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,7 on behalf of the IRIS study group

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 or less, impaired consciousness, asphyxia, seizures, assisted ventilation, sepsis, meningitis, bronchopulmonary dysplasia, intraventricular haemorrhage, periventricular leukomalacia, 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%) vs 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.
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.\(^12\)\(^13\)

Which third trimester screening strategy is most effective in detecting fetal growth restriction is controversial.\(^14\)\(^15\) 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 15 participants) 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...
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. 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. 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. 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).

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. 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, (or a birth weight between the 2.3rd and...
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), 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 χ² 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...
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 (SPSS 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 (86.4% 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 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 group 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.
60 Midwifery practices

6495 Women allocated to control strategy

6282 Women allocated to intervention strategy

347 Excluded

456 Excluded

1308 Women providing baseline characteristics

1031 Women providing baseline characteristics

320 Failed linkage to or missing Perined data

305 Failed linkage to both Perined and hospital data†

26 Missing outcome data

169 Failed linkage to both Perined and hospital data†

118 Severe adverse perinatal outcome

27 Missing outcome data

5979 Available data on severe adverse perinatal outcomes

6922 No severe adverse perinatal outcome

7067 Available data on severe adverse perinatal outcomes

7645 No severe adverse perinatal outcome

5472 Severe adverse perinatal outcome

26 Missing outcome data

Excluded

Not meeting inclusion criteria

Declined use of medical data

Included in one or more of the primary analyses

Included in one or more of the primary analyses

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

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (n=59)*</th>
<th>Start date of intervention strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 May 2015 (n=19)*</td>
</tr>
<tr>
<td>Mean (SD) No of midwives per practice</td>
<td>5.0 (2.0)</td>
<td>5.2 (2.1)</td>
</tr>
<tr>
<td>Ultrasound biometry in midwifery practice (No (%))</td>
<td>19 (32)</td>
<td>9 (47)</td>
</tr>
<tr>
<td>Ultrasound biometry in sonography centre (No (%))</td>
<td>40 (68)</td>
<td>10 (53)</td>
</tr>
<tr>
<td>Mean (SD) No of women in 2013</td>
<td>422.2 (226.0)</td>
<td>400.3 (155.0)</td>
</tr>
<tr>
<td>Mean (SD) nulliparous women (%)</td>
<td>46.7 (9.0)</td>
<td>44.7 (11.5)</td>
</tr>
<tr>
<td>Mean (SD) antepartum referrals to hospital care in 2013 (%)</td>
<td>37.5 (11.5)</td>
<td>37.1 (12.1)</td>
</tr>
<tr>
<td>Median (interquartile range) birth in midwife led care in 2013</td>
<td>88.0 (57.0; 128.0)</td>
<td>83.0 (58.0; 118.0)</td>
</tr>
<tr>
<td>Mean (SD) home births among midwife led births in 2013 (%)</td>
<td>49.3 (20.2)</td>
<td>49.6 (17.6)</td>
</tr>
<tr>
<td>Customised fundal height chartst (No (%))</td>
<td>27 (46)</td>
<td>9 (47)</td>
</tr>
</tbody>
</table>

*One practice dropped out in April 2015.
†None reported use of non-customised charts for fundal height measurements.
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 2 | Personal and baseline characteristics of participants. Data are numbers (percentages) unless stated otherwise

<table>
<thead>
<tr>
<th>Characteristics*</th>
<th>Intervention strategy (n=7067)</th>
<th>Control strategy (n=5979)</th>
<th>Total (n=13046)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nulliparous</td>
<td>3368 (47.7)</td>
<td>2928 (49.0)</td>
<td>6296 (48.3)</td>
</tr>
<tr>
<td>Multiparous</td>
<td>3632 (51.4)</td>
<td>3004 (50.2)</td>
<td>6636 (50.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>67 (0.9)</td>
<td>47 (0.8)</td>
<td>114 (0.9)</td>
</tr>
<tr>
<td>Mean (SD) maternal age (years)</td>
<td>31.0 (4.5)</td>
<td>31.0 (4.3)</td>
<td>31.0 (4.4)</td>
</tr>
<tr>
<td>Body mass index:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5</td>
<td>232 (3.3)</td>
<td>185 (3.1)</td>
<td>417 (3.2)</td>
</tr>
<tr>
<td>18.5-25.0</td>
<td>4583 (64.9)</td>
<td>4025 (67.3)</td>
<td>8608 (66.0)</td>
</tr>
<tr>
<td>&gt;25.0</td>
<td>2149 (30.4)</td>
<td>1707 (28.5)</td>
<td>3856 (29.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>103 (1.5)</td>
<td>62 (1.0)</td>
<td>165 (1.3)</td>
</tr>
<tr>
<td>Ethnicity:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>5096 (72.1)</td>
<td>4684 (78.4)</td>
<td>9780 (75.0)</td>
</tr>
<tr>
<td>Other Western</td>
<td>766 (10.8)</td>
<td>576 (9.6)</td>
<td>1342 (10.3)</td>
</tr>
<tr>
<td>Non-Western</td>
<td>1202 (17.0)</td>
<td>714 (11.9)</td>
<td>1916 (14.7)</td>
</tr>
<tr>
<td>Missing</td>
<td>3 (0.0)</td>
<td>3 (0.0)</td>
<td>6 (0.1)</td>
</tr>
<tr>
<td>Education:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>3770 (53.3)</td>
<td>3190 (53.4)</td>
<td>6960 (53.3)</td>
</tr>
<tr>
<td>Medium</td>
<td>2650 (34.7)</td>
<td>2115 (35.4)</td>
<td>4765 (35.0)</td>
</tr>
<tr>
<td>Low</td>
<td>700 (9.9)</td>
<td>588 (9.8)</td>
<td>1288 (9.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>167 (2.1)</td>
<td>86 (1.4)</td>
<td>253 (1.8)</td>
</tr>
<tr>
<td>Work status:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>5763 (81.5)</td>
<td>5043 (84.3)</td>
<td>10806 (82.8)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1188 (16.8)</td>
<td>863 (14.4)</td>
<td>2051 (15.7)</td>
</tr>
<tr>
<td>Missing</td>
<td>116 (1.6)</td>
<td>73 (1.2)</td>
<td>189 (1.4)</td>
</tr>
<tr>
<td>Marital status:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or cohabiting</td>
<td>6457 (91.4)</td>
<td>5604 (93.7)</td>
<td>12061 (92.4)</td>
</tr>
<tr>
<td>Not cohabiting</td>
<td>352 (5.0)</td>
<td>198 (3.3)</td>
<td>550 (4.2)</td>
</tr>
<tr>
<td>Single</td>
<td>167 (2.4)</td>
<td>121 (2.0)</td>
<td>288 (2.2)</td>
</tr>
<tr>
<td>Missing</td>
<td>91 (1.3)</td>
<td>56 (0.9)</td>
<td>147 (1.1)</td>
</tr>
<tr>
<td>Maternal smoking, alcohol use, or recreational drug use:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2283 (32.3)</td>
<td>1956 (32.7)</td>
<td>4239 (32.5)</td>
</tr>
<tr>
<td>No</td>
<td>4759 (67.3)</td>
<td>4004 (67.0)</td>
<td>8763 (67.2)</td>
</tr>
<tr>
<td>Smoking during pregnancy:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1003 (14.2)</td>
<td>835 (14.0)</td>
<td>1838 (14.1)</td>
</tr>
<tr>
<td>No</td>
<td>6048 (85.6)</td>
<td>5129 (85.8)</td>
<td>11177 (85.7)</td>
</tr>
<tr>
<td>Missing</td>
<td>16 (0.2)</td>
<td>15 (0.2)</td>
<td>31 (0.2)</td>
</tr>
<tr>
<td>Alcohol consumption during pregnancy:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1679 (23.8)</td>
<td>1481 (24.8)</td>
<td>3160 (24.2)</td>
</tr>
<tr>
<td>No</td>
<td>5362 (75.9)</td>
<td>4482 (75.0)</td>
<td>9844 (75.5)</td>
</tr>
<tr>
<td>Missing</td>
<td>26 (0.4)</td>
<td>16 (0.3)</td>
<td>42 (0.3)</td>
</tr>
<tr>
<td>Recreational drug use during pregnancy:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>86 (1.2)</td>
<td>70 (1.2)</td>
<td>156 (1.2)</td>
</tr>
<tr>
<td>No</td>
<td>6963 (98.5)</td>
<td>5896 (98.6)</td>
<td>12859 (98.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>18 (0.3)</td>
<td>13 (0.2)</td>
<td>31 (0.2)</td>
</tr>
<tr>
<td>Medical indication for fetal ultrasound determined at inclusion:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1324 (18.7)</td>
<td>1121 (18.7)</td>
<td>2445 (18.7)</td>
</tr>
<tr>
<td>No</td>
<td>5743 (81.3)</td>
<td>4858 (81.3)</td>
<td>10601 (81.3)</td>
</tr>
<tr>
<td>Type of medical indications determined at inclusion:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundal height measurements unreliable*</td>
<td>137 (1.9)</td>
<td>119 (2.0)</td>
<td>256 (2.0)</td>
</tr>
<tr>
<td>Previous SGA neonate*</td>
<td>176 (2.5)</td>
<td>161 (2.7)</td>
<td>337 (2.6)</td>
</tr>
<tr>
<td>Other (not biometry related) indication†</td>
<td>1012 (14.3)</td>
<td>844 (14.1)</td>
<td>1856 (14.2)</td>
</tr>
<tr>
<td>Fetal sex:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>3480 (49.2)</td>
<td>2923 (48.7)</td>
<td>6403 (48.9)</td>
</tr>
<tr>
<td>Male</td>
<td>3585 (50.7)</td>
<td>3055 (50.9)</td>
<td>6640 (50.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>2 (0)</td>
<td>1 (0)</td>
<td>3 (0)</td>
</tr>
</tbody>
</table>

SGA=small for gestational age (birth weight <10th centile according to Dutch birth weight curve). *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.
Comparison with other studies

Our trial addressed important shortcomings of previous studies. 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. 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. 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, the specificity in detecting 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, the specificity in detecting SGA was significantly higher in the control strategy (all \(P < 0.001\)).

1. McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks’ gestation and at 34-36 weeks’ gestation (P=0.38).

2. 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).

3. Numbers for routine biometry scans at 28-30 weeks’ gestation and at 34-36 weeks’ gestation differ slightly because of missing values.

4. Sensitivities of AC presented as percentages in the table differ slightly because of missing values.

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

<table>
<thead>
<tr>
<th>Ultrasound variables</th>
<th>Intervention strategy (n=6909)</th>
<th>Control strategy (n=5498)</th>
<th>Routine fetal biometry at 28-30 weeks’ gestation (n=6909)*</th>
<th>Routine fetal biometry at 34-36 weeks’ gestation (n=6888)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SGA</strong> (n=556)</td>
<td>AC&lt;P10 or slow growth</td>
<td>179 (65 to 78)</td>
<td>179 (65 to 78)</td>
<td>179 (65 to 78)</td>
</tr>
<tr>
<td><strong>No SGA</strong> (n=6353)</td>
<td>AC&lt;P10 or slow growth</td>
<td>177 (61 to 78)</td>
<td>177 (61 to 78)</td>
<td>177 (61 to 78)</td>
</tr>
<tr>
<td><strong>SGA</strong> (n=407)</td>
<td>AC&lt;P10 or slow growth</td>
<td>122 (156 to 53)</td>
<td>122 (156 to 53)</td>
<td>122 (156 to 53)</td>
</tr>
<tr>
<td><strong>No SGA</strong> (n=5901)</td>
<td>AC&lt;P10 or slow growth</td>
<td>122 (156 to 53)</td>
<td>122 (156 to 53)</td>
<td>122 (156 to 53)</td>
</tr>
<tr>
<td><strong>SGA</strong> (n=556)</td>
<td>Slow growth</td>
<td>76 (22 to 34)</td>
<td>76 (22 to 34)</td>
<td>76 (22 to 34)</td>
</tr>
<tr>
<td><strong>No SGA</strong> (n=6353)</td>
<td>Slow growth</td>
<td>76 (22 to 34)</td>
<td>76 (22 to 34)</td>
<td>76 (22 to 34)</td>
</tr>
</tbody>
</table>

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. 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.

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. 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. 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, the specificity in detecting SGA was significantly higher in the control strategy (all \(P < 0.001\)).

1. McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks’ gestation and at 34-36 weeks’ gestation (P=0.38).

2. 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).

3. Numbers for routine biometry scans at 28-30 weeks’ gestation and at 34-36 weeks’ gestation differ slightly because of missing values.

4. Sensitivities of AC presented as percentages in the table differ slightly because of missing values.

5. \(\chi^2\) tests showed that sensitivities of AC 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 \(P < 0.001\)).

6. McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks’ gestation and at 34-36 weeks’ gestation (P=0.38).

7. 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).

8. Numbers for routine biometry scans at 28-30 weeks’ gestation and at 34-36 weeks’ gestation differ slightly because of missing values.

9. Sensitivities of AC 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 \(P < 0.001\)).

10. McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks’ gestation and at 34-36 weeks’ gestation (P=0.38).

11. 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).

12. McNemar test revealed no differences in sensitivity between the fetal biometry scan at 28-30 weeks’ gestation and at 34-36 weeks’ gestation (P=0.38).

13. 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).
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. 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 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. 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 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. 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 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. 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 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. 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 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. 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 was higher for the second...
showed an increase in identifying the risk of stillbirth when customised fetal growth charts were used.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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 there was a suspicion of SGA was associated with a higher incidence of initiated delivery by the provider.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{43} Moreover, women are aware of fetal movements, which are a sign of fetal wellbeing. A change in fetal activity

---

Table 5 | Maternal outcomes and peripartum interventions

<table>
<thead>
<tr>
<th>Maternal outcomes and peripartum interventions</th>
<th>Total No*</th>
<th>Intervention strategy</th>
<th>Control strategy</th>
<th>Odds ratio (95% CI) or Adjusted odds ratio (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite of spontaneous vaginal birth without intervention</td>
<td>12490</td>
<td>6663, 2974 (44.6)</td>
<td>5827, 2650 (45.5)</td>
<td>0.97 (0.90 to 1.04)</td>
</tr>
<tr>
<td>Induction of labour‡</td>
<td>12984</td>
<td>7034, 1118 (15.9)</td>
<td>5950, 813 (13.7)</td>
<td>1.19 (1.08 to 1.32)</td>
</tr>
<tr>
<td>Vacuum or forceps assisted birth</td>
<td>13044</td>
<td>7065, 538 (7.6)</td>
<td>5979, 506 (8.5)</td>
<td>0.89 (0.79 to 1.01)</td>
</tr>
<tr>
<td>Caesarean section:</td>
<td>13044</td>
<td>7065</td>
<td>5979</td>
<td>1.01 (0.91 to 1.12)</td>
</tr>
<tr>
<td>Primary</td>
<td>414 (5.9)</td>
<td>342 (5.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>555 (7.9)</td>
<td>472 (7.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>969 (13.7)</td>
<td>814 (13.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Augmentation of labour</td>
<td>12713</td>
<td>6894, 1902 (27.6)</td>
<td>5819, 1839 (31.6)</td>
<td>0.82 (0.76 to 0.89)</td>
</tr>
<tr>
<td>Pharmacological pain relief</td>
<td>12337</td>
<td>6582, 2786 (42.3)</td>
<td>5755, 2378 (41.3)</td>
<td>1.04 (0.97 to 1.12)</td>
</tr>
</tbody>
</table>

*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) 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.
¶Maternal peripartum morbidity and mortality

---

RESEARCH
might be a sign of fetal growth restriction.\textsuperscript{45} Fetal
death, often associated with fetal growth restriction, is
usually preceded by reduced fetal movements.\textsuperscript{46}
More research is needed, however, to determine what
type of information women should receive about the
wellbeing of the fetus.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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
0.5% 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.\textsuperscript{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.\textsuperscript{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
Gynaecologists),\textsuperscript{22} 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 Boechorststraat 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

7 Department of Obstetrics and Gynaecology, University Medical
Centre, Nijmegen, Netherlands; Chantal W P M
Hukkelhoven, Wageningen University and Research, Wageningen,
Netherlands; William A Groberman, 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; Marloes de Koon, University of Groningen, Groningen, Netherlands; Maya Kruit, Dutch Society of Obstetrics and Gynaecology NVOG, Utrecht, Netherlands; Anneke Kwee, University Medical Centre Utrecht, Utrecht, Netherlands; Sabina Ledda, Midwifery practice het Paleto/BN, Rotterdam, Netherlands; Harry N Lafeber, Amsterdam University Medical Centre, Leiden, Netherlands; Jan 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; Marion Nieuwenhuijse, Academie Verloskunde Maastricht, Maastricht, Netherlands; Guid Oei, Maxima Medical Centre, Eindhoven, Netherlands; Cees Oudejans, Amsterdam University Medical Centre, Amsterdam, Netherlands; K Marieke Paarlbregt, Gelse Hospitals, location Apeldoorn, Netherlands; Aris T Papageorghiou, University of Oxford, Oxford, UK; Uma M. Reddy, Yale University, New Haven, CT, USA; Paul De Reu, Prenatala Screeningscentrum “de Meier”, Eindhoven, Netherlands; Mariëls Rijnders, TNO, Leiden, Netherlands; Alieke de Roon-Imeerzeel, Royal Dutch Organization of Midwives, Netherlands; Connie Scheelee, University Medical Centre Utrecht/BN, 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; Piem 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 Westermann, Amsterdam University Medical Centre, Amsterdam, Netherlands; Kristel C Zeeuw, University of Amsterdam, Amsterdam, Netherlands; and Jun Jim Zhang, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Contributors: JH and VW contributed equally to the manuscript and are first authors. AD, HFvdH, PI, VV, and EP conceived and designed the study. ADJ, PI, PI, U, VW, and JH were involved in the development, implementation, or data collection of the study. JH, JW, and VW conducted data analyses. JH, VV, and ADJ drafted the manuscript and interpreted the results. All authors critically reviewed 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 authorizes 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), AD 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 Centre Amsterdam (reference No 2013.409).

Data sharing: No additional data available.

The manuscript’s guarantor (AD) 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.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.


Web appendix 1: study protocol