Effectiveness of routine third trimester ultrasonography to reduce adverse perinatal outcomes in low risk pregnancy (the IRIS study): nationwide, pragmatic, multicentre, stepped wedge cluster randomised trial

Jens Henrichs,[1] Viki Verfaille,[1,2] Petra Jellema,[1] Laura Viester,[1] Eva Pajkrt,[3] Janneke Wilschut,[4] Henriëtte E van der Horst,[5] Arie Franx,[6] Ank de Jonge,[1] on behalf of the IRIS study group

For numbered affiliations see end of the article.

Correspondence to: A de Jonge ank.dejonge@amsterdamumc.nl (or @midwifescience on Twitter; ORCID 0000-0002-5384-3744)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2019;367:l5517
http://dx.doi.org/10.1136/bmj.l5517

Accepted: 27 August 2019

ABSTRACT

OBJECTIVES

To investigate the effectiveness of routine ultrasonography in the third trimester in reducing adverse perinatal outcomes in low risk pregnancies compared with usual care and the effect of this policy on maternal outcomes and obstetric interventions.

DESIGN

Pragmatic, multicentre, stepped wedge cluster randomised trial.

SETTING

60 midwifery practices in the Netherlands.

PARTICIPANTS

13 046 women aged 16 years or older with a low risk singleton pregnancy.

INTERVENTIONS

60 midwifery practices offered usual care (serial fundal height measurements with clinically indicated ultrasonography). After 3, 7, and 10 months, a third of the practices were randomised to the intervention strategy. As well as receiving usual care, women in the intervention strategy were offered two routine biometry scans at 28-30 and 34-36 weeks' gestation. The same multidisciplinary protocol for detecting and managing fetal growth restriction was used in both strategies.

MAIN OUTCOME MEASURES

The primary outcome measure was a composite of severe adverse perinatal outcomes: perinatal death, Apgar score <4, impaired consciousness, asphyxia, seizures, assisted ventilation, septicaemia, meningitis, bronchopulmonary dysplasia, intraventricular haemorrhage, periventricular leucomalacia, or necrotising enterocolitis. Secondary outcomes were two composite measures of severe maternal morbidity, and spontaneous labour and birth.

RESULTS

Between 1 February 2015 and 29 February 2016, 60 midwifery practices enrolled 13 520 women in mid-pregnancy (mean 22.8 (SD 2.4) weeks' gestation). 13 046 women (intervention n=7067, usual care n=5979) with data based on the national Dutch perinatal registry or hospital records were included in the analyses. Small for gestational age at birth was significantly more often detected in the intervention group than in the usual care group (179 of 556 (32%) *v* 78 of 407 (19%), P<0.001). The incidence of severe adverse perinatal outcomes was 1.7% (n=118) for the intervention strategy and 1.8% (n=106) for usual care. After adjustment for confounders, the difference between the groups was not significant (odds ratio 0.88, 95% confidence interval 0.70 to 1.20). The intervention strategy showed a higher incidence of induction of labour (1.16, 1.04 to 1.30) and a lower incidence of augmentation of labour (0.78, 0.71 to 0.85). Maternal outcomes and other obstetric interventions did not differ between the strategies.

CONCLUSION

In low risk pregnancies, routine ultrasonography in the third trimester along with clinically indicated ultrasonography was associated with higher antenatal detection of small for gestational age fetuses but not with a reduced incidence of severe adverse perinatal outcomes compared with usual care alone. The findings do not support routine ultrasonography in the third trimester for low risk pregnancies.

TRIAL REGISTRATION

Netherlands Trial Register NTR4367.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Fetal growth restriction is a risk factor for perinatal mortality and morbidity and cardiovascular disease and neurodevelopmental disorders in adulthood

Routine ultrasonography in the third trimester detects neonates who are small for gestational age (SGA) significantly more often than usual care using serial fundal height measurements combined with clinically indicated ultrasonography

Evidence that routine ultrasonography in the third trimester reduces the incidence of severe adverse perinatal outcomes is lacking

WHAT THIS STUDY ADDS

In low risk pregnancies, routine ultrasonography in the third trimester combined with clinically indicated ultrasonography was associated with greater antenatal detection of SGA neonates and induction of labour but was not associated with a reduction in severe adverse perinatal outcomes compared with usual care

Based on these findings, routine ultrasonography has no benefit (or harm) to the neonate but was associated with a moderately increased incidence of induction of labour

These findings do not support routine ultrasonography in the third trimester for low risk pregnancies

Introduction

Fetal growth restriction, defined as failure to reach fetal growth potential, occurs in 7-10% of pregnancies.[1 2] Causes include smoking, exposure to teratogens, maternal malnutrition, infections, genetic and structural disorders, and placental insufficiency.[3] Fetal

growth restriction is a major risk factor for perinatal morbidity and perinatal death, including sudden intrauterine unexplained death.[3-5] It is also associated with an increased risk of diseases in adulthood, such as neurodevelopmental and cardiovascular disorders.[2 6] Monitoring fetal growth and managing suspected growth restriction are therefore key objectives of antenatal care.[7] The terms fetal growth restriction and small for gestational age (SGA) are often used interchangeably although they do differ.[3] SGA is defined as the statistical deviation of fetal size or birth weight from a population based reference, with a predefined threshold that is usually the 10th centile.[8] Antenatal SGA is indicated by fetal abdominal circumference or estimated fetal weight below the 10th centile.[9] Because fetal growth restriction is not easily measured but most often occurs in SGA fetuses,[10] SGA is generally used as a proxy for growth restriction. Because many SGA infants can be constitutionally small but healthy, this proxy is unsatisfactory.[11] To improve the detection of fetal growth restriction prenatally, repeated measurements of fetal growth are recommended.[12-14]

Which third trimester screening strategy is most effective in detecting fetal growth restriction is controversial.[11 15] Routine ultrasonography in the third trimester detects SGA at birth more often than usual care, which comprises serial fundal height measurements combined with clinically indicated ultrasonography. Sensitivities of routine third trimester ultrasonography range from 22% to 57%.[9 11 16] Nevertheless, evidence based guidelines in many Western countries, including the Netherlands, do not recommend routine biometry scans in the third trimester because in previous clinical trials perinatal outcomes were not positively affected.[7 14 15 17 18] Also, when evaluating the introduction of a new screening programme, adverse effects, such as unnecessary medical care, should be considered.[19] Major limitations, however, have been identified in earlier trials.[15] Firstly, almost all trials were underpowered to detect clinically significant differences in severe perinatal outcomes. Secondly, in many trials, only the ultrasound screening strategy was described and the subsequent clinical management of suspected fetal growth restriction was unclear.[15] Biometry screening alone cannot prevent adverse perinatal outcomes unless screening is combined with effective clinical management.[11] Thirdly, the ultrasound technology used in most of the earlier randomised studies is outdated.[15]

The Dutch Ministry of Health considered introducing routine ultrasonography in the third trimester of pregnancy but was unable to decide on the effectiveness of this screening approach owing to lack of evidence. We therefore conducted a large pragmatic trial, the IUGR Risk Selection (IRIS) study, to evaluate the effects of offering routine ultrasonography in the third trimester to low risk pregnant women in the Netherlands. For this trial, we developed a multidisciplinary protocol based on consensus for detecting and managing suspected fetal growth restriction.[13] We chose a cluster randomised design to roll-out the intervention and to avoid contamination bias due to the women's preferences for or against ultrasound scans.[20] The stepped wedge design facilitated the participation of a large number of midwifery practices, even if they had a preference for one of the screening strategies. With this design, each practice first applied usual care and then switched to offering routine ultrasonography in the third trimester at a defined moment during the study, depending on the randomisation scheme. In this stepped wedge cluster randomised trial we evaluated the effectiveness of routine ultrasonography in the third trimester combined with usual care (ie, serial fundal height measurements with clinically indicated ultrasonography) in reducing severe adverse perinatal outcomes in low risk pregnancies compared with usual care alone. Both approaches included a multidisciplinary protocol for detecting and treating fetal growth restriction. We also examined the effect of the intervention on maternal outcomes and obstetric interventions.

Methods

Study design and participants

The IRIS study was a nationwide, stepped wedge cluster randomised trial conducted in 60 primary care midwifery practices in the Netherlands in low risk pregnant women. The study design has been previously described.[20]

In the Netherlands, hospitals provide secondary and tertiary antenatal care, whereas primary care midwives are independent medical practitioners qualified to provide full maternity care for women with uncomplicated low risk pregnancies.[21] Midwifery practices were invited to participate in the IRIS study at meetings, at postgraduate courses, and through social media and professional journals.[20] Practices were included if the midwives had completed the postgraduate registration training in the detection of fetal growth restriction based on the guideline of the Royal Dutch Association of Midwives (KNOV).[17] Biometry scans were performed in sonography centres or in midwifery practices. Some sonographers worked in both primary care centres and hospitals and others worked in primary care only. Participating practices signed an agreement showing their commitment to the study protocol.

Sonographers conducted third trimester biometry according to the guidelines of the Dutch Society of Obstetrics and Gynaecology (NVOG).[22 23] Sonographers who participated in the IRIS study were experienced in performing biometry and held a certificate for structural anomaly screening (73% of 154 participating sonographers) or passed a biometry quality test before the trial (27%), based on four biometry scans assessed by two experienced sonographers; had successfully completed a module on fetal biometry from a national Dutch medical e-learning education programme (see www.medicaleducation.nl); and used ultrasound equipment according to the standards of the Dutch Society of Obstetrics and Gynaecology.[14] Two independent and experienced sonographers who were

doi: 10.1136/bmj.l5517 | BMJ 2019;367:l5517 | the bmj

board members of the Dutch Professional Organisation of Sonographers carried out quality assessments of the sonographers during the trial.

From 1 February 2015 to 29 February 2016, pregnant women in the participating midwifery practices who fulfilled the inclusion criteria were informed about the study and given a trial information leaflet by their midwife during the first consultation. After the 20 week anomaly screening had been conducted if desired, the women were invited to take part in the study. Inclusion criteria for women with a low risk pregnancy were: antenatal care in a participating midwifery practice at enrolment, age 16 years or older, a singleton pregnancy, no major obstetric or medical risk factors, and a reliable expected date of delivery based on a dating scan or a reliable first day of the last menstrual period.[14] Participants provided written informed consent for data usage.

The control strategy (usual care) comprised fetal growth monitored by serial fundal height measurements and ultrasonography if clinically indicated. In addition to their usual care, women in the intervention strategy received two biometry ultrasound scans in the third trimester, at 28-30 and 34-36 weeks' gestation, to detect fetal growth restriction.

Randomisation and masking

Midwifery practices formed the unit of cluster randomisation. At the onset of data collection on 1 February 2015 all the midwifery practices (n=60) carried out the control strategy, with a third sequentially crossing over to the intervention strategy at 3, 7, and 10 months (fig 1). Practices were stratified before randomisation into large and small practices, with the median practice size (300 women annually) as a cut-off. A stratified computer generated randomisation sequence determined the order in which practices changed from the control to intervention strategy. An independent statistician performed randomisation on anonymous data from the midwifery practices. Because of the nature of the intervention, it was not possible to blind participants, care providers, and researchers to group allocation.

Procedure

The logistics of the study and enrolment procedures were piloted in January 2015. The pilot data were not included in the analyses. In the intervention and control strategies, we used prenatal SGA and slow fetal abdominal growth as indicators for suspected fetal growth restriction. We defined prenatal SGA as a fetal abdominal circumference below the 10th centile based on a population based Dutch reference growth curve.[24] Slow fetal abdominal growth was defined as a decrease in abdominal circumference of at least 20 centiles (eg, from the 70th to 50th centile, with a minimum interval of two weeks) on the Dutch reference curve.[13 24] A volume of amniotic fluid of less than 2 cm in the deepest vertical pocket was also an indication of suspected fetal growth restriction. In both strategies, women with suspected fetal growth restriction were referred to obstetrician led care for further diagnosis and management. Women remained in the strategy that their midwifery practice was allocated to on enrolment. In both strategies, women with suspected fetal growth restriction were referred to obstetrician led care for further diagnosis and management. Suspected fetal growth restriction was detected and managed based on a protocol specifically developed for this study in a Delphi study incorporating recommendations from national and international guidelines (see appendix 1).[13 14 17 25]

Information on the characteristics of the participating midwifery practices was collected by electronic survey before the start of the study in January 2015. At inclusion in the study, women completed a survey of questions on personal characteristics, anthropometric measurements, smoking status, and intake of alcohol and recreational drugs.

We retrieved clinical data on care processes, perinatal outcomes, and maternal outcomes from the database of the Netherlands Perinatal Registry (Perined), which collects healthcare data from midwife led and obstetrician led care in the Netherlands.[26] Data on ultrasound scans were obtained from the databases of the midwifery practices and the participating sonography centres.

For suspected severe adverse perinatal outcomes based on the Perined database, five trained research assistants retrieved detailed clinical data from hospital files using standard case report forms. Hospital files were selected in cases of perinatal death, a low Apgar score (<4) at five minutes, a birth weight less than the 2.3rd centile,[27] (or a birth weight between the 2.3rd and

No of practices	1 Feb to 30 Apr 2015	1 May to 31 Aug 2015*	1 Sep to 30 Nov 2015	1 Dec 2015 to 29 Feb 2016
20†	1039	1568	1142	1225
20	1059	1184	1061	1230
20	980	1061	825	1146

Intervention strategy (n=7372) Control strategy (n=6148)

Fig 1 | Study sample in stepped wedge cluster randomised trial, showing numbers recruited in each cluster and period. *Crossover postponed after one month because of fewer than expected inclusions †One midwifery practice dropped out in April 2015, after the first randomisation

5th centile and) neonatal admission for more than three days, and referral to a neonatologist if the admission data were missing or were not clearly registered in Perined. To warrant adequate retrieval of information from the neonatal records by the research assistants, an experienced neonatologist helped to operationalise the perinatal outcomes, collected using the standard case report forms.

Outcomes

The primary outcome was a dichotomous composite measure of 12 adverse perinatal outcomes occurring up to seven days after birth: perinatal death between 28 weeks' gestation and seven days after birth; Apgar score <4 at five minutes; impaired consciousness (coma, stupor, or decreased response to pain); asphyxia, with arterial base excess of cord blood less than −12 mmol/L; seizures on at least two occasions within 72 hours of birth; assisted ventilation by endotracheal tube for more than 24 hours started within 72 hours of birth; septicaemia confirmed by blood culture; meningitis confirmed by culture of cerebrospinal fluid; bronchopulmonary dysplasia requiring oxygen after 36 weeks' gestation and confirmed by radiography; intraventricular haemorrhage grade 3 or 4 confirmed by ultrasonography or autopsy; cystic periventricular leucomalacia confirmed by ultrasonography; or necrotising enterocolitis confirmed by radiography, surgery, or autopsy.

Secondary neonatal outcomes were congenital abnormalities, birth weight, gestational age, prematurity (<37 weeks' gestation), SGA at birth (birth weight <10th centile), large for gestational age (birth weight >90th centile),[27] and neonatal mortality from eight to 28 days after birth.

Two dichotomous maternal composite outcomes were defined as secondary outcomes. The first composite outcome was at least one of four maternal adverse peripartum outcomes: maternal death within 42 days of birth, hypertensive disorders or pre-eclampsia (diastolic blood pressure ≥95 mm Hg with or without proteinuria, or ≥90 mm Hg with proteinuria), postpartum haemorrhage of 1000 mL or more, or anal sphincter damage. The second composite outcome was spontaneous labour and birth, defined as a spontaneous vaginal birth with no induction or augmentation of labour, no drug pain relief, no vacuum or forceps assisted birth, and no caesarean section. Other secondary outcomes were the individual components of the perinatal and maternal composite outcomes and the secondary outcomes non-cephalic presentation at the start of labour in midwife led care and birth in midwife led or obstetrician led care.

Statistical analysis

Perined data suggested an incidence of 1.54% for the severe adverse perinatal composite outcome in low risk pregnant women in the Netherlands. We defined a clinically significant reduction of severe adverse perinatal outcome in the intervention strategy as 1.54% to 1.0%. With an α of 5% and 80% power, inclusion of 13 536 women was required. Assuming an intracluster correlation coefficient of 0.0003 based on previous literature,[28] and an a priori assumed average cluster size (ie, practice size of 250 women annually), we aimed to include 15 000 pregnant women (7500 for each strategy) to be able to take possible clustering effects into account.[20]

We performed double entry analyses in a 5% sample of the hospital record forms collected. Based on the hospital records, we also calculated the number of women who received additional ultrasound and Doppler scans in obstetrician led care after they were referred for suspected fetal growth restriction by midwives because of fetal abdominal circumference below the 10th centile or slow fetal abdominal growth.

To estimate the diagnostic accuracy of the two third trimester screening strategies to detect SGA at birth (birth weight <10th centile based on the Dutch reference curve),[27] we calculated sensitivity, specificity, and positive and negative predictive values of fetal abdominal circumference below the 10th centile, slow fetal abdominal growth, or a combination of the two.[29] We compared sensitivity and specificity rates between the intervention and control strategies using the χ^2 test and between the scans at 28-30 weeks' gestation and 34-36 weeks' gestation in the intervention strategy using the McNemar test.

As a first step, we conducted univariable logistic regression analyses to see if routine ultrasonography in the third trimester was associated with a reduction in severe adverse perinatal outcomes and adverse secondary neonatal and maternal outcomes. Then we conducted multilevel multivariable logistic regression analyses for the dichotomous primary and secondary outcomes. For continuous secondary outcomes, we ran multivariable linear mixed models. Because of the cluster randomised design, we included midwifery practice as a random effect in the multilevel regression models. Time of inclusion, divided into four groups according to the crossover from usual care to the intervention strategy, was considered as a fixed factor. As the study condition (intervention versus usual care) and time of inclusion were strongly correlated (Pearson's r=0.73, P<0.001), indicating collinearity,[30] we did not include this fixed factor in the multilevel multivariable logistic (or linear) regression analyses. Also, we adjusted our main analyses for potential confounders selected a priori and based on previous literature: maternal age; body mass index; smoking, alcohol, or recreational drug use; parity; educational level; employment status; marital status; infant's sex; and size of the midwifery practice (≤300 or >300 women annually).[31 32] Analyses were performed on complete case analysis given that less than 5% of the data on confounders were missing. We performed a multilevel analysis only if the expected number of events per cluster was at least one, as advocated previously.[33] We used an intention to treat approach. We then conducted a fully adjusted post hoc sensitivity analysis for the primary outcome, comparing women in the intervention strategy, who received two routine

doi: 10.1136/bmj.l5517 | BMJ 2019;367:l5517 | the bmj

ultrasound scans, with women in the control strategy, who received no ultrasound scan. Neonates born before the second routine ultrasound scan were excluded from the additional analysis. We set the level of significance at P<0.05. Statistical analyses were performed with the Statistical Package for Social Science (SPPS V.22; IBM, Chicago, IL) and R (V.3.4.3).

Patient and public involvement

A patient representative was a member of the project group that drafted the grant proposal and design of the IRIS study and of the sounding board of the IRIS study providing feedback to design aspects and discussing study results. Client organisations will be involved in communicating the findings of the study to the general public.

Results

From 1 February 2015, 60 midwifery practices participated in the IRIS study (about 12% of practices in the Netherlands). Nineteen practices performed biometry scans and the others referred women to one of the 18 sonography centres involved in the study. After the first randomisation in April 2015, one midwifery practice withdrew from the study because of time constraints. The remaining 59 practices participated in the study until 29 February 2016. As recruitment was slower than anticipated, the predefined recruitment period of one year was extended, and hence the second group of midwifery practices crossed over to the intervention strategy one month later than planned (fig 1). Twenty practices provided the intervention in the second period (May to August 2015), 40 in the third period (September to November 2016), and 59 in the fourth period (December 2015 to February 2016) (fig 1).

A total of 14 323 pregnant women were invited to participate in the IRIS study (fig 2). Six women did not fulfil the inclusion criteria and 797 refused to participate. The remaining 13 520 women were enrolled in mid-pregnancy (mean 22.8 (SD 2.4) weeks' gestation) and provided baseline characteristics. Neonates of the participating women were born between June 2015 and August 2016. Data from 13 024 (96.3%) women and neonates were linked to data in the Perined database. Data were retrieved from hospital records for 2339 cases, selected for additional in-depth data collection. In total, 13 046 women with Perined data or data from hospital records, or both were included in one or more analyses, with 5979 women in the usual care strategy and 7067 women in the intervention strategy. Data on severe adverse perinatal outcomes were available for 12 993 of 13 046 (99.6%) women, 7040 in the intervention strategy and 5953 in the control strategy (fig 2).

Double entry analyses on hospital case report forms of 111 women were carried out. The overall incidence of error in data entry was 3.2% (2.6% for neonatal data and 3.7% for maternal data).

Table 1 shows the baseline characteristics of the participating midwifery practices. Table 2 shows the personal and clinical baseline characteristics of the participants. Women in the intervention strategy had significantly more ultrasound scans than women in the control strategy (mean 2.04 (SD 0.75) *v* 0.88 (0.96), P<0.001). For the indication biometry, the values were lower (mean 1.84 (0.82) *v* 0.72 (0.90), respectively, P<0.001). Of 5840 women (82.6% of 7067 women) in the intervention strategy, who were not referred to obstetrician led care before 37 weeks' gestation, 3.0% (n=177) did not receive a third trimester ultrasound scan. These 5840 non-referred pregnant women had a mean number of 1.91 (SD 0.8) scans for the indication biometry. Routine ultrasound scans were performed at mean gestational ages of 28.9 (SD 0.6) and 34.7 (SD 0.6) weeks. Of 5049 women (84.4% of 5979 women) receiving usual care, who were not referred to obstetrician led care before 37 weeks' gestation, 41.0% (n=2072) did not receive a third trimester ultrasound scan.

Data from hospital files on fetuses with a suspected severe adverse perinatal outcome were used to analyse the level of adherence to the multidisciplinary protocol for diagnosing and managing fetal growth restriction. Of the pregnant women (n=107) referred to obstetrician led care because of a fetal abdominal circumference below the 10th centile, 97% (74 of 76) in the intervention strategy had additional ultrasound scans compared with 97% (30 of 31 women) receiving usual care. For 103 of these 107 women, 97% (71 of 73) in the intervention strategy and 87% (26 of 30) in the usual care group had Doppler scans. For women referred for slow fetal abdominal growth (n=162), 91% (112 of 123) in the intervention strategy had additional ultrasounds compared with 90% (35 of 39) in the usual care strategy. For 132 of 162 women referred for slow fetal abdominal growth, 78% (78 of 100) in the intervention strategy and 72% (23 of 32) in the usual care strategy had Doppler scans. None of these results was significantly different between the two strategies.

Table 3 shows the diagnostic accuracy for detecting SGA at birth (birth weight <10th centile) for both screening strategies. In the intervention strategy, more SGA neonates (22%) had an abdominal circumference below the 10th centile compared with SGA neonates in the usual care strategy (13%; P<0.001). Also, in the intervention strategy, significantly more SGA neonates (32%) had an abdominal circumference below the 10th centile or slow fetal growth compared with SGA neonates in the usual care strategy (19%; P<0.001) and specificity differed significantly (90% and 97%, respectively; P<0.001) (table 3).

The overall incidence of a severe adverse perinatal composite outcome was 1.7% (n=224); 1.7% (n=118) for the intervention strategy and 1.8% (n=106) for the usual care strategy (table 4). In a multilevel multivariable logistic regression analysis, routine ultrasonography in the third trimester was not associated with a significant reduction in severe adverse perinatal composite outcome (adjusted odds ratio 0.88, 95% confidence interval 0.70 to 1.20). The post hoc sensitivity analysis showed similar results

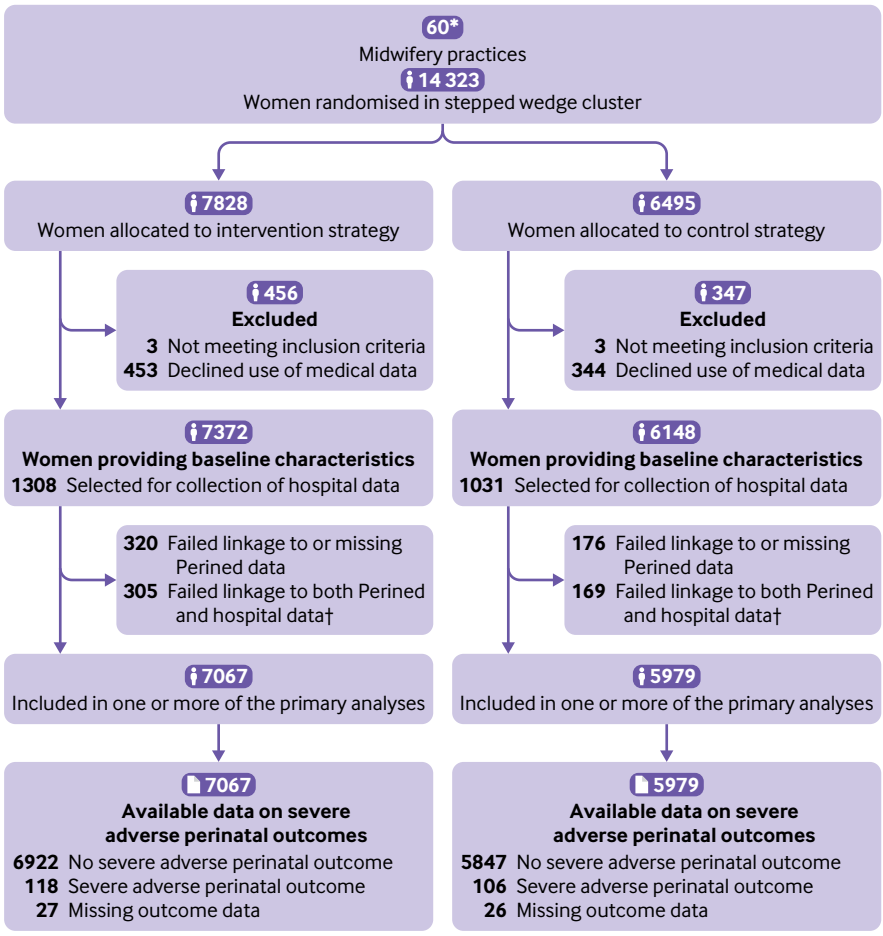

Fig 2 | Flow chart of IRIS study. *One midwifery practice withdrew from the study before crossover to the intervention strategy (59 practices participated in the study). †For 22 women (ie, 15 women in the intervention strategy and seven women in the control strategy) with missing data in the Perined database, data were available from hospital records

(0.83, 0.59 to 1.34). Secondary neonatal outcomes were also not significantly different between the two strategies. No significant differences were found in maternal morbidity and mortality between the groups (table 5). In a multilevel multivariable logistic regression analysis, routine ultrasonography in the third trimester was not related to the composite outcome of maternal peripartum morbidity or mortality (1.06, 0.95 to 1.18), or spontaneous labour and birth (1.00, 0.92 to 1.08). Routine ultrasonography in the third trimester was not associated with medical interventions in the peripartum period, with two exceptions (table 5). In a multilevel multivariable logistic regression adjusted for confounders, routine ultrasonography in the third trimester was associated with a higher incidence of induction of labour (1.16, 1.04 to 1.30) and a lower incidence of augmentation of labour (0.78, 0.71 to 0.85). Although higher numbers of births were observed in obstetrician led care in the intervention strategy compared with usual care strategy (65.0% *v* 63.3%; table 5), this association was not significant in a multilevel multivariable logistic regression adjusted for confounders (1.05, 0.96 to 1.14).

Table 1 | Characteristics of participating midwifery practices

Characteristics	Total (n=59)*	Start date of intervention strategy		
		1 May 2015 (n=19)*	1 September 2015 (n=20)	1 December 2015 (n=20)
Mean (SD) No of midwives per practice	5.0 (2.0)	5.2 (2.1)	5.0 (2.4)	4.9 (1.7)
Ultrasound biometry in midwifery practice (No (%))	19 (32)	9 (47)	3 (15)	7 (35)
Ultrasound biometry in sonography centre (No (%))	40 (68)	10 (53)	7 (85)	/13 (65)
Mean (SD) No of women in 2013	422.2 (226.0)	400.3 (155.0)	421.1 (252.3)	444.1 (261.7)
Mean (SD) nulliparous women (%)	46.7 (9.0)	44.7 (11.5)	46.9 (7.3)	48.4 (7.4)
Mean (SD) antepartum referrals to hospital care in 2013 (%)	37.5 (11.5)	37.1 (12.1)	37.8 (8.7)	37.6 (13.7)
Median (interquartile range) birth in midwife led care in 2013	88.0 (57.0; 128.0)	83.0 (58.0; 118.0)	87.5 (59.3; 128.3)	95.5 (52.5; 139.3)
Mean (SD) home births among midwife led births in 2013 (%)	49.3 (20.2)	49.6 (17.6)	52.4 (22.1)	45.9 (20.8)
Customised fundal height charts† (No (%))	27 (46)	9 (47)	6 (30)	12 (60)

*One practice dropped out in April 2015.
†None reported use of non-customised charts for fundal height measurements.

doi: 10.1136/bmj.l5517 | *BMJ* 2019;367:l5517 | the bmj

Table 2 | Personal and baseline characteristics of participants. Data are numbers (percentages) unless stated otherwise

Characteristics*	Intervention strategy (n=7067)	Control strategy (n=5979)	Total (n=13 046)
Parity:			
Nulliparous	3368 (47.7)	2928 (49.0)	6296 (48.3)
Multiparous	3632 (51.4)	3004 (50.2)	6636 (50.9)
Missing	67 (0.9)	47 (0.8)	114 (0.9)
Mean (SD) maternal age (years)	31.0 (4.5)	31.0 (4.3)	31.0 (4.4)
Body mass index:			
<18.5	232 (3.3)	185 (3.1)	417 (3.2)
18.5-25.0	4583 (64.9)	4025 (67.3)	8608 (66.0)
>25.0	2149 (30.4)	1707 (28.5)	3856 (29.6)
Missing	103 (1.5)	62 (1.0)	165 (1.3)
Ethnicity:			
Dutch	5096 (72.1)	4684 (78.4)	9780 (75.0)
Other Western	766 (10.8)	576 (9.6)	1342 (10.3)
Non-Western	1202 (17.0)	714 (11.9)	1916 (14.7)
Missing	3 (0.0)	5 (0.1)	8 (0.1)
Education:			
High	3770 (53.3)	3190 (53.4)	6960 (53.3)
Medium	2450 (34.7)	2115 (35.4)	4565 (35.0)
Low	700 (9.9)	588 (9.8)	1288 (9.9)
Missing	147 (2.1)	86 (1.4)	233 (1.8)
Work status:			
Employed	5763 (81.5)	5043 (84.3)	10 806 (82.8)
Unemployed	1188 (16.8)	863 (14.4)	2051 (15.7)
Missing	116 (1.6)	73 (1.2)	189 (1.4)
Marital status:			
Married or cohabiting	6457 (91.4)	5604 (93.7)	12 061 (92.4)
Not cohabiting	352 (5.0)	198 (3.3)	550 (4.2)
Single	167 (2.4)	121 (2.0)	288 (2.2)
Missing	91 (1.3)	56 (0.9)	147 (1.1)
Maternal smoking, alcohol use, or recreational drug use:			
Yes	2283 (32.3)	1956 (32.7)	4239 (32.5)
No	4759 (67.3)	4004 (67.0)	8763 (67.2)
Missing	25 (0.4)	19 (0.3)	44 (0.3)
Smoking during pregnancy:			
Yes	1003 (14.2)	835 (14.0)	1838 (14.1)
No	6048 (85.6)	5129 (85.8)	11 177 (85.7)
Missing	16 (0.2)	15 (0.2)	31 (0.2)
Alcohol consumption during pregnancy:			
Yes	1679 (23.8)	1481 (24.8)	3160 (24.2)
No	5362 (75.9)	4482 (75.0)	9844 (75.5)
Missing	26 (0.4)	16 (0.3)	42 (0.3)
Recreational drug use during pregnancy:			
Yes	86 (1.2)	70 (1.2)	156 (1.2)
No	6963 (98.5)	5896 (98.6)	12 859 (98.6)
Missing	18 (0.3)	13 (0.2)	31 (0.2)
Medical indication for fetal ultrasound determined at inclusion:			
Yes	1324 (18.7)	1121 (18.7)	2445 (18.7)
No	5743 (81.3)	4858 (81.3)	10 601 (81.3)
Type of medical indications determined at inclusion:			
Fundal height measurements unreliable*	137 (1.9)	119 (2.0)	256 (2.0)
Previous SGA neonate*	176 (2.5)	161 (2.7)	337 (2.6)
Other (not biometry related) indication†	1012 (14.3)	844 (14.1)	1856 (14.2)
Fetal sex:			
Female	3480 (49.2)	2923 (48.7)	6403 (48.9)
Male	3585 (50.7)	3055 (50.9)	6640 (50.9)
Missing	2 (0)	1 (0)	3 (0)

SGA=small for gestational age (birth weight <10th centile according to Dutch birth weight curve).[27]
*Percentages do not always add up to 100% because of combinations of multiple medical indications for ultrasonography in one pregnancy.
†Other clinical indications for ultrasonography.

Discussion

In this large, pragmatic, nationwide, stepped wedge cluster randomised trial in low risk pregnant women, using a multidisciplinary protocol for detecting and managing fetal growth restriction, routine ultrasonography in the third trimester improved prenatal detection of neonates who were small for gestational age (SGA) compared with usual care. But this approach did not result in a significantly lower incidence of severe adverse perinatal outcomes. Routine ultrasonography was not associated with significantly improved secondary neonatal outcomes or secondary maternal composite peripartum outcomes. Routine ultrasonography was associated with a higher incidence of induction of labour.

Table 3 | Diagnostic accuracy of two screening strategies with small for gestational age (SGA) at birth as outcome

	Intervention strategy (n=6909)		Control strategy (n=5498)		Routine fetal biometry at 28-30 weeks' gestation (n=6909)*		Routine fetal biometry at 34-36 weeks' gestation (n=6888)*	
Ultrasound variables	SGA (n=556)	No SGA (n=6353)	SGA (n=407)	No SGA (n=5091)	SGA (n=556)	No SGA (n=6353)	SGA (n=554)	No SGA (n=6334)
Total (No)								
AC <P10 or slow growth	179	651	78	153	—	—	—	—
AC <P10	122	156	53	60	74	126	63	44
Slow growth	76	522	34	100	—	—	—	—
Sensitivity (% (95% CI))								
AC <P10 or slow growth	32 (28 to 36)		19 (15 to 23)†		—	—	—	—
AC <P10	22 (19 to 26)		13 (10 to 17)†		13 (11 to 16)		11 (9 to 14)‡	
Slow growth	14 (11 to 17)		8 (6 to 11)†		—	—	—	—
Specificity (% (95% CI))								
AC <P10 or slow growth	90 (89 to 91)		97 (96 to 97)†		—	—	—	—
AC <P10	98 (97 to 98)		99 (98 to 99)†		98 (97 to 98)		99 (99 to 100)§	
Slow growth	92 (91 to 93)		98 (97 to 98)†		—	—	—	—
Positive predictive value (% (95% CI))								
AC <P10 or slow growth	22 (19 to 24)		34 (28 to 40)		—	—	—	—
AC <P10	44 (39 to 49)		47 (38 to 56)		37 (31 to 44)		59 (50 to 68)	
Slow growth	13 (10 to 15)		25 (19 to 33)		—	—	—	—
Negative predictive value (% (95% CI))								
AC <P10 or slow growth	94 (93 to 94)		94 (93 to 94)		—	—	—	—
AC <P10	93 (93 to 94)		93 (93 to 94)		93 (92 to 93)		93 (92 to 93)	
Slow growth	92 (92 to 93)		93 (92 to 93)		—	—	—	—

For diagnostic accuracy analyses, 639 cases with incomplete data on third trimester ultrasound scans were excluded. Intervention strategy=prenatal screening for fetal growth restriction based on routine biometry scans in third trimester in combination with serial fundal height measurements and ultrasonography if clinically indicated. Control strategy=prenatal screening for fetal growth restriction based on serial fundal height measurements and ultrasonography only if clinically indicated.
SGA=small for gestational age (birth weight <10th centile (<P10) of Dutch national (Perined) birth weight curve[27]); AC <P10=abdominal circumference estimated by fetal biometry scan <P10 on Dutch growth curve[24]; slow growth=decrease of ≥20 centiles of abdominal circumference on Dutch growth curve with interval of at least two weeks from gestational age ≥26 weeks and 0 days.[13]
*Numbers for routine biometry scans at 28-30 weeks' gestation and at 34-36 weeks' gestation differ slightly because of missing values.
†χ^2 tests showed that sensitivities of AC <P10 or slow growth (P<0.001), AC <P10 (P<0.001), and slow growth (P=0.010) were significantly higher in the intervention strategy than in the control strategy, whereas specificities of these measures were significantly higher in the control strategy (all P<0.001).
‡A McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks' gestation and 34-36 weeks' gestation (P=0.38).
§A McNemar test showed a higher specificity of the fetal biometry scan at 34-36 weeks' gestation compared with the scan at 28-30 weeks' gestation (P<0.001).

Comparison with other studies

Our trial addressed important shortcomings of previous studies.[15] Modern ultrasound equipment was used, sonographers met predefined quality criteria, and a multidisciplinary protocol was applied. Nevertheless, our findings are in line with a previous meta-analysis, which failed to show better perinatal outcomes in women who received routine ultrasound scans after 24 weeks' gestation, based on 13 previous trials (n=34 980).[15]

Possible explanations for our findings are: routine ultrasound fetal biometry is ineffective in detecting fetal growth restriction and preventing subsequent adverse outcomes in low risk pregnancies; adding routine ultrasound scans in the third trimester to usual care does not yield major benefits because women receiving such care already undergo one clinically indicated ultrasound scan on average in the third trimester; the quality of ultrasonography was insufficient; and using fetal abdominal circumference below the 10th centile (in combination with biometric measures of slow growth) on a population based curve is ineffective in detecting fetal growth restriction, and better methods are required.

The number of babies with a birth weight below the 10th centile was around 8% in our study. Previous research suggests that most of these babies are likely to be constitutionally small rather than growth restricted and would not be at increased risk of severe adverse perinatal outcomes.[11] Women assigned to the usual care strategy had one clinically indicated ultrasound scan on average in the third trimester of pregnancy. Moreover, nearly one in five women in the intervention and usual care strategies had an indication for an ultrasound scan in the third trimester that was identified at inclusion in the study. Routine antenatal ultrasonography might therefore have little or no added benefit in detecting SGA neonates at risk of adverse outcomes compared with clinically indicated ultrasonography as part of usual care in the third trimester.

Another explanation for our findings might be that the quality of ultrasound scans was too low. Similar to the Pregnancy Outcome Prediction (POP) study, we found that sensitivity rates were higher for the intervention strategy with routine ultrasonography compared with usual care strategy with clinically indicated ultrasonography, although specificity rates were lower.[11] Thus our findings suggest that repeated ultrasonography measures increase the detection of SGA but are also accompanied by higher false positive rates. In the intervention strategy, for an abdominal circumference below the 10th centile or slow growth in abdominal circumference the sensitivity in detecting birth weight below the 10th centile was 32% and

Table 4 | Severe adverse perinatal (composite) outcome and secondary neonatal outcomes

Perinatal and neonatal outcomes	Total No*	No, No of events (%)		Odds ratio (95% CI) or P value†	Adjusted odds ratios or β linear regression coefficient (95% CI)‡
		Intervention strategy	Control strategy		
Severe adverse perinatal outcome					
Composite of severe adverse perinatal outcome§	12 993	7040, 118 (1.7)	5953, 106 (1.8)	0.94 (0.72 to 1.23)	0.88 (0.70 to 1.20)
Perinatal death, 28 weeks' gestational age to 7 days postnatal:	13 043	7066, 14 (0.2)	5977, 15 (0.3)	0.79 (0.38 to 1.64)	NA
Antepartum death		8 (0.1)	9 (0.2)		
Intrapartum death		1 (0)	0 (0)		
Neonatal death, 0-7 days postnatal		5 (0.1)	6 (0.1)		
Apgar score <4 at five minutes	13 039	7066, 21 (0.3)	5976, 27 (0.5)	0.66 (0.37 to 1.16)	NA
Impaired consciousness	12 995	7040, 7 (0.1)	5955, 9 (0.2)	0.66 (0.25 to 1.77)	NA
Asphyxia (cord blood arterial base excess ≤12 mmol/L)	13 012	7047, 74 (1.1)	5965, 70 (1.2)	0.89 (0.64 to 1.24)	0.88 (0.62 to 1.23)
Seizures on ≥2 occasions <72 hours after birth	12 997	7041, 6 (0.1)	5956, 6 (0.1)	0.85 (0.27 to 2.62)	NA
Assisted ventilation >24 hours by endotracheal tube <72 hours after birth	12 997	7040, 27 (0.4)	5957, 19 (0.3)	1.20 (0.67 to 2.17)	NA
Septicaemia ascertained by blood culture	12 997	7041, 7 (0.1)	5956, 4 (0.1)	1.48 (0.43 to 5.06)	NA
Meningitis ascertained by cerebrospinal fluid culture	12 997	7041, 1 (0)	5956, 2 (0)	0.60	NA
Bronchopulmonary dysplasia: need for oxygen ≥36 weeks' gestational age, confirmed by radiography	12 997	7041, 0 (0)	5956, 1 (0)	0.46	NA
Intraventricular haemorrhage grade 3 or 4, confirmed by ultrasonography/autopsy	12 995	7041, 0 (0)	5954, 1 (0)	0.46	NA
Cystic periventricular leucomalacia, confirmed by ultrasonography/autopsy	12 997	7041, 0 (0)	5956, 0 (0)	NA	NA
Necrotising enterocolitis, confirmed by radiography/surgery/autopsy	12 997	7041, 1 (0)	5956, 1 (0)	NA	NA
Secondary neonatal outcome					
Mean (SD) birth weight (g)	13 035	7065, 3478 (518)	5970, 3487 (511)	0.35	−1.75 (−20.3 to 16.8)
SGA at birth	13 007	7046, 572 (8.1)	5961, 447 (7.5)	1.09 (0.96 to 1.24)	1.03 (0.90 to 1.18)
LGA at birth	13 007	7046, 748 (10.6)	5961, 640 (10.7)	0.99 (0.88 to 1.10)	0.98 (0.87 to 1.10)
Mean (SD) gestational age at birth (weeks)	13 043	7066, 39.7 (1.61)	5977, 39.7 (1.57)	0.98	0.02 (−0.04 to 0.07)
Prematurity (<37 weeks):	13 043	7066, 283 (4.0)	5977, 251 (4.2)	0.95 (0.80 to 1.13)	0.93 (0.77 to 1.12)
32-37 weeks		249 (3.5)	225 (3.8)		
<32 weeks		34 (0.5)	26 (0.4)		
Neonatal death, 8-28 days after birth¶	1520	759, 1 (0.1)	761, 1 (0.1)	NA	NA
Detection of congenital abnormalities	13 038	7062, 167 (2.4)	5976, 117 (2.0)	1.21 (0.96 to 1.54)	0.97 (0.74 to 1.26)

Presented outcome variables were prespecified in the Netherlands Trial Registry (NTR4367).
NA=not applicable (numbers too small); SGA=small for gestational age (birth weight <10th centile (<P10) of Dutch (Perined) birth weight curve[27]); LGA=large for gestational age (birth weight >P90 of Dutch (Perined) birth weight curve[27]).
*Numbers differ because of missing values.
†P values based on independent *t* tests for continuous outcome variables and Fisher's exact tests for categorical variables because of empty cells or expected cell counts <5.
‡Adjusted odds ratios and 95% confidence intervals were calculated by multilevel, multivariable logistic regression adjusted for clustering and for potential confounders, including maternal age; body mass index; smoking, alcohol, or recreational drug use; and midwifery practice size; parity; educational level; employment status; marital status; and sex of infant. In the various multilevel, multivariable models, the amount of missing values for potential confounders was ≤4.4%.
§One or more of perinatal death, Apgar score <4 at five minutes, impaired consciousness, asphyxia, seizures, >24 hour assisted ventilation, haemorrhage; meningitis, bronchopulmonary dysplasia, intraventricular haemorrhage, cystic periventricular leucomalacia, or necrotising enterocolitis.
¶Information based only on hospital records.

the positive predictive value was 22%. The positive predictive value of an abdominal circumference below the 10th centile was higher for the second routine scan (59%) at 34-36 weeks' gestation than for the first scan at 28-30 weeks' gestation (37%), whereas negative predictive values were similar, in line with the findings of the POP study.[11] Thus late third trimester scans seem to have more diagnostic accuracy than earlier ones. Also, sensitivity was 22% for an abdominal circumference below the 10th centile in the intervention strategy, similar to the findings of a large nationwide French population based study (n=14 100).[9] Single centre studies with smaller sample sizes, however, showed better accuracy, which might be because a limited number of dedicated sonographers performed the ultrasound scans.[11 16] Sensitivity and specificity rates vary with the prevalence of disease.[34] As our study was conducted in a low risk population, a lower sensitivity and higher specificity might be expected than in the general population. Although the accuracy of ultrasonography in detecting low birth weight is higher than that of fundal height measurement, systematic errors in the prediction of SGA neonates using fetal abdominal circumference or estimated fetal weight limit its accuracy because these predictions are based on an estimation with an algorithm.[35] Nevertheless, even if not used routinely, fetal biometry is frequently performed. Although quality assurance systems have been developed for the anomaly scan, much less attention is paid to developing systems to guarantee the quality of fetal biometry in the third trimester.[36] Given that low birth weight is associated with adverse perinatal outcomes, the quality of fetal biometry should be maintained to the highest standard.

Even if the quality of ultrasonography is improved, the most appropriate screening test for fetal growth restriction is not clear. The centiles for measurements of fetal biometry, estimated fetal weight and birth weight, depend on the charts used. Some growth charts are universal and prescriptive, which means they show optimal growth in a healthy population, whereas others are customised for maternal characteristics, such as parity and ethnicity. So far, a consensus on the best charts for fetal growth and birth weight has not been reached.[37] A recent large retrospective study

Table 5 | Maternal outcomes and peripartum interventions

Maternal outcomes and peripartum interventions	Total No*	No, No of events (%)		Odds ratio (95% CI) or P value	Adjusted odds ratio (95% CI)†
		Intervention strategy	Control strategy		
Peripartum interventions and mode of delivery					
Composite of spontaneous vaginal birth without intervention‡	12 490	6663, 2974 (44.6)	5827, 2650 (45.5)	0.97 (0.90 to 1.04)	1.00 (0.92 to 1.08)
Induction of labour§	12 984	7034, 1118 (15.9)	5950, 813 (13.7)	1.19 (1.08 to 1.32)	1.16 (1.04 to 1.30)
Vacuum or forceps assisted birth	13 044	7065, 538 (7.6)	5979, 506 (8.5)	0.89 (0.79 to 1.01)	0.90 (0.78 to 1.04)
Caesarean section:	13 044	7065	5979	1.01 (0.91 to 1.12)	1.00 (0.90 to 1.11)
Primary		414 (5.9)	342 (5.7)		
Secondary		555 (7.9)	472 (7.9)		
Total		969 (13.7)	814 (13.6)		
Augmentation of labour	12 713	6894, 1902 (27.6)	5819, 1839 (31.6)	0.82 (0.76 to 0.89)	0.78 (0.71 to 0.85)
Pharmacological pain relief	12 337	6582, 2786 (42.3)	5755, 2378 (41.3)	1.04 (0.97 to 1.12)	1.02 (0.94 to 1.11)
Maternal peripartum morbidity and mortality					
Composite outcome of maternal morbidity/mortality¶	11 787	6327, 1123 (17.7)	5460, 910 (16.7)	1.08 (0.98 to 1.19)	1.06 (0.95 to 1.18)
Maternal death, up to 42 days after birth	13 044	7065, 0 (0)	5979, 0 (0)	NA	NA
Hypertensive disorders	11 638	6235, 539 (8.6)	5403, 445 (8.2)	1.05 (0.93 to 1.20)	1.06 (0.91; 1.22)
Postpartum haemorrhage	13 044	7065, 472 (6.7)	5979, 402 (6.7)	0.99 (0.87 to 1.14)	0.98 (0.85 to 1.14)
Third or fourth degree perineal trauma	13 044	7065, 186 (2.6)	5979, 134 (2.2)	1.18 (0.94 to 1.48)	1.17 (0.92 to 1.47)
Other secondary maternal outcomes					
Birth in obstetrician led care	13 018	7047, 4580 (65.0)	5971, 3778 (63.3)	1.08 (1.00 to 1.16)	1.05 (0.96 to 1.14)
Non-cephalic presentation at start of labour	12 922	6998, 203 (2.9)	5924, 198 (3.3)	0.86 (0.71 to 1.05)	0.87 (0.71 to 1.07)

Presented outcome variables were prespecified in the Netherlands Trial Registry (NTR4367) except for hypertension and pre-eclampsia, which were combined into one variable—that is, hypertensive disorders were defined as a diastolic blood pressure ≥95 mm Hg with or without proteinuria or ≥90 mm Hg with proteinuria, because of the small number of cases (n=12) with proteinuria and diastolic blood pressure ≥90 mm Hg.
NA=not applicable (numbers were too small).
*Total numbers differ because of missing values.
†Adjusted odds ratios and 95% confidence intervals were calculated by multilevel, multivariable logistic regression adjusted for clustering, midwifery practice size (potential fixed factor), and potential confounders, including maternal age; body mass index; smoking, alcohol, or recreational drug use; parity; educational level; employment status; marital status; sex of infant; and midwifery practice size. In the various multilevel, multivariable models, the amount of missing values for potential confounders was ≤4.4%.
‡Defined as spontaneous vaginal birth, without induction or augmentation of labour, and with no drug pain relief, vacuum or forceps assisted birth, or caesarean section.
§Based on available data (n=12 971) for both induction of labour and prematurity status, labour was induced in 12% (n=33) of 283 premature infants in the intervention strategy and 12% (n=29) of 251 premature infants in the control strategy.
¶Composite outcome was one or more of maternal peripartum death within 42 days after delivery, maternal hypertension (with or without proteinuria), postpartum haemorrhage, and third or fourth degree perineal trauma.

showed an increase in identifying the risk of stillbirth when customised fetal growth charts were used.[38] Nonetheless, in the prospective POP study, compared with universal charts, customised charts did not result in an increased association between estimated fetal weight below the 10th centile and neonatal morbidity.[11] Before the start of our study, the Royal Dutch Association of Midwives issued guidelines on fetal growth restriction and recommended customised fetal growth curves. Technical difficulties prevented integration of this approach into many midwifery practices, whereas hospitals did not use customised curves. The applied multidisciplinary protocol therefore recommended serial measurements of fundal height but not plotting them on customised growth curves.

Even if birth weight can be estimated accurately, many small babies are constitutionally small but healthy.[11] In the POP study, about 70% of fetuses with an estimated fetal weight below the 10th centile were not growth restricted and had similar perinatal outcomes compared to those with a greater estimated fetal weight.[11] Disadvantages associated with routine ultrasound scans in the third trimester might be increased levels of emotional distress in women because of an inaccurate suspicion of fetal growth restriction and increased exposure to additional diagnostic tests, monitoring, and obstetric interventions.[39 40] That the incidence of most obstetric interventions was not significantly different between the groups is reassuring but we found a higher incidence of induction of labour associated with the intervention strategy, with no evidence of better perinatal outcomes. A study also showed that a suspicion of SGA was associated with a higher incidence of initiated delivery by the provider.[9] The incidence of augmentation of labour was lower in the intervention strategy. Oxytocin would, however, have been used as part of the induction of labour strategy but this would not have been recorded separately in the Perined database. Inducing labour artificially is more invasive than augmentation of labour that has started spontaneously, and overuse of induction of labour in the absence of clear beneficial effects is a growing concern.[41 42] Overall, the findings of this pragmatic trial do not support a policy of routine ultrasound screening in the third trimester for low risk pregnant women to reduce severe adverse perinatal outcomes.

As estimated fetal weight and abdominal circumference alone are not good markers of fetal growth restriction, more sensitive methods are needed. These include other ultrasound markers of fetal compromise, such as Doppler indices.[18] The POP study showed that the combination of ultrasonography in the third trimester and measurement of placental biomarkers in the mother's blood (the soluble fms-like tyrosine kinase 1:placental growth factor ratio) strongly predicted adverse pregnancy outcomes related to fetal growth restriction, suggesting that biomarkers might be useful in detecting growth restriction.[43] Moreover, women are aware of fetal movements, which are a sign of fetal wellbeing. A change in fetal activity

might be a sign of fetal growth restriction.[44] Fetal death, often associated with fetal growth restriction, is usually preceded by reduced fetal movements.[44] More research is needed, however, to determine what type of information women should receive about the wellbeing of the fetus.[44] Further research is also needed to address the histopathological mechanisms that might underlie the association between fetal growth restriction and perinatal death, including sudden intrauterine unexplained death, and to improve preventive strategies.[45 46]

Strengths and limitations of this study

Our study has strengths and limitations. The cluster randomised design controlled for unknown confounders at the cluster level and limited contamination between the study strategies, which might occur in individual randomised trials. The stepped wedge design reduced confounding owing to differences between midwifery practices because each practice applied the control and intervention strategy for some of the time. Sonographers met predefined quality criteria, and a multidisciplinary protocol was developed for detecting and managing fetal growth restriction to achieve the best quality care possible in a pragmatic nationwide study.[13 20]

We did not achieve our required sample size of 15 000 women. Owing to the stepped wedge design, it was not possible to extend the data collection period because the midwifery practices had adopted the intervention strategy at the end of the study period. We cannot therefore completely rule out that the study lacked the statistical power to determine if routine ultrasonography has a beneficial or harmful effect on perinatal outcomes compared with usual care. Although we found a difference of only 0.1% between the two strategies, it is unlikely that this difference would have met the preset meaningful difference of 0.54% had the sample size been larger. Also, we used registration data as an initial screening for potential severe adverse perinatal outcomes. Information was also obtained from many hospital records, but for most women only routine registration data for adverse outcomes were available. Because of the inherent limitations of these data, several outcomes might have been misclassified as normal, resulting in an underestimation of the primary outcome for both strategies. But we do not expect that this has biased the comparison between the two strategies as the incidence of adverse outcomes was similar to our estimations. Furthermore, because of the collinearity of time of inclusion period and study condition, we were unable to adjust for time. The effect estimates of our main analyses might therefore be overestimated.

Finally, our study was conducted in one country (the Netherlands) where primary antenatal care of uncomplicated pregnancies is provided by midwives who are educated, trained, and officially registered as independent health practitioners.[21] When risk factors or complications occur, women are referred to obstetrician led care. For about 90% of women in the Netherlands, antenatal care is midwife led initially, and about 50% of women start labour in midwife led care.[26] Also, most of the recommendations of the multidisciplinary protocol for diagnosing and managing suspected fetal growth restriction in our study are similar to international guidelines in other countries (eg, the Royal College of Obstetricians and Gynaecologist),[25] making our results relevant to low risk populations in other international care contexts.

Conclusions

Our pragmatic nationwide trial found that routine ultrasonography in the third trimester of pregnancy and with a multidisciplinary protocol for detecting and treating fetal growth restriction was associated with a moderately increased antenatal detection of SGA neonates and induction of labour. This strategy was not, however, associated with a reduction in the incidence of severe adverse perinatal outcomes in low risk pregnancies compared with usual care including clinically indicated ultrasonography. Based on our findings, we cannot recommend routine ultrasonography in the third trimester in low risk pregnancies. Challenges for future research are to identify the most appropriate fetal growth and birth weight charts and to develop more sensitive and effective methods to detect fetal growth restriction. Such methods include other ultrasound markers of fetal compromise, maternal and placental biomarkers, and maternal awareness of fetal wellbeing.

AUTHOR AFFILIATIONS

[1]Amsterdam University Medical Centre, Vrije Universiteit Amsterdam, Department of Midwifery Science, AVAG/Amsterdam Public Health, Van der Boechhorststraat 7, 1081 BT Amsterdam, Netherlands

[2]Dutch Professional Organisation of Sonographers, Woerden, Netherlands

[3]Amsterdam University Medical Centre, University of Amsterdam, Department of Obstetrics and Gynaecology, Amsterdam Reproduction and Development, Amsterdam, Netherlands

[4]Amsterdam University Medical Centre, Vrije Universiteit Amsterdam, Department of Epidemiology and Biostatistics, Amsterdam Public Health, Amsterdam, Netherlands

[5]Amsterdam University Medical Centre, Vrije Universiteit Amsterdam, Department of General Practice and Elderly Care Medicine, Amsterdam Public Health, Amsterdam, Netherlands

[6]Department of Obstetrics and Gynaecology, Erasmus Medical University Centre, Rotterdam, Netherlands

We thank the participating women, midwives, obstetricians, sonographers, IRIS study data management and research assistants team, and staff at participating midwifery practices and hospitals in the Netherlands; Inge Zonneberg, Wessel Ganzevoort, and Sanne Gordijn for their medical advice during the development of the case report forms for the hospital record data collection; Tony Sheldon for the language editing of the manuscript; and the Perinatal Registry of the Netherlands (Perined) for the use of their database.

The IRIS study group: Anneloes L van Baar, Utrecht University, Utrecht, Netherlands; Joke M J Bais, Medical Centre Alkmaar, Alkmaar, Netherlands; Gouke J Bonsel, University Medical Centre Utrecht, Utrecht, Netherlands; Judith E Bosmans, VU University Amsterdam, Amsterdam, Netherlands; Jeroen van Dillen, Radboud University Medical Centre, Nijmegen, Netherlands; Noortje T L van Duijnhoven, Radboud University Medical Centre, Nijmegen, Netherlands; William A Grobman, Northwestern University, Chicago, IL, USA; Henk Groen, University of Groningen, Groningen, Netherlands; Chantal W P M Hukkelhoven, Wageningen University and Research, Wageningen, Netherlands; Trudy Klomp, Amsterdam University Medical Centre/

AVAG, Amsterdam, Netherlands; Marjolein Kok, Amsterdam University Medical Centre, Amsterdam, Netherlands; Marlou L de Kroon, University of Groningen, Groningen, Netherlands; Maya Kruijt, Dutch Society of Obstetrics and Gynaecology NVOG, Utrecht, Netherlands; Anneke Kwee, University Medical Centre Utrecht, Utrecht, Netherlands; Sabina Ledda, Midwifery practice het Palet/BEN, Rotterdam, Netherlands; Harry N Lafeber, Amsterdam University Medical Centre, Netherlands; Jan M M van Lith, Leiden University Medical Centre, Leiden, Netherlands; Ben Willem Mol, University of Adelaide, Adelaide, Australia; Bert Molewijk, Amsterdam University Medical Centre, Amsterdam, Netherlands; Marianne Nieuwenhuijze, Academie Verloskunde Maastricht, Maastricht, Netherlands; Guid Oei, Maxima Medical Centre, Eindhoven, Netherlands; Cees Oudejans, Amsterdam University Medical Centre, Amsterdam, Netherlands; K Marieke Paarlberg, Gelre Hospitals, location Apeldoorn, Netherlands; Aris T Papageorghiou, University of Oxford, Oxford, UK; Uma M Reddy, Yale University, New Haven, CT, USA; Paul De Reu, Prenataal Screenigscentrum "de Meierij," Eindhoven, Netherlands; Marlies Rijnders, TNO, Leiden, Netherlands; Alieke de Roon-Immerzeel, Royal Dutch Organisation of Midwives, Netherlands; Connie Scheele, University Medical Centre Utrecht/BEN, Utrecht, Netherlands; Sicco A Scherjon, University Medical Centre Groningen, Groningen, Netherlands; Rosalinde Snijders, Amsterdam University Medical Centre, Amsterdam, Netherlands; Marc E Spaanderman, University Medical Centre Maastricht, Maastricht, Netherlands; Pim W Teunissen, Amsterdam University Medical Centre, Amsterdam, Netherlands; Hanneke W Torij, Rotterdam University of Applied Sciences, Rotterdam, Netherlands; Tanja G Vrijkotte, Amsterdam University Medical Centre, Amsterdam, Netherlands; Myrte Westerneng, Amsterdam University Medical Centre, Amsterdam, Netherlands Kristel C Zeeman, University of Amsterdam, Amsterdam, Netherlands; and Jun Jim Zhang, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Contributors: JH and VV contributed equally to the manuscript and are first authors. AdJ, AF, HEvdH, PJ, VV, and EP conceived and designed the study. AdJ, PJ, LV, VV, and JH were involved in the development, implementation, or data collection of the study. JH, JW, and VV conducted data analyses. JH, VV, and AdJ drafted the manuscript and interpreted the results. All authors critically revised the manuscript for important intellectual content, approved its final version, and agree to be accountable for all aspects of the work presented in this manuscript. AdJ is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: This study was supported by a grant from the Netherlands Organisation for Health Research and Development (ZonMw; grant No 209030001). This funding source had no role in study design, data collection, data analysis, data interpretation, writing of the scientific article, or the decision to submit the paper for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: for the current study (the IRIS study), AdJ and JW received funding from the Netherlands Organisation for Health Research and Development; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The IRIS study was approved by the Dutch Institutional Review Board of the VU Medical University Centre Amsterdam (reference No 2013.409).

Data sharing: No additional data available.

The manuscript's guarantor (AdJ) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

1 Figueras F, Gratacós E. Update on the diagnosis and classification of fetal growth restriction and proposal of a stage-based management protocol. *Fetal Diagn Ther* 2014;36:86-98. doi:10.1159/000357592
2 Crispi F, Crovetto F, Gratacos E. Intrauterine growth restriction and later cardiovascular function. *Early Hum Dev* 2018;126:23-7. doi:10.1016/j.earlhumdev.2018.08.013
3 Kesavan K, Devaskar SU. Intrauterine growth restriction: postnatal monitoring and outcomes. *Pediatr Clin North Am* 2019;66:403-23. doi:10.1016/j.pcl.2018.12.009
4 Flenady V, Koopmans L, Middleton P, et al. Major risk factors for stillbirth in high-income countries: a systematic review and meta-analysis. *Lancet* 2011;377:1331-40. doi:10.1016/S0140-6736(10)62233-7
5 Frøen JF, Gardosi JO, Thurmann A, Francis A, Stray-Pedersen B. Restricted fetal growth in sudden intrauterine unexplained death. *Acta Obstet Gynecol Scand* 2004;83:801-7. doi:10.1111/j.0001-6349.2004.00602.x
6 Batalle D, Eixarch E, Figueras F, et al. Altered small-world topology of structural brain networks in infants with intrauterine growth restriction and its association with later neurodevelopmental outcome. *Neuroimage* 2012;60:1352-66. doi:10.1016/j.neuroimage.2012.01.059
7 NICE. Antenatal care for uncomplicated pregnancies. Clinical guideline [CG62], 2008. www.rcog.org.uk/en/guidelines-research-services/guidelines/antenatal-care/
8 McCowan L, Horgan RP. Risk factors for small for gestational age infants. *Best Pract Res Clin Obstet Gynaecol* 2009;23:779-93. doi:10.1016/j.bpobgyn.2009.06.003
9 Monier I, Blondel B, Ego A, Kaminiski M, Goffinet F, Zeitlin J. Poor effectiveness of antenatal detection of fetal growth restriction and consequences for obstetric management and neonatal outcomes: a French national study. *BJOG* 2015;122:518-27. doi:10.1111/1471-0528.13148
10 Marconi AM, Ronzoni S, Bozzetti P, Vailati S, Morabito A, Battaglia FC. Comparison of fetal and neonatal growth curves in detecting growth restriction. *Obstet Gynecol* 2008;112:1227-34. doi:10.1097/AOG.0b013e31818bdc7e
11 Sovio U, White IR, Dacey A, Pasupathy D, Smith GCS. Screening for fetal growth restriction with universal third trimester ultrasonography in nulliparous women in the Pregnancy Outcome Prediction (POP) study: a prospective cohort study. *Lancet* 2015;386:2089-97. doi:10.1016/S0140-6736(15)00131-2
12 Gordijn SJ, Beune IM, Thilaganathan B, et al. Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound Obstet Gynecol* 2016;48:333-9. doi:10.1002/uog.15884
13 Verfaille V, de Jonge A, Mokkink L, et al, IRIS study group. Multidisciplinary consensus on screening for, diagnosis and management of fetal growth restriction in the Netherlands. *BMC Pregnancy Childbirth* 2017;17:353. doi:10.1186/s12884-017-1513-3
14 Terwischa van Scheltinga JA, Scherjon SA, van Dillen J. *NVOG-richtlijn Foetale groeirestrictie (FGR).* [Dutch Association of Obstetrics and Gynaecology (NVOG)-guideline Fetal Growth Restriction (FGR)] Dutch Association of Obstetrics and Gynaecology, 2017.
15 Bricker L, Medley N, Pratt JJ. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *Cochrane Database Syst Rev* 2015;(6):CD001451. doi:10.1002/14651858.CD001451.pub4
16 De Reu PA, Smits LJ, Oosterbaan HP, Nijhuis JG. Value of a single early third trimester fetal biometry for the prediction of birth weight deviations in a low risk population. *J Perinat Med* 2008;36:324-9. doi:10.1515/JPM.2008.057
17 Beentjes M, Roon-Immerzeel A, Zeeman K. *Opsporing van foetale groeivertraging KNOV-standaard.* [Detecting intrauterine growth restriction KNOV-standard] Utrecht, Netherlands: Koninklijke Nederlandse Organisatie van Verloskundigen (KNOV), 2013.
18 McCowan LM, Figueras F, Anderson NH. Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. *Am J Obstet Gynecol* 2018;218(2S):S855-68. doi:10.1016/j.ajog.2017.12.004
19 Haynes B, Sacket DL, Gordon GH, et al. *Clinical epidemiology: how to do clinical practice research.* Lippincott Williams and Wilkins, 2016.
20 Henrichs J, Verfaille V, Viester L, et al, IRIS Study Group. Effectiveness and cost-effectiveness of routine third trimester ultrasound screening for intrauterine growth restriction: study protocol of a nationwide stepped wedge cluster-randomized trial in The Netherlands (The IRIS Study). *BMC Pregnancy Childbirth* 2016;16:310. doi:10.1186/s12884-016-1104-8
21 Prins M. *Praktische verloskunde.* [Practical maternity care] Bohn Stafleu van Loghum, 2009. doi:10.1007/978-90-313-8144-9
22 NVOG. *Foetale echoscopie.* [Fetal ultrasonography] NVOG, 2009.
23 NVOG. *Foetale biometrie.* [Fetal biometry] NVOG, 2008.
24 Verburg BO, Steegers EA, De Ridder M, et al. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008;31:388-96. doi:10.1002/uog.5225
25 RCOG. *Small-for-gestational-age fetus, investigation and management (Green-top Guideline No 31).* Royal College of Obstetricians and Gynaecologists, 2013.

26 Dijs-Elsinga J, Groenendaal F, van Huis AM, et al. *Perined. Perinatale Zorg in Nederland 2014.* [Perined. Perinatal Care in the Netherlands 2014] Perined, 2015.
27 Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM. New Dutch reference curves for birthweight by gestational age. *Early Hum Dev* 2009;85:737-44. doi:10.1016/j.earlhumdev.2009.09.008
28 Pagel C, Prost A, Lewycka S, et al. 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. doi:10.1186/1745-6215-12-151
29 Altman DG, Bland JM. Diagnostic tests 2: Predictive values. *BMJ* 1994;309:102. doi:10.1136/bmj.309.6947.102
30 Dormann CF, Elith J, Bacher S, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 2013;36:27-46. doi:10.1111/j.1600-0587.2012.07348.x
31 Jaddoe VW, Verburg BO, de Ridder MA, et al. Maternal smoking and fetal growth characteristics in different periods of pregnancy: the generation R study. *Am J Epidemiol* 2007;165:1207-15. doi:10.1093/aje/kwm014
32 Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull World Health Organ* 1987;65:663-737.
33 Moineddin R, Matheson FI, Glazier RH. A simulation study of sample size for multilevel logistic regression models. *BMC Med Res Methodol* 2007;7:34. doi:10.1186/1471-2288-7-34
34 Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;16:981-91. doi:10.1002/(SICI)1097-0258(19970515)16:9<981::AID-SIM510>3.0.CO;2-N
35 Lappen JR, Myers SA. The systematic error in the estimation of fetal weight and the underestimation of fetal growth restriction. *Am J Obstet Gynecol* 2017;216:477-83. doi:10.1016/j.ajog.2017.02.013
36 Cavallaro A, Ash ST, Napolitano R, et al. Quality control of ultrasound for fetal biometry: results from the INTERGROWTH-21st Project. *Ultrasound Obstet Gynecol* 2018;52:332-9. doi:10.1002/uog.18811
37 Ohuma EO, Njim T, Sharps MC. Current issues in the development of foetal growth references and standards. *Curr Epidemiol Rep* 2018;5:388-98. doi:10.1007/s40471-018-0168-6
38 Francis A, Hugh O, Gardosi J. Customized vs INTERGROWTH-21st standards for the assessment of birthweight and stillbirth risk at term. *Am J Obstet Gynecol* 2018;218(2S):S692-9. doi:10.1016/j.ajog.2017.12.013
39 Åhman A, Runestam K, Sarkadi A. Did I really want to know this? Pregnant women's reaction to detection of a soft marker during ultrasound screening. *Patient Educ Couns* 2010;81:87-93. doi:10.1016/j.pec.2009.12.011
40 Viaux-Savelon S, Dommergues M, Rosenblum O, et al. Prenatal ultrasound screening: false positive soft markers may alter maternal representations and mother-infant interaction [correction in: *PLoS One* 2014;9:e91494]. *PLoS One* 2012;7:e30935. doi:10.1371/journal.pone.0030935
41 Chamberlain G, Zander L. ABC of labour care: induction. *BMJ* 1999;318:995-8. doi:10.1136/bmj.318.7189.995
42 Aaboud M, Aad G, Abbott B, et al, Atlas Collaboration. Performance of the ATLAS trigger system in 2015. *Eur Phys J C Part Fields* 2017;77:317. doi:10.1140/epjc/s10052-017-4852-3
43 Gaccioli F, Sovio U, Cook E, Hund M, Charnock-Jones DS, Smith GCS. Screening for fetal growth restriction using ultrasound and the sFLT1/PlGF ratio in nulliparous women: a prospective cohort study. *Lancet Child Adolesc Health* 2018;2:569-81. doi:10.1016/S2352-4642(18)30129-9
44 Mangesi L, Hofmeyr GJ, Smith V, Smyth RM. Fetal movement counting for assessment of fetal wellbeing. *Cochrane Database Syst Rev* 2015;(10):CD004909. doi:10.1002/14651858.CD004909.pub3
45 Ottaviani G, Buja LM. Anatomopathological changes of the cardiac conduction system in sudden cardiac death, particularly in infants: advances over the last 25 years. *Cardiovasc Pathol* 2016;25:489-99. doi:10.1016/j.carpath.2016.08.005
46 Lavezzi AM, Johanson CE. Editorial: New approaches to the pathogenesis of sudden intrauterine unexplained death and sudden infant death syndrome. *Front Neurol* 2017;8:441. doi:10.3389/fneur.2017.00441

Web appendix 1: study protocol