scientific committee AJN | Erasmus Medical Center, Rotterdam
Trudy Klomp, advisor	MSc, Midwife-researcher and representative of MRNN | Department of Midwifery Science, AVAG - EMGO+, VU University Medical Center, Amsterdam

Members national IRIS-study group

Midwives
P. de Reu, Boxtel
M. Nieuwenhuijze, head of Research Department Midwifery Science, Midwifery Academy Maastricht
M. Rijnders, senior midwife researcher, TNO
A. de Roon, department of guideline development, KNOV
H. Torij, associate lecturer Pregnancy & Childbirth, Midwifery Academy Rotterdam

Obstetricians
M. Haak, president working group fetal ultrasound, NVOG, LUMC Leiden
J. Bais, Alkmaar
K. Bloemenkamp, Leiden
M. Paarlberg, Apeldoorn, president of the International Society of Psychosomatic Obstetrics and Gynaecology
J. van Dillen, Nijmegen
M. Spaanderman, MUMC J.
van Eyck, Zwolle

Professor obstetricians
J. van der Post, AMC
C. de Groot, VUmc
F. Vandenbussche, Nijmegen
P. P. van den Berg, UMCG J.
Nijhuis, MUMC
J. G. van Roosmalen, VUmc
G. Oei, Veldhoven J.
van Lith, LUMC E.
Steegers, EMC
S. Scherjon, UMCG, DIGITAT study

Neonatologists
S. Veen, LUMC
H. Lafeber, professor of neonatology, VUmc

Other
J. Brug, professor of Epidemiology
J. Twisk, professor in statistics, VUMC
T. Abma, professor and research director at the department of Medical Humanities, VUmc
I. de Groot-Loeve, sonographer, private practice primary care, member of BEN
R. Verwey, chair consumerorganisation 'Het Ouderschap', G.
Bonsel, epidemiologist, professor EMC
K. van der Pal, epidemiologist, TNO.
A. van Baar, psychologist, Professor Clinical Child and Family studies, Utrecht University
T. Roseboom, medical biologist/ epidemiologist, AMC
B. Molewijk, clinical ethicist and senior researcher/ program director Moral Deliberation, VUmc
N. van Duijnhoven, senior researcher health sciences, Nijmegen
C. Hukkelhoven, senior epidemiologist, PRN, Utrecht
C. Oudejans, Professor Clinical Chemistry, VUmc
B. Mol, professor of Obstetrics and Gynecology at The Robinson Institute, Adelaide (Australia)

Members international IRIS-study group

E. Hutton, professor of Midwifery Science at AVAG/EMGO+, VUmc and associate professor at department of Obstetrics and Gynecology at McMaster University, Hamilton (Canada)

J. Gardosi, professor Maternal and Perinatal Health, Warwick and Director West Midlands Perinatal Institute, Birmingham (UK)

U. Reddy, medical officer, Pregnancy and Perinatology Branch, National Institute of Health, Bethesda (USA)

W.A. Grobman, department of Obstetrics and Gynaecology, Northwestern University Chicago (U.S.A.) J. Zhang, MOE -Shanghai Key Laboratory of Children's Environmental Health, Shanghai Jiao Tong University School of Medicine (Shang Hai)

A. Papageorghiou, senior fellow in fetal medicine, INTERGROWTH-21st Project Team, University of Oxford (UK)

SUMMARY

BACKGROUND: Of all babies that die after 25 or more weeks gestation, 40% are small-for-gestational-age (SGA). In the Netherlands third trimester ultrasound (US) screening is increasingly being used to monitor foetal growth even though evidence on its effectiveness or cost-effectiveness is lacking. The proposed study fulfils the urgent need to evaluate the value of third trimester US for monitoring foetal growth among low risk women in primary care. If shown to be effective, routine third trimester US will contribute to reducing the national perinatal mortality and severe morbidity rate.

DESIGN: A nationwide stepped wedge cluster randomised trial in which primary care midwifery practices will be randomised.

PARTICIPANTS: 15,000 women who receive/continue care in the participating midwifery practice after the 20 week structural ultrasound and who have a singleton pregnancy.

INTERVENTION AND CONTROL STRATEGIES: In all midwifery practices, growth will be monitored using standardised symphysis fundal height (SFH) measurement. In the intervention strategy two routine US examinations will be performed (between 28-30 weeks and 34-36 weeks), while in the control strategy US examination will only be performed when clinically indicated. In both groups the IRIS consensus-based protocol will be followed if intrauterine growth retardation is suspected.

PRIMARY OUTCOMES: The clinical primary outcome is a dichotomous composite measure 'severe adverse perinatal outcome' up to 7 days after birth, including: perinatal death; Apgar score below 4 at 5 minutes after birth; impaired consciousness; need for assisted ventilation for more than 24 hours; asphyxia; septicaemia; meningitis; bronchopulmonary dysplasia; intraventricular haemorrhage; cystic periventricular leukomalacia; neonatal seizures or necrotizing enterocolitis. Also direct and indirect costs are primary outcomes.

SAMPLE SIZE: A total of 15,000 women in 60 midwifery practices; 7,500 women per strategy.

MAIN DATA ANALYSIS AND ECONOMIC EVALUATION: Multivariable logistic regression analyses, taking into account the clustered design. The economic evaluation will consist of a cost-effectiveness and a cost-utility analysis and will be performed from both a health care provider and societal perspective. We will base all primary analyses on intention to treat.

DURATION: 48 months.

Problem definition

DETECTION OF SMALL-FOR-GESTATIONAL-AGE

Of all babies that die after 25 or more weeks gestation, 40% are small-for-gestational-age (SGA) [De Reu et al., 2010]. SGA refers to a foetus or neonate who has failed to achieve a specific biometric or estimated weight threshold by a specific gestational age [Voskamp et al., 2013]. SGA is merely a statistical construct [Gardosi 2009] and does not differentiate between physiological and pathological smallness. Not all small foetuses are being growth retarded, some are constitutionally small but healthy. A growth restricted foetus may, on the other hand, not be SGA [Voskamp et al., 2013]. In the majority of cases, intrauterine growth restriction (IUGR) is the result of placental insufficiency, which is caused by dysfunction of the fetal-placental perfusion [von Beckerath 2013].

Research into detection of IUGR is classified as a top priority [Bonsel et al, 2010; Flenady et al., 2011]. In primary midwifery care, the main method to monitor foetal growth is by abdominal palpation. Bais and colleagues (2004) estimated that the sensitivity to detect growth < P10 is 21% and the specificity 96%, meaning that this screening is not very effective.

THIRD TRIMESTER ULTRASOUND SCREENING

Another method that can be used to monitor foetal growth is routine third trimester ultrasound (US). US has a higher sensitivity for SGA with abdominal circumference and estimated foetal weight being the best predictors [De Reu et al., 2010]. US screening may also have other benefits, such as the detection of congenital abnormalities or non-cephalic presentation.

In the Netherlands third trimester US screening is increasingly being used routinely even though meta-analyses of randomised controlled trials among low risk populations have not shown better perinatal outcomes among these women [Alfirevic et al., 2010a,b; Bricker et al., 2009]. Reasons why these studies failed to show advantages of routine US among low risk women include:

- most studies were underpowered to detect clinically relevant differences in severe perinatal outcome;
- studies are rather old and US technology used is now outdated;
- studies suffer from considerable methodological shortcomings, including contamination – US scans were also frequently performed in the control group;
- heterogeneity between studies in number and timing of US scans
- in many studies, only the screening procedure was described but not the subsequent actions that were taken. It is important that the only variable that differs between the intervention and control group is the use of routine third trimester US or not.

ADVERSE EFFECTS OF US SCREENING

Introducing a screening programme always has potential negative effects [Sackett et al., 2011]. For example, if a cut-off point of P10 is used for the definition of SGA, about 10% of women will be classified as ‘high risk’. Additional monitoring may raise anxiety and unnecessary interventions may occur, most notably elective delivery. Nonetheless, many of the babies of these ‘high risk women’ will be constitutionally small and will not be compromised.

Furthermore, US exposure or increased medical interventions may lead to short or long term effects in the neonate. Therefore, several authors stress the importance of examining long term neonatal outcomes [Alfirevic et al., 2010a,b; Bricker et al., 2009]. In an Australian randomised controlled trial, the Raine study, the incidence of SGA was higher in the group with serial US examinations (birth weight < P10, RR 1.84, 95% CI 1.17 to 2.89 and < P3, RR 1.46, 95% CI 1.14 to 1.87) [Newnham et al.,

1993], although no differences were found in child development at eight years of age [Newnham et al., 2004].

The increase in US examinations has led to a considerable rise in health care costs. The Dutch Minister of Health announced a cut in the budget for midwifery care of 10 million euro's because of overspending on US examinations. Also health insurance companies have indicated that they are worried about the increase of US in pregnancy without good evidence of its advantages and possible disadvantages.

IRIS-STUDY

At the moment, routine third trimester US for detection of SGA is still a hotly debated issue which means the time is ripe for a large study into this screening method. If this momentum passes, routine third trimester US will be common practice and a trial will no longer be feasible. We urgently need a high quality randomised controlled trial among low risk women examining whether routine third trimester US leads to improved perinatal outcomes and is cost-effective. Such a study should include sufficient participants to detect changes in severe perinatal outcome, use state-of-the art US equipment and explicitly explore the optimal timing of routine third trimester US. The proposed study, the IRIS-study, meets all of these criteria.

In the IRIS study we will evaluate whether detecting IUGR by routine third trimester US screening leads to improved perinatal outcomes through timely intervention, based on a multidisciplinary consensus-based protocol. This protocol will be developed and evaluated in sub-study A.

As screening inevitably leads to possible harms such as raised anxiety, unnecessary investigations and obstetric interventions and costs, we will evaluate in sub-study B ethical dilemmas with regard to unexpected findings and incorrect suspicion of IUGR and to explore how explore how professionals and women can deal with these issues.

The IRIS study will demonstrate whether routine US and subsequent protocolized management in the third trimester of pregnancy is a cost-effective way of reducing the rate of severe perinatal outcomes. If routine US is shown to be effective, its use can contribute to reducing the national perinatal mortality and severe morbidity rate.

IMPLEMENTATION ACTIVITIES IN THE STUDY

At the start of the project, we will develop a consensus-based protocol for detection and management of suspected IUGR in the IRIS study. This in itself is a tremendously useful spin-off of the project. At the moment, there is no (multidisciplinary) consensus in the country on when to repeat US, when to refer from primary to secondary care and vice versa, frequency of performing ultrasound examinations, or how to manage a deteriorating fetal condition. We will develop the IRIS consensus-based protocol by a Delphi study (IRIS sub-study A).

FEASIBILITY OF DATA COLLECTION

To facilitate data collection on the vast number of women required in the study (15,000), we will use existing databases as much as possible. We will work closely with Perinatal Registration Netherlands (PRN).

Aims and objectives

AIM

The main aim of the study is to evaluate the effectiveness and cost-effectiveness of routine US screening in the third trimester of pregnancy among low risk women in reducing severe adverse perinatal outcome. We aim to conduct this study before routine US in the third trimester is common practice among low risk women in the Netherlands.

OBJECTIVES

1. To evaluate whether routine US screening between 28-30 weeks and between 34-36 weeks gestation added to standardised SFH measurement and subsequent protocolised management among low risk women in primary care leads to a reduction in severe adverse perinatal outcome compared to use of US on indication added to standardised SFH measurement and subsequent protocolised management
2. To evaluate whether routine US screening between 28-30 weeks and between 34-36 weeks gestation added to standardised SFH measurement and subsequent protocolised management among low risk women in primary care is cost-effective compared to use of US on indication added to standardised SFH measurement and subsequent protocolised management
3. To develop a consensus-based IRIS protocol for detection and management of suspected SGA, including the role of third trimester US, and to study professionals' adherence to this protocol. This objective will be studied in sub-study A.
4. To examine ethical dilemmas with regard to unexpected findings and incorrect suspicion of IUGR and to explore how professionals and women can deal with these issues. This objective will be studied in sub-study B.

Study methods

STUDY DESIGN

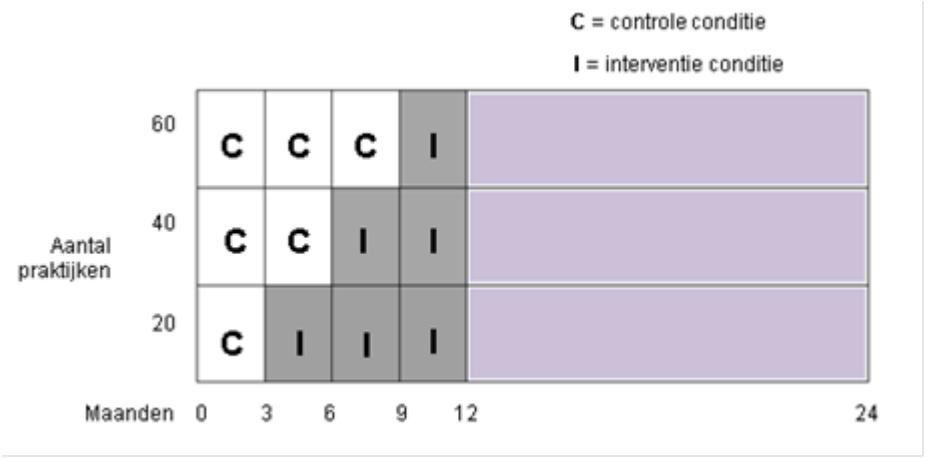
The IRIS study is designed as a stepped wedge cluster randomised trial (c-RCT, Figure 1). A stepped wedge design is a type of crossover design in which different clusters cross over (switch treatments) at different time points. Clusters cross over in one direction only; typically from control to intervention. The time at which a cluster starts with the intervention is randomized; more than one cluster may start the intervention at a particular time point [Hussey 2007]. The advantages of a stepped wedge design are: a) all practices will offer the intervention for a shorter or longer period of time, which may encourage them to take part, b) practices will serve as their own control for a variable time period which will reduce the variation in characteristics and practice management between the intervention and control strategy group. Note: in contrast to many other stepped wedge c-RCTs the women participating in our study will receive the intervention strategy OR control strategy.

Midwifery practice will be the unit of randomisation. Randomisation per practice rather than per midwife or client has the advantage of minimising contamination and therefore maximising contrast between the intervention and control strategy group.

All midwifery practices (n=60) will start in the control strategy and provide obstetric care as usual. At intervals of 3 months, 33,3% of all practices (n=20) will change status from control to intervention group and will start offering the intervention, i.e. routine biometry ultrasound in the third trimester of pregnancy. To balance the number of women in the intervention and control group, practices will be stratified prior to randomisation in large and small practices, with the average practice size as the cut-

off point (i.e. 250 clients a year). Within these two strata, a computer-generated random sequence will determine the order in which practices move from control to intervention status. This randomisation will take place after practices have been recruited for participation in the study and before data collection begins.

Figure 1: The IRIS stepped wedge design | UC = control strategy; INT= intervention strategy; US= ultrasonography



Clients will be recruited during months 1-12 (at 20-22 weeks gestation). During months 13-16 clients who are recruited last will receive their routine US (at 28-30 weeks and 34-36 weeks gestation). In month 17 the last IRIS study babies will be born. Follow-up measurement will take place in months 18-23.

SELECTION CRITERIA, RECRUITMENT, AND INFORMED CONSENT

Midwifery practices

Practices are eligible if they meet the IRIS study quality criteria of fetal ultrasonography. Furthermore, only practices where all midwives have received (or: agree to soon receive) the post registration training on the new KNOV guideline 'detection of IUGR' will be included in our trial.

Midwifery practices may show interest to participate in our RCT via a question in our nationwide survey on detection and management of suspected IUGR (sub-study A). Additionally, we will attend meetings of regional maternity care networks (the so-called VSVs) to inform and invite members to participate in our randomized trial. Other methods to inform midwifery practices about the IRIS study will be the post registration training sessions within the context of the new KNOV guideline; via the newly founded Midwifery Research Network Netherlands (MRNN) which is part of the midwifery consortium; articles in the journals of the national professional organisations; via social media.

All midwifery practices that show interest will be visited by a researcher who will inform them about the study's aims and procedures, will check if the practice fulfils the inclusion criteria, and if so, ask the midwives to sign a contract to show their commitment to the study protocol, in particular to the use of US according to randomisation status and use of the IRIS consensus-based protocol.

Pregnant women

In participating practices, pregnant women will be handed an information leaflet on the IRIS-study by their midwife during the first consultation after the 20 week screening test. Inclusion criteria: receiving/continuing care in the participating midwifery practice after the 20 week screening test and

having a singleton pregnancy. The midwife will ask the woman if she is willing to participate in the IRIS study and if so, the midwife will ask her to sign the consent form. The form asks the woman's permission for using her personal data for merging data, for collecting detailed information from the woman's and her baby's client records, and for approaching her in the context of future sub-studies. An extra consent form will be added among the subsample of women who will fill in questionnaires. Consent forms will contain name, address, postal code, telephone number, and email and will be kept separate from other data files in a secure place.
Inclusion criteria: receiving

Women who do not have had a dating US will only be included if they have a reliable estimated data of delivery (EDD) based on the first day of their last menstrual period.

The NVOG guideline on US dating of pregnancy will be used for estimation of the EDD.

CONTROL AND INTERVENTION US STRATEGIES

Women in both US strategies will receive routine care as follows:

- Standardised SFH measurement
- Information about life style factors that may influence foetal growth, such as smoking, use of alcohol and drugs, and labour circumstances.
- Advise to report a reduction in fetal movements.
- US examinations on indication

Subsequent protocolised management will be according to the IRIS consensus-based protocol (to be developed in sub-study A).

Intervention US strategy

Additionally, two routine third trimester US examinations will be performed, one between 28 and 30 weeks and one 6 weeks after the first US. By performing two US examinations the growth pattern can be monitored. If a baby does not follow the expected curve, this may indicate a lack of growth even before the abdominal circumference or the estimated fetal weight is low. Subsequent protocolised management will be accomplished according to the IRIS consensus based guidelines.

PRIMARY OUTCOME - SEVERE ADVERSE PERINATAL OUTCOME

The primary clinical outcome is the occurrence of any severe adverse perinatal outcome up to 7 days after birth (yes/no). It concerns a dichotomous composite measure with 'yes' being defined as the presence of one or more of the following:

1. Antepartum, intrapartum or neonatal death occurring from 28 weeks gestation
2. Apgar score below 4 at 5 minutes after birth;
3. Coma, stupor or decreased response to pain;
4. Asphyxia, defined as cord blood arterial base excess of less than minus 12;
5. Neonatal seizures defined as clonic movements which cannot be stopped by holding the limb, occurring on two or more occasions before 72 hours of age;
6. Assisted ventilation for more than 24 hours via endotracheal tube initiated within 72 hours after birth;
7. Septicaemia, ascertained by a positive blood culture;
8. Meningitis, ascertained by positive cerebrospinal fluid culture;

9. Bronchopulmonary dysplasia (BPD), defined as need for oxygen at a postnatal gestational age from 36 completed weeks as well as an X -ray compatible with BPD;
10. Intraventricular hemorrhage, defined as grade 3 or 4 and diagnosed by cranial US or at autopsy;
11. Cystic periventricular leukomalacia (PVL), diagnosed by cranial US or at autopsy showing periventricular cystic changes in the white matter excluding subependymal and choroid plexus cysts;
12. Necrotizing enterocolitis, defined as either perforation of intestine, pneumatosis intestinalis or air in the portal vein, diagnosed by X-ray, surgery, or at autopsy;

Cord blood arterial base excess was chosen as an indicator of asphyxia because, unlike pH, it is linearly related to the degree of metabolic acidosis which is a better indicator of asphyxia than measures of respiratory acidosis [Ross & Gala, 2002]. There was some discussion in the expert and project group on the use of Apgar scores in the composite outcome. Apgar score in itself is not comparable to the other outcomes, which are more predictive of long term severe adverse outcomes. However, in the Netherlands, a quarter of all women still give birth at home. Suspected SGA is a reason for referral to specialist care and routine US may lead to more SGA babies being born in hospital. Nevertheless, unexpected poor outcomes may occur at home and blood gases immediately after birth will then not be available. In hospital too, cord blood samples may be forgotten or go missing. We have chosen a low cutoff point for Apgar score at five minutes (below 4) in which case most babies will have other adverse outcomes as well. In a recent Dutch RCT on the management of pregnancies complicated by SGA, Apgar scores were also part of the composite outcome [Boers et al., 2010]. Low birth weight is not part of the primary outcome because this is not a measure of morbidity but rather a factor that is associated with morbidity such as respiratory problems. Also, premature labour may sometimes be induced to prevent stillbirth even though prematurity is associated with low birth weight.

PRIMARY OUTCOME - COSTS

Direct costs will include costs related to pregnancy related healthcare use, such as consultations with the midwife, referrals to specialist care, US scans, laboratory tests; CTG monitoring; hospital admission; interventions during labour; admission to neonatal unit. We will measure absenteeism and presenteeism (indirect costs) by the Productivity Cost Questionnaire (PCQ) [IMTA, Rotterdam]

To value health care utilization, standard costs published in the Dutch costing guidelines will be used [Hakkaart-van Roijen et al., 2011]. Medication use will be valued using prices of the Royal Dutch Society for Pharmacy. The friction cost approach will be used to estimate indirect costs using Dutch age and sex specific lost productivity costs [Koopmanschap & Rutten, 1996; Hakkaart-van Roijen et al., 2011].

SECONDARY OUTCOMES

<u>Composite outcomes</u>

1. Spontaneous vaginal birth without intervention, i.e. a birth without any of the following:
    - induction of labour other than amniotomy
    - vacuum/ forceps
    - caesarean section
    - augmentation of labour
    - pharmacological pain relief: epidural anaesthesia or use of opioids

2. Maternal morbidity, defined as the presence of one or more of the following:

   - maternal death within 42 days after giving birth
   - hypertension, defined as highest diastolic blood pressure equal or greater than 95
   - pre-eclampsia, defined as diastolic blood pressure equal or greater than 90 and proteinuria >= 300 mg/l
   - postpartum haemorrhage > 1000 mL
   - third or fourth degree perineal trauma

Singular outcomes

- elements of composite primary outcome
- elements of the two composite secondary outcomes
- neonatal mortality and severe morbidity between 7th and 28th day after birth
- detection of congenital abnormalities
- life threatening congenital conditions born in primary care
- home and hospital birth in primary care
- birth weight: mean; % <P5; % >P95;
- gestational age at birth: mean; % < 37 week
- non-cephalic presentations (when labour started) in primary care
- general quality of life, measured by the EQ-5D-5L [Herdman et al., 2011]
- symptoms of depression, measured by the EPDS [J.L. Cox et al., 1987; Dutch translation: V.J.M. Pop et al., 1992]
- Continuity of health care, measured by the NCQ [A.A. Uijen et al., 2011]
- Satisfaction with health care, measured by the PCQ [S.E.M. Truijens et al., 2014]

OTHER OUTCOMES

Descriptive variables and predictors of outcome

Characteristics of midwifery practice, as presented in the latest LVR report

- number of midwives working in the practice
- number of clients per year
- proportion of nulliparous and multiparous women
- proportion of women per ethnic group
- proportion of home births
- rate of referral to secondary care

Characteristics of clients, collected by the midwife after the informed consent procedure:

- mother's date of birth, ethnic background, years of education, weight and height
- smoking (never, discontinued before current pregnancy, discontinued in first trimester, discontinued later in pregnancy; less than 10 cigarettes a day, 10-20, more than 20)
- alcohol use (no, discontinued before current pregnancy, discontinued in first trimester, discontinued later in pregnancy; sporadically, 0-2 units a day, more than 2)
- drug use (no, discontinued before current pregnancy, discontinued in first trimester, discontinued later in pregnancy; occasionally, daily, weekly, monthly)
- work status

Process measures

- % women declining participation
- % protocol violations
- % disagreements in primary outcome on verification by research assistant
- 3rd trimester US: intra-observer variability, inter-observer variability
- accuracy US: sensitivity, specificity of routine US to detect SGA
- Opinion of midwives towards effective and ineffective elements of intervention strategy

DATA COLLECTION

Existing databases

For all 15,000 women we will extract data from the following existing databases: 1) PRN database; 2) ultrasound centres' databases; 3) hospitals' patient records. These databases will be used to collect data on the primary clinical outcome 'severe adverse perinatal outcome' (e.g. Apgar score), obstetric variables, US scans (number, indication and findings, including biometry measures) and care processes (like number of days in hospital).

Forms and surveys

To collect data on variables that are not included in existing databases we will use the following forms and surveys.

Firstly, prior to onset of the study midwifery practices will fill in a short questionnaire on midwife practice characteristics as published in the practice's most recent LVR report.

Secondly, midwives will be asked to collect data on a small set of important prognostic variables for all women having signed the informed consent form (for example, on ethnicity, smoking, see p. 14).

Thirdly, a random sample of 450 women receiving the intervention US strategy and 450 women receiving the control US strategy will be asked to complete questionnaires at 22 and 32 weeks gestation during pregnancy, and at 6 weeks and 6 months after EDD (n= 15 per midwifery practice). Additionally, a non-random sample consisting of 600 women in whom IUGR is suspected by ultrasonography (300 intervention and 300 control women) will be asked to complete questionnaires (n= 10 per midwifery practice). Assuming a non-response and drop-out rate of 33% we aim to collect follow-up data on 1000 women. Questionnaires will collect detailed information on healthcare utilization related to the pregnancy, type of work, absenteeism and presenteism (PCQ), quality of life (EQ-5D-5L), satisfaction with health care (PCQ), continuity of health care (NCQ) and depression (EPDS). To minimise non-response the baseline questionnaire will be handed to the women by their midwife. Women will receive an email with a link to online follow-up questionnaires. Non-responders will be contacted by telephone and/or receive a postal follow-up questionnaire. To enhance participation from women from various ethnic backgrounds and socio-economic positions interviewers who speak Dutch, English, Turkish and/or Moroccan will be available to fill in the questionnaire by telephone if women need assistance.

Lastly, research midwives/nurses will collect detailed information from hospital medical records on 1) clinical outcomes and care process of babies who are referred to a pediatrician for neonatal admission; 2) clinical outcomes and care process of babies with severe adverse perinatal outcome according to data of existing databases and 3) pregnancy related health care utilization of women participating in the questionnaire study (n=1000).

DATA MONITORING AND EVALUATION

Data will be checked for missing values and irregularities. US management of midwifery practices will be checked at least two times during the recruitment period.

STATISTICAL ANALYSES

Sample size of full trial

Based on PRN data the expected rate of severe adverse perinatal outcome in the study population is 1.54 %. Unfortunately, national or international consensus is lacking on what reduction in severe perinatal outcome is feasible and clinically relevant. The IRIS study group agreed to strive for a reduction from 1.54% to 1.0% in the primary outcome. With 80% power and a significance level of 0.05, we need to include 13,536 women. However, given the clustered design this sample size needs adjustment for dependency of data.

Pagel et al. (2011) note that intra cluster correlation (ICC) estimates for perinatal health outcomes in primary care are scarcely available in literature; most come from hospital-based studies and may therefore not reflect outcomes in the community. To change this situation, the authors estimated the ICC and the coefficient of variation for a range of outcomes using data from five community-based c-RCTs in three low-income countries. In their paper they present five ICCs for neonatal mortality ranging from 0.0003 to 0.002. Additionally, they performed a simulation exercise to investigate the impact of cluster size and number of clusters on the reliability of estimates of the coefficient of variation for rare outcomes. Their results show that estimates of ICC are associated with the prevalence of the outcome of interest, the nature of the outcome of interest (mortality or behavioural) and the size and number of clusters. They advise that when planning future trials, published estimates of ICCs from larger clusters are probably safer to use and a range of possible ICCs should be used for sample size calculations.

Given the above mentioned ICCs for neonatal mortality, the formula to correct for clustering [1 + (n-1) * ICC] with n = 250 (i.e. average cluster size) gives a required sample size for the IRIS study of 14,547 (ICC 0.0003) to 20,304 (ICC 0.002). For two reasons we expect the ICC in the IRIS study to be much more similar to 0.0003 than to 0.002. Our first argument is that the ICCs reported in the paper of Pagel et al. (2011) are based on prevalence rates ranging from 1.5% to 5.9% which is higher than the expected severe perinatal outcome in the IRIS study. As ICCs are associated with the prevalence of the outcome of interest [Pagel et al., 2011] we may expect the ICC in the IRIS study to be even lower than 0.0003. Our second argument is that midwife practices will offer both the control strategy (i.e. US on indication) and, later on, the intervention strategy (i.e. routine US). This will reduce the variation in characteristics and practice management between the intervention and control strategy group and thereby lower the size of the ICC .

Taking into account an ICC of 0.0003, the required sample size is 14,547 women. As not all women who will sign the informed consent form may receive care in the participating midwifery practice at the start of the third trimester, we adjusted this figure to a total of 15,000.

Effectiveness analyses

Firstly, we will assess the similarity between baseline characteristics of women having had the intervention strategy and women having had the control strategy. Secondly, we will compare baseline characteristics of drop-outs and completers by using logistic regression analysis. Thirdly, we will use multivariable logistic multilevel analyses to investigate the effect of routine US screening on perinatal outcome, adjusting for possible clustering of observations and confounders (e.g. ethnicity, maternal age, weight and height, smoking, alcohol use, drug use, blood pressure). We will base our analyses on intention to treat and set the level of significance at $p<0.05$.

We will perform a subgroup analysis for women without US indication and for nulliparous women.

In addition to the intention to treat analysis we will also perform a per-protocol analysis, restricted to the women in the intervention group who received the two US examinations and to the women in the control group who did not receive any screening US.

Cost-effectiveness and cost-utility analyses

Firstly, we will conduct a pregnancy related cost-effectiveness analysis from a health care provider perspective. The time horizon of this analysis is from 26 weeks of gestation until one week after the date of birth. In this analysis, pregnancy related data of all 15,000 women will be included, using the PRN data. Detailed cost data collected by the research midwives on the 600 random and 400 selected women will be used to estimate the pregnancy related direct costs for the whole group of 15,000 women by using Bayesian techniques in combination with Monte Carlo simulation. These costs will be related to a composite of severe adverse perinatal outcome to estimate the Incremental Cost-Effectiveness Ratios (ICER).

Secondly, we will conduct a cost-effectiveness analysis from a societal perspective. The time horizon of this analysis is from 22 weeks gestation until 6 months after the EDD. In this analysis, pregnancy related data of all 15,000 women will be included, using the PRN data. Detailed cost data collected on the 600 random and 400 selected women will be used to estimate the pregnancy related direct costs for the whole group of 15,000 women by using Bayesian techniques in combination with Monte Carlo simulation. Furthermore, the combination of direct and indirect costs will be related to a composite of severe adverse perinatal outcome to estimate the Incremental Cost- Effectiveness Ratios (ICER).

Finally, we will conduct a cost-utility analysis from a societal perspective. The time horizon of this analysis is from 22 weeks gestation until 6 months after the expected date of birth. In this analysis, solely data of the 600 randomly selected women participating in the survey study will be included. To estimate outcomes more precisely for women in the random sample with suspected IUGR, we will additionally use the data of the 400 women in the selective sample. Costs will be related to Quality-Adjusted Life-Years based on the EQ-5D-5L using the Dutch tariff to estimate the Incremental Cost-Utility Ratio (ICUR).[Lamers et al., 2005] In this analysis, missing cost and effect data will be imputed using multiple imputation according to the MICE algorithm developed by Van Buuren and colleagues (1999).

Both cost-effectiveness and cost-utility analyses will be done according to the intention-to-treat principle. ICERs and the ICUR will be calculated by dividing the difference in mean costs between the two groups by the difference between the two groups in effects and in QALYs, respectively. Bias-corrected and accelerated bootstrapping with 5000 replications will be used to estimate 95% confidence intervals around cost differences and the uncertainty surrounding the ICERs and ICUR. Uncertainty surrounding the ICERs and ICUR will be graphically presented on cost-effectiveness planes. Cost-effectiveness acceptability curves will also be estimated using the net benefit framework.[Stinett & Mullahy, 1998] Cost-effectiveness acceptability curves show the probability that routine third trimester ultrasound screening to detect IUGR is cost-effective in comparison with usual care for a range of different ceiling ratios thereby showing decision uncertainty.

We will perform a subgroup analysis for women without US indication and for nulliparous women.

We will explore the robustness of the results of our intention to treat analysis by a per-protocol sensitivity analysis. Furthermore, we will repeat the pregnancy related cost-effectiveness analysis from a health care provider perspective and the cost-effectiveness analysis from a societal perspective using solely the data of the 600 randomly selected women participating in the survey study.

REFERENCES

Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in high-risk pregnancies. Cochrane Database Syst Rev 2010a;(1):CD007529.

Alfirevic Z, Stampalija T, Gyte GM. Fetal and umbilical Doppler ultrasound in normal pregnancy. Cochrane Database Syst Rev 2010b;(8):CD001450.

Bais JM, Eskes M, Pel M, Bonsel GJ, Bleker OP. Effectiveness of detection of intrauterine growth retardation by abdominal palpation as screening test in a low risk population: an observational study. Eur J Obstet Gynecol Reprod Biol 2004 Oct 15;116(2):164-9.

Boers KE, Vijgen SM, Bijlenga D, van der Post JA, Bekedam DJ, Kwee A, van der Salm PC, van Pampus MG, Spaanderman ME, de Boer K, Duvekot JJ, Bremer HA, Hasaart TH, Delemarre FM, Bloemenkamp KW, van Meir CA, Willekes C, Wijnen EJ, Rijken M, le Cessie S, Roumen FJ, Thornton JG, van Lith JM, Mol BW, Scherjon SA; DIGITAT study group. Induction versus expectant monitoring for intrauterine growth restriction at term: randomised equivalence trial (DIGITAT). BMJ 2010;341:c7087.

Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Signalementstudie 'Zwangerschap en geboorte'. Lijnen in de Perinatale Sterfte. Rotterdam: Erasmus MC; 2010

Bricker L, Mahsud-Dornan S, Dornan JC. Detection of foetal growth restriction using third trimester ultrasound. Best Pract Res Clin Obstet Gynaecol 2009 Dec;23(6):833-44

Cox JL, Holden JM, Sagovsky R. Detection of Postnatal Depression Development of the 10-item Edinburgh Postnatal Depression Scale. B J Psychiat 1987,150,782-786.

De Reu PA, Oosterbaan HP, Smits LJ, Nijhuis JG. Avoidable mortality in small-for-gestational-age children in the Netherlands. J Perinat Med 2010 May;38(3):311-8

Flenady V, Middleton P, Smith GC, Duke W, Erwich JJ, Khong TY, et al. Stillbirths: the way forward in high-income countries. Lancet 2011 May 14;377(9778):1703-17

Gardosi J. Intrauterine growth restriction: new standards for assessing adverse outcome. Best Pract Res Clin Obstet Gynaecol. 2009 Dec;23(6):741-9.

Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek: methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Geactualiseerde versie 2010 [Dutch manual for costing in economic evaluations]. Diemen, CvZ, 2011.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., Badia, X.: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual. Life Res. 20(10), 1727–1736 (2011)

Hussey, Hughes. Design and analysis of stepped wedge cluster randomized trials. Con Clin Tials 2007;28:182-91.

Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. Pharmacoeconomics 1996;10(5):460-6.

Lamers LM, Stelmeier PF, McDonnell J, Krabbe PF, van Busschback JJ. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff. NTvG 2005;149(28):1574-8

Newnham JP, Evans SF, Michael CA, Stanley FJ, Landau LI. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. Lancet 1993 Oct 9;342(8876):887-91.

Newnham JP, Doherty DA, Kendall GE, Zubrick SR, Landau LL, Stanley FJ. Effects of repeated prenatal ultrasound examinations on childhood outcome up to 8 years of age: follow-up of a randomised controlled trial. Lancet 2004 Dec 4;364(9450):2038-44.

Pagel C, Prost A, Lewycka S, Das S, Colbourn T, Mahapatra R, Azad K, Costello A, Osrin D. Intracluster correlation coefficients and coefficients of variation for perinatal outcomes from five cluster-randomised controlled trials in low and middle-income countries: results and methodological implications. Trials 2011; 12:151.

Pop VJ, Komproe IH, van Son MJ. Characteristics of the Edinburgh Post Natal Depression Scale in The Netherlands. J Affect Disord. 1992 Oct;26(2):105-10.

Ross MG, Gala R. Use of umbilical artery base excess: algorithm for the timing of hypoxic injury. Am J Obstet Gynecol 2002 Jul;187(1):1-9.

Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology. How to do clinical practice research. 3rd ed. Lipincott Williams and Wilkins; 2011

Stinett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998:18(2 Suppl):68-80.

Truijens SEM, Pommer AM, van Runnard Heimel PJ, Verhoeven CJM, Oei SG, Pop VJM. Development of the Pregnancy and Childbirth Questionnaire (PCQ): evaluation quality of care as perceived by women who recently gave birth. Eu J Obs & Gyn Rep Biol 174 (2014) 35-40.

Uijen AA, Schellevis FG, van den Bosch WJ, Mokkink HG, van Weel C, Schers HJ. Nijmegen Continuity Questionnaire: development and testing of a questionnaire that measures continuity of care. J Clin Epidemiol 2011 Dec;64(12):1391-9.

van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. StatMed 1999;18:681-694.

von Beckerath AK, Kollmann M, Rotky-Fast C, Karpf E, Lang U, Klaritsch P. Perinatal complications and long-term neurodevelopmental outcome of infants with intrauterine growth restriction. Am J Obstet Gynecol. 2013 Feb;208(2):130

Voskamp BJ, Kazemier BM, Ravelli AC, Schaaf J, Mol BW, Pajkrt E. Recurrence of small-for-gestational-age pregnancy: analysis of first and subsequent singleton pregnancies in The Netherlands. Am J Obstet Gynecol. 2013 May;208(5):374