Perinatal death in a term fetal growth restriction randomized controlled trial

DIGITAT study group; van Wyk, Linda; Boers, Kim E; Gordijn, Sanne J; Ganzevoort, Wessel; Bremer, Henk A; Kwee, Anneke; Delemarre, Friso M C; van Pampus, Maria G; Bloemenkamp, Kitty W M

[Link to publication in University of Groningen/UMCG research database](#)

Original Research

Perinatal death in a term fetal growth restriction randomized controlled trial: the paradox of prior risk and consent

Check for updates

Linda van Wyk, MD, PhD; Kim E. Boers, MD, PhD; Sanne J. Gordijn, MD, PhD; Wessel Ganzevoort, MD, PhD; Henk A. Bremer, MD, PhD; Anneke Kwee, MD, PhD; Friso M. C. Delemarre, MD, PhD; Maria G. van Pampus, MD, PhD; Kitty W. M. Bloemenkamp, MD, PhD; Frans J. M. E. Roumen, MD, PhD; Jan M. M. van Lith, MD, PhD; Ben W. J. Mol, MD, PhD; Jim G. Thornton, MD, PhD; Sicco A. Scherjon, MD, PhD; Saskia le Cessie, PhD; On behalf of the DIGITAT Study Group

BACKGROUND: The disproportionate intrauterine growth intervention trial at term was an intention to treat analysis and compared labor induction with expectant monitoring in pregnancies complicated by fetal growth restriction at term and showed equivalence for neonatal outcomes.

OBJECTIVE: To evaluate trial participation bias and to examine the generalizability of the results of an obstetrical randomized trial.

STUDY DESIGN: We used data from participants and nonparticipants of a randomized controlled trial—the disproportionate intrauterine growth intervention trial at term (n=1116) —to perform a secondary analysis. This study compared induction of labor and expectant management in women with term growth restriction. Data were collected in the same manner for both groups. Baseline characteristics and neonatal and maternal outcomes were compared. The primary outcome was a composite measure of adverse neonatal outcome. Secondary outcomes were delivery by cesarean delivery and instrumental vaginal delivery; length of stay in the neonatal intensive care, neonatal ward, and the maternal hospital; and maternal morbidity.

RESULTS: Nonparticipants were older, had a lower body mass index, had a higher level of education, smoked less, and preferred expectant management. The time between study inclusion and labor onset was shorter in participants than in nonparticipants. Notably, 4 perinatal deaths occurred among nonparticipants and none among participants. Among nonparticipants, there were more children born with a birthweight below the third centile. The nonparticipants who had expectant management were monitored less frequently than the participants in both the intervention and the expectant arm.

CONCLUSION: We found less favorable outcomes and more perinatal deaths in nonparticipants. Protocol-driven management, differences between participants and nonparticipants, or the fact that nonparticipants had a preference for expectant management might explain the findings.

Key words: fetal growth restriction, Hawthorne effect, intrauterine growth restriction, randomized controlled trial

Introduction

The factors that influence patient participation in trials are widely discussed.[1–4] Patients' choices to participate are, for example, influenced by individual preferences and socioeconomic background.[1,5] Participation in a clinical trial can lead to the so-called "Hawthorne effect."[6]

The original Hawthorne effect was first described between 1924 and 1933 where productivity of factory workers was increased while they—as part of a study to improve quality of the production process—were more strictly supervised.[7,8] It is unknown what the exact

Related editorial, 100243.

EDITOR'S CHOICE

Hawthorne effect is, in general, on the outcome of clinical studies. Some studies report better outcome (in both arms of a study) than expected, which is then explained by the Hawthorne effect.[9–11] Other studies report no such thing as the Hawthorne effect.[12–14] A factor that could influence a possible Hawthorne effect is that characteristics of people who consent to participate in clinical trials often differ from patients who decline participation. Recruitment to clinical trials is influenced by socioeconomic status, and less educated women are often less willing to participate.[15–19]

The Hawthorne effect can result in better outcomes, not only in the intervention group, but also in the control group of the trial. This positive effect suggests that participating in a clinical trial may have an effect on the behavior both of doctors and patients. Doctors may feel that they are being watched and therefore could act differently and

therefore adhere more strictly to protocols, perhaps leading to earlier or other interventions that could improve outcome. Patients may be more aware of risk factors because of the fact that they may be better informed regarding their condition owing to the study information provided before the trial. They may also feel that they are being watched. Subsequently, they may improve healthy behavior and the patient delay in complaints or problems may be less, leading to better outcomes.

In many pregnancy complications, it was unknown whether induction of labor with the disadvantage of less maturation of a fetus or expectant management with the disadvantage of prolonged pathologic condition is the better option. We considered a randomized trial that compared these 2 management options in term fetal growth restriction (FGR)—the disproportionate intrauterine growth intervention trial at term (DIGITAT). A unique feature of the DIGITAT was that

AJOG MFM at a Glance

Why was this study conducted?
This study aimed to investigate possible factors contributing to the generalizability of a randomized trial in term fetal growth restriction by comparing perinatal characteristics and outcome between randomized and nonrandomized patients.

Key findings
Although maternal characteristics in general were more favorable in the nonrandomized participants, their perinatal outcome was less optimal. All 4 perinatal deaths occurred in the nonrandomized group. It remains debatable whether these findings can be contributed to the characteristics of the participants or their doctors or are related to the so-called "Hawthorne effect."

What does this add to what is known?
To assess generalizability of findings from a randomized trial, data on the nonrandomized patients also have to be evaluated and discussed.

the data of the nonparticipants were also collected in exactly the same manner as of the participants.

The DIGITAT was an intention to treat analysis and compared labor induction with expectant monitoring in pregnancies complicated by FGR at term and showed equivalence for neonatal outcomes.[20,21]

This study aimed to evaluate trial participation bias and to analyze generalizability of the results by (1) assessing whether and how baseline characteristics of nonparticipants differed from participants and by (2) comparing study outcomes of the 2 groups. We also aimed to (3) explore a possible Hawthorne effect among trial participants.

Materials and Methods
Study design and patients
The design of the DIGITAT has been described elsewhere. The trial compared induction of labor with an expectant management in term pregnancies complicated by antenatal suspected FGR. Participants allocated to the induction of labor group were induced within 48 hours of randomization. Participants allocated to the expectant monitoring group were monitored until the onset of spontaneous labor with daily fetal movement counts and twice weekly heart rate tracings, ultrasound examination, maternal blood pressure measurement, assessment of proteinuria, laboratory tests of liver and kidney functions, and full blood count.[22]

Eligible patients who declined randomization, but agreed to follow-up of their medical data, were treated according to local protocols and patient preferences and at the discretion of the attending obstetrician. The nonparticipants of this study all officially declined participation after they received written and oral information about the study. As prespecified in the study protocol, the data of the nonparticipants who consented to data use were collected in the same way as of the participants, parallel to the trial. In the nonparticipants, women were considered to be monitored expectantly if labor was not induced and no primary cesarean delivery was performed within 48 hours after inclusion in the study. The primary outcome was a composite measure of adverse neonatal outcome. This was defined as death before hospital discharge, 5-minute Apgar score of <7, umbilical artery pH of <7.05, or admission to the neonatal intensive care unit. Secondary outcomes were delivery by cesarean delivery and instrumental vaginal delivery, length of stay in the neonatal intensive care or neonatal ward, length of stay in the maternal hospital, and maternal morbidity.

Statistical analysis
Continuous variables were summarized as means with standard deviations or medians with interquartile ranges. Continuous variables were compared using Student's t test or the nonparametric Mann-Whitney U test and presented as differences in means with 95% confidence intervals (CIs). The chi-square test and Fisher exact test were used to compare categorical variables and presented as risk differences with 95% CIs. If >5% of the observations were missing, this was indicated in the footnote of the table.

Propensity score methods were used to compare outcomes adjusting for group imbalances.[16] The propensity is the probability for women to participate in the randomized trial given their baseline characteristics (Table 1). This propensity score was calculated for all patients using logistic regression. Mean differences and risk differences were adjusted for the propensity score in linear regression models and additive risk models. Multiple imputation was used to handle missing data in the baseline variables with imputation models with baseline and outcome variables. Statistical analysis was performed using Statistical Package for the Social Sciences software (version 22.0; IBM Corp, Chicago, IL) and R (version 2.10.0; R Foundation for Statistical Computing, Vienna, Austria.), using the package Multivariate Imputation by Chained Equations.[17]

Results
For the DIGITAT, 1116 women were registered as eligible for participation—650 women were randomized, and 466 women declined randomization. Of the women who declined randomization, 452 consented to follow-up of their medical data and were included in the study for follow-up of their medical data. Of these nonparticipants, 58 were induced within 48 hours and 6 had a primary cesarean delivery within 48 hours. The other 377 women were monitored expectantly (Figure).

The baseline characteristics of the 2 groups are presented in Table 1. Nonparticipants were older, had a lower body mass index (BMI), had a higher level of education, and were less likely to smoke. Nonparticipants did not have more serious growth restriction than the participants, and no differences were seen in hypertensive pregnancy complications.

TABLE 1
Demographic and baseline characteristics of participants and nonparticipants

Characteristic	Participants (n=650)	Nonparticipants (n=452)	Difference in percentage or mean (95% CI)
Maternal age	27 (23—31)	31 (30—34)	−3.3 (−3.9 to 2.6)[a]
Body mass index at study entry[b]	23.3 (20—25)	21.2 (20—24)	1.0 (0.4—1.6)[a]
Gestational age at inclusion	263 (258—269)	262 (258—269)	0.9 (−1.0 to 2.7)
Nulliparous	383 (59)	275 (61)	−2.1 (−7.9 to 3.8)
Caucasian[c]	507 (82)	344 (83)	−1.0 (−5.7 to 3.7)
Education[d]			
Higher professional	63 (16)	93 (38)	−23 (−30 to 1.6)[a]
Maternal smoking[e]	265 (44)	114 (27)	17 (11—23)[a]
Blood pressure at booking, mm Hg[f]			
Systolic	115 (105—120)	115 (110—120)	−0.8 (−2.4 to 0.8)
Diastolic	69 (60—75)	70 (60—75)	−1.2 (−2.4 to 0.045)
Blood pressure at study entry, mm Hg			
Systolic	120 (110—130)	120 (110—130)	−0.8 (−2.5 to 1.0)
Diastolic	72 (65—80)	75 (70—85)	−1.4 (−2.8 to 0.0)
Women with gestational hypertension	28 (4.3)	25 (5.5)	−1.2 (−3.9 to 1.4)
Women with preeclampsia	45 (6.9)	27 (6.0)	0.9 (−2.1 to 4.0)
Fetal abdominal circumference at inclusion, mm	288 (278—297)	289 (278—299)	1 (−2.0 to 2.2)

Data are presented as median (interquartile range, 25th to 75th percentile) or number (percentage).

CI, confidence interval; *DIGITAT*, disproportionate intrauterine growth intervention trial at term.

[a] $P<.001$, $P<.05$; [b] n=570 for participants, n=364 for nonparticipants; [c] n=616 for participants, n=413 for nonparticipants; [d] n=401 for participants, n=242 for nonparticipants; [e] DIGITAT, n=605 for participants, n=114 for nonparticipants; [f] DIGITAT, n=619 for participants, n=411 for nonparticipants.

van Wyk et al. Outcomes of nonparticipants of an obstetrical randomized controlled trial. AJOG MFM 2020.

The nonparticipants and/or their doctors had a strong preference for expectant management (Figure). Only 64 women (15%) were induced or had a primary cesarean delivery within 48 hours. Subsequently, labor started spontaneously more often resulting in prolonged pregnancy (Table 2) in nonparticipants. More cesarean deliveries were performed in the nonparticipant group, and more nonparticipants developed pregnancy-induced hypertension or preeclampsia after study inclusion than participants.

To compare the differences in fetal monitoring strategies between participants and nonparticipants during the trial period, we compared the number of fetal heart rate (FHR) tracings and the number of antenatal visits. Nonparticipants have net more antenatal visits and FHR tracings than participants

owing to the prolongation of pregnancy by awaiting spontaneous delivery. When comparing the number of FHR tracings and antenatal visits between the nonparticipants and participants of the expectant monitoring groups, per day of expectant management, we found significantly more FHR tracing and a tendency to more antenatal visits in the participants.

Neonatal outcomes are presented in Table 3. More babies of nonparticipants were severely growth restricted (<p3). No perinatal deaths occurred among participants. In contrast, 4 perinatal deaths occurred among nonparticipants (Fisher exact test, $P=.03$). Notably, 2 intrauterine deaths occurred after following an expectant monitoring policy, 1 at 40+2 and the other 1 at 41+4 weeks' gestation, with time to delivery after study entry of 14 and 24 days,

respectively, and birthweight of 3026 and 2780 grams, respectively. The first patient had 8 FHR tracings and 5 antenatal visits while waiting for spontaneous delivery. Clinical and pathologic examination of the fetus and placenta revealed that fetal death was probably caused by a partial placental abruption. The second patient was monitored with 10 FHR tracings and 6 antenatal visits. Postmortem examination of the fetus and placenta showed that this stillbirth was associated with FGR. The third baby died with a birthweight of 2130 grams after induction and emergency cesarean delivery because of placental abruption at 37+2 weeks' gestation. The mother was included 1 day before delivery and declined randomization, but gave permission for follow-up of medical data. However, the suspicion of FGR was raised already at 35+6 weeks' gestation.

FIGURE
Flow diagram of the process of the study
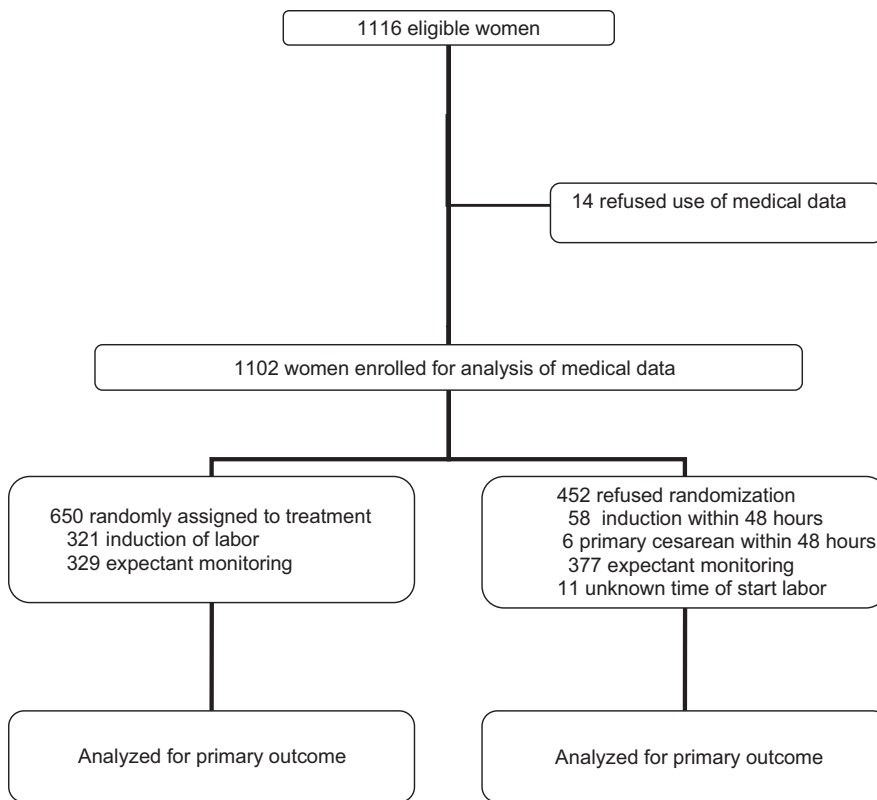
van Wyk et al. Outcomes of nonparticipants of an obstetrical randomized controlled trial. AJOG MFM 2020.

This child died after a long hospital admission (>150 days) owing to serious complications of severe asphyxia. The fourth child was delivered at 40+2 weeks' gestation with a birthweight of 2665 grams and had been discharged from the hospital 12 days after study entry in apparent good health but died unexpectedly from an intracranial hemorrhage owing to an arteriovenous malformation at the age of 28 days.

As reported in the original trial report, 1 participating woman who had induction of labor died unexpectedly at home 10 days after delivery. She had had an uncomplicated delivery at 38+4 weeks' gestation. No cause was found at the postmortem examination. No maternal deaths occurred in the nonparticipant group.

We observed no significant differences in composite adverse neonatal outcome between the participants and nonparticipants (6.5% vs 8.3%). Neonatal admissions and length of stay in the hospital were also similar between the groups (Table 3).

Discussion
Principal findings
Nonparticipants were healthier at baseline on important clinical characteristics (BMI, education, and less smoking). Most nonparticipant women with suspected FGR and/or their doctors preferred an expectant management policy. Notably, 4 perinatal deaths occurred among nonparticipants compared with 0 among participants. The percentage of children with a severe growth restriction (<p3) was higher among nonparticipants owing to prolongation of pregnancy.

Results
Women who are less educated and with a lower socioeconomic status are often less willing to participate in clinical trials.[15-19] However, in our study, we find the opposite; that is, highly educated women were less willing to be randomized for the trial. This could be explained by the fact that Dutch women are in general very reluctant when it comes to interventions during pregnancy. This phenomenon was also observed in our study where most nonparticipants preferred an expectant management.

Despite the fact that the participants had less favorable characteristics at baseline, their perinatal mortality rate was lower than the nonparticipating patients. This finding was also observed in low birthweight infants (defined as birthweight of <2500 grams), who—when the mothers smoked—had a lower mortality rate.[18] This so-called birthweight paradox can be explained by the fact that smoking causes FGR in otherwise healthy infants, whereas FGR in nonsmoking women is possible caused by other, possibly more severe, pathology. Another possible explanation could also be that the doctor is more aware of the risk factors in a pregnant woman who smokes and therefore is more alert.[23,24]

All 4 perinatal deaths in our study were in the nonparticipant group. Notably, 3 of these 4 deaths were in pregnancies with a gestational age of more than 281 days (40 weeks) and monitored expectantly. When looking at fetal monitoring strategies (FHR tracings and antenatal visits), these 4 pregnancies were monitored intensively, similar to the monitoring of the trial participants. However, we speculate that 2 of these deaths and perhaps the third death could have been prevented by induction of labor at an earlier gestational age.

Even though we found adequate fetal monitoring in the pregnancies where a perinatal death occurred, when comparing all the participants with the nonparticipants, participants had more FHR tracings, which can been seen as a proxy of standard of care, per day of expectant management and also more antenatal visits than nonparticipants. During the time of the study, no national protocols existed for monitoring strategies for pregnancies complicated by FGR, and policy was mainly based on the

TABLE 2
Pregnancy outcome, fetal monitoring, and onset of labor

	Participants (n=650)	Nonparticipants (n=452)	Differences in percentage or mean	Adjusted difference in percentage or mean (95% CI)[a]
Time between inclusion and onset of labor, d	2.4 (0.5–10.4)	8.0 (4.0–17.0)	−4.4 (−5.4 to −3.3)	−4.5 (−5.6 to 3.4)[b]
Induction	0.5 (0.4–1.4)	1.0 (1.0–1.0)	0.4 (−0.5 to 1.3)	0.6 (−0.6 to 1.8)
Expectant management	9.6 (5.3–15.5)	11.0 (6.0–18.0)	−1.2 (−2.4 to 0.1)	−0.6 (−1.9 to 0.8)
Gestational age at delivery, d	271 (263–279)	275 (268–282)	−3.4 (−4.6 to −2.2)[b]	−3.0 (−4.3 to −1.72)[b]
Induction	266 (261–271)	266 (261–272)	−0.1 (−2.2 to 2.1)	1.3 (−1.5 to 4.0)
Expectant management	277 (269–283)	277 (270–282)	0.1 (−1.2 to 1.4)	0.6 (−0.9 to 2.0)
Number of FHR tracings per d in women with expectant management	0.27 (0.0–0.52)	0.22 (0.00–0.38)	0.07 (−0.04 to 0.19)	0.08 (0.01– 0.14)[c]
Number of antenatal visits per d in women with expectant management	0.29 (0.10–0.48)	0.23 (0.11–0.33)	0.05 (−0.05 to 0.15)	0.06 (−0.00 to 0.12)
Onset of labor				
Spontaneously	163 (25)	197 (44)	−19 (−24 to −13)[b]	−19 (−24 to −13)[b]
Planned cesarean delivery	13 (2.0)	13 (2.9)	−0.9 (−2.8 to 1.0)	−1.8 (−3.5 to −0.1)[c]
Induction	472 (73)	241 (53)	19 (14–25)[b]	21 (15–27)[b]
Mode of delivery				
Spontaneously	506 (77.8)	328 (72.7)	5.1 (−0.1 to 10.3)	4.8 (−0.9 to 10.6)
Vaginal instrumental delivery	54 (8.3)	521 (11.5)	−3.2 (−6.9 to 0.4)	−2.0 (−6.0 to 1.9)
Cesarean delivery	90 (14)	71 (16)	−1.9 (−6.2 to 2.4)	−3.1 (−7.8 to 1.5)
Indications for cesarean delivery				
Suspected fetal distress (+/− arrest of labor)	77 (85.6)	59 (84.3)	1.3 (−10 to 13)	−0.4 (−13 to 13)
Arrest of labor	7 (7.8)	5 (7.1)	0.6 (−7.6 to 8.8)	0.9 (−8.4 to 10.2)
Other	6 (6.7)	6 (8.6)	−1.9 (−10 to 6.4)	−1.5 (−12 to 9.1)
Maternal death	1 (0.2)	0		N/A
Progression to gestational hypertension	7 (1.1)	13 (2.9)	−1.8 (−3.5 to −0.1)	−1.2 (−4.5 to 2.1)
Progression to preeclampsia	38 (5.8)	38 (8.4)	−2.6 (−5.7 to 0.6)	−2.1 (−3.8 to −0.3)[c]
Postpartum hemorrhage	25 (3.9)	23 (5.1)	−1.2 (−3.8 to 1.3)	−0.8 (−3.5 to 2.0)
Thromboembolic events	0 (0)	1 (0.2)		N/A
Placental abruption	1 (0.2)	2 (0.4)		N/A

Data are presented as median (interquartile range, 25th to 75th percentile) or number (percentage). For binary outcomes, CIs were not calculated when there were <5 events per group.

CI, confidence interval; *FHR*, fetal heart rate; *N/A*, not applicable.

[a] Difference corrected with propensity score (inverse probability weighting); [b] *P*<.001; [c] *P*<.05.

van Wyk et al. Outcomes of nonparticipants of an obstetrical randomized controlled trial. Am J Obstet Gynecol MFM 2020.

preference of the attending obstetrician and the patient and local protocols. Participants to the trial were required to adhere to a strict study protocol. For the nonparticipants, the management was dependent on the preference of the attending obstetrician and local protocols.

Although nonparticipants were healthier at baseline, we nevertheless found a higher incidence of a birthweight of <p3 in the nonparticipants. We have shown, in a follow-up study after 2 years, among trial participants that more severe FGR (<p3) is independently associated with a less optimal neurodevelopmental and behavioral outcome, independent from management—induction of labor or an expectant management. This outcome was not measured in the nonparticipant group, but the finding of more severe growth restriction among nonparticipants may therefore have long-term implications.[26]

TABLE 3
Neonatal outcomes

	Participants (n=650)	Nonparticipants (n=452)	Difference in percentage or mean (95% CI)[a]	Adjusted difference in percentage or mean (95% CI)[a]
Birthweight, g	2485 (2235–2750)	2528 (2270–2810)	−28 (−75 to 19)	−17 (−69 to 34)
Percentiles[b]				
<p2.3	131 (20)	128 (29)	−8.5 (−14 to −3.2)[c]	−8.7 (−14 to −3.0)[c]
p2.3 to p5	148 (23)	106 (24)	−0.9 (−6.0 to 4.2)	−1.1 (−6.7 to 4.5)
p5 to p10	155 (24)	106 (24)	0.2 (−5.0 to 5.3)	1.0 (−4.7 to 6.6)
p10 to p25	165 (25)	76 (17)	8.4 (3.6–13.3)[c]	8.2 (2.9–14)[c]
>p25	50 (7.7)	31 (6.9)	0.8 (−2.4 to 3.9)	1.2 (−2.2 to 4.7)
Composite adverse neonatal outcome	37 (6.5)	32 (8.3)	−1.8 (−5.2 to 1.6)	−2.4 (−6.1 to 1.3)
Fetal deaths or neonatal deaths[d]	0 (0)	4 (0.9)		N/A
Apgar score after 5 min of <7	9 (1.4)	10 (2.2)	−0.8 (−2.5 to 0.8)	−1.0 (−2.5 to 0.4)
Arterial pH of <7.05	14 (2.5)	9 (2.4)	0.1 (−1.9 to 2.1)	0.2 (−2.0 to 2.4)
Admission to intensive care	22 (3.4)	16 (3.5)	−0.2 (−2.4 to 2.0)	−1.1 (−3.0 to 0.9)
Neonatal admission				
High care or medium care	275 (43)	195 (43)	−0.7 (−6.7 to 5.2)	−2.7 (−9.2 to 3.8)
Length of stay, d	4.0 (2.0–8.0)	4.0 (2.0–7.0)	0.2 (−0.9 to 1.4)	0.2 (−1.1 to 1.5)

Data are presented as median (interquartile range, 25th to 75th percentile) or number (percentage). For binary outcomes, CIs were not calculated when there were <5 events per group.

Apgar, appearance, pulse, grimace, activity, and respiration; *CI*, confidence interval; *N/A*, not applicable.

[a] Difference corrected for propensity score; [b] According to Dutch fetal growth charts; [c] *P*<.001, *P*<.05; [d] Defined as death within 28 days after birth or before hospital discharge.

van Wyk et al. Outcomes of nonparticipants of an obstetrical randomized controlled trial. AJOG MFM 2020.

In our trial, we saw a more favorable outcome among participants (less severely growth restricted, no perinatal deaths) in both the intervention and expectantly monitored groups. Apart from slightly more FHR tracings in the participants, we found no clear evidence of a Hawthorne effect. We cannot dismiss the possibility that at least part of the cause of the adverse outcomes was that most nonparticipants seemed to favor an expectant management policy, subsequently leading to relatively more advanced gestational ages and more growth restriction.

Clinical implications
The primary DIGITAT[20] advises that it is rational to choose induction of labor to prevent stillbirths. The secondary analyses of the DIGITAT[25,26] conclude that induction of labor beyond 38 weeks' gestation seems to be the most favorable management policy. The differences we found between the participants and nonparticipants could be explained by the fact that most nonparticipants preferred an expectant management. This finding could be an extraargument for induction of labor before 40 weeks' gestation in term pregnancies complicated by FGR.

Research implications
FGR was defined using a population based neonatal growth curve and without using an international accepted definition. For future studies, using an international accepted definition of FGR should be used and also hemodynamic fetal parameters should be used for the antenatal identification of FGR based on placental insufficiency. Such a trial is now running in the Netherlands at the moment. In addition, a long-term follow-up study is important to draw definitive conclusions on perinatal management. Such a study—at the age of 11 years and older—is actually done at the moment both in randomized and nonrandomized patients.

Strengths and limitations
Data from randomized and nonrandomized patients were included prospectively in the same database. A weakness is that fetal Doppler studies were not performed. A larger sample size is needed to adequately adjust for possible confounding when comparing the number of deaths between the randomized and nonrandomized patients.

As in all randomized controlled trials, we expect that there is a group of patients who were eligible for randomization in the study, but they were not approached by the clinician for whatever reason. We unfortunately do not have information about this group of patients.

Conclusion
We found that the nonparticipants in the DIGITAT have a worse outcome than the participants, despite the fact that these women were healthier at baseline. This shows the extreme importance of collecting data of those declining

randomization in the same way as for patients participating in a trial. We have no strong evidence of the presence of the Hawthorne effect. The fact that most nonparticipants preferred expectant management and thereby prolonged the possible undernourished fetal environment could explain the less favorable outcomes in these women and remains an argument for timely induction of labor. ∎

References

1. Ford JG, Howerton MW, Lai GY, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. Cancer 2008;112:228–42.

2. Lakerveld J, Ijzelenberg W, van Tulder MW, et al. Motives for (not) participating in a lifestyle intervention trial. BMC Med Res Methodol 2008;8:17.

3. Vist GE, Hagen KB, Devereaux PJ, Bryant D, Kristoffersen DT, Oxman AD. Outcomes of patients who participate in randomised controlled trials compared to similar patients receiving similar interventions who do not participate. Cochrane Database Syst Rev 2007;2: MR000009.

4. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 2005;365:82–93.

5. Vermaire JH, van Loveren C, Poorterman JH, Hoogstraten J. Non-participation in a randomized controlled trial: the effect on clinical and non-clinical variables. Caries Res 2011;45: 269–74.

6. Wickström G, Bendix T. The "Hawthorne effect"–what did the original Hawthorne studies actually show? Scand J Work Environ Health 2000;26:363–7.

7. Noland EW. Hawthorne revisited. By Henry A. Landsberger. Ithaca, New York: The New York state school of Industrial and Labour Relations, 1958, 132 pp. Soc Forces 1959;37:361–4.

8. Parsons HM. What happened at Hawthorne?: new evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. Science 1974;183:922–32.

9. Feil PH, Grauer JS, Gadbury-Amyot CC, Kula K, McCunniff MD. Intentional use of the Hawthorne effect to improve oral hygiene compliance in orthodontic patients. J Dent Educ 2002;66:1129–35.

10. Evans R, Joseph-Williams N, Edwards A, et al. Supporting informed decision making for prostate specific antigen (PSA) testing on the web: an online randomized controlled trial. J Med Internet Res 2010;12:e27.

11. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne effect: a randomised, controlled trial. BMC Med Res Methodol 2007;7:30.

12. Ertem IO, Votto N, Leventhal JM. The timing and predictors of the early termination of breastfeeding. Pediatrics 2001;107:543–8.

13. Fernald DH, Coombs L, DeAlleaume L, West D, Parnes B. An assessment of the Hawthorne effect in practice-based research. J Am Board Fam Med 2012;25:83–6.

14. Fox NS, Brennan JS, Chasen ST. Clinical estimation of fetal weight and the Hawthorne effect. Eur J Obstet Gynecol Reprod Biol 2008;141:111–4.

15. Mills N, Metcalfe C, Ronsmans C, et al. A comparison of socio-demographic and psychological factors between patients consenting to randomisation and those selecting treatment (the ProtecT study). Contemp Clin Trials 2006;27:413–9.

16. Vind AB, Andersen HE, Pedersen KD, Jørgensen T, Schwarz P. Baseline and follow-up characteristics of participants and non-participants in a randomized clinical trial of multifactorial fall prevention in Denmark. J Am Geriatr Soc 2009;57:1844–9.

17. Petersen MK, Andersen KV, Andersen NT, Søballe K. "To whom do the results of this trial apply?" External validity of a randomized controlled trial involving 130 patients scheduled for primary total hip replacement. Acta Orthop 2007;78:12–8.

18. Sidani S, Miranda J, Epstein D, Fox M. Influence of treatment preferences on validity: a review. Can J Nurs Res 2009;41:52–67.

19. Travers J, Marsh S, Williams M, et al. External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? Thorax 2007;62:219–23.

20. Boers KE, Vijgen SM, Bijlenga D, et al. Induction versus expectant monitoring for intrauterine growth restriction at term: randomised equivalence trial (DIGITAT). BMJ 2010;341: c7087.

21. Bijlenga D, Boers KE, Birnie E, et al. Maternal health-related quality of life after induction of labor or expectant monitoring in pregnancy complicated by intrauterine growth retardation beyond 36 weeks. Qual Life Res 2011;20: 1427–36.

22. Boers KE, Bijlenga D, Mol BW, et al. Disproportionate intrauterine growth intervention trial at term: DIGITAT. BMC Preg Childbirth 2007;7:12.

23. Wilcox AJ. On the importance–and the unimportance–of birth weight. Int J Epidemiol 2001;30:1233–41.

24. Wilcox AJ. Commentary: on the paradoxes of birth weight. Int J Epidemiol 2003;32:632–3.

25. van Wyk L, Boers KE, van der Post JA, et al. Effects on (neuro)developmental and behavioral outcome at 2 years of age of induced labor compared with expectant management in intrauterine growth-restricted infants: long-term outcomes of the DIGITAT trial. Am J Obstet Gynecol 2012;206:406.e1–7.

26. Boers KE, van Wyk L, van der Post JA, et al. Neonatal morbidity after induction vs expectant monitoring in intrauterine growth restriction at term: a subanalysis of the DIGITAT RCT. Am J Obstet Gynecol 2012;206:344.e1–7.

Author and article information

From the Leiden University Medical Center, Leiden, the Netherlands (Drs van Wyk, van Lith, and le Cessie); Haaglanden Medical Center, the Hague, the Netherlands (Drs van Wyk and Boers); University Medical Center Groningen, Groningen, the Netherlands (Drs van Wyk, Scherjon, and Gordijn); Academic Medical Center, Amsterdam, the Netherlands (Dr Ganzevoort); Reinier de Graaf Gasthuis, Delft, the Netherlands (Dr Bremer); University Medical Center Utrecht, Utrecht, the Netherlands (Drs Kwee and Bloemenkamp); Elkerliek Hospital, Helmond, the Netherlands (Dr Delemarre); Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands (Dr van Pampus); Atrium Medical Center, Heerlen, the Netherlands (Dr Roumen); Monash University Health Services, Clayton, Australia (Dr Mol); and University of Nottingham, Nottingham, United Kingdom (Dr Thorton).

Corresponding author: Linda van Wyk, MD, PhD. vwyklinda@gmail.com