Potential Improvement of Pregnancy Outcome through Prenatal Small for Gestational Age Detection

8 authors, including:

Bart Jan Voskamp
Academisch Medisch Centrum Universiteit van Amsterdam
16 PUBLICATIONS   **106** CITATIONS

SEE PROFILE

Corine J M Verhoeven
Amsterdam University Medical Center
67 PUBLICATIONS   **569** CITATIONS

SEE PROFILE

Katrien Oude Rengerink
University Medical Center Utrecht
134 PUBLICATIONS   **1,202** CITATIONS

SEE PROFILE

Anita C.J. Ravelli
Amsterdam UMC
195 PUBLICATIONS   **9,171** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Pre-Eclampsia Intervention with Esomeprazole (PIE) trial: a double blind randomised control trial to treat early onset pre-eclampsia View project

Prognostic models in obstetrics View project

Potential Improvement of Pregnancy Outcome through Prenatal Detection

Bart Jan Voskamp, MD[1,*]   Daphne H. Beemsterboer, MD[1,*]   Corine J. M. Verhoeven, PhD[2,3]
Katrien Oude Rengerink, MSc[1]   Anita C. J. Ravelli, PhD[1]   Jannet J. H. Bakker, PhD[1]
Ben Willem J. Mol, MD, PhD[1]   Eva Pajkrt, MD, PhD[1]

[1] Department of Obstetrics and Gynecology, Academic Medical
Center, Amsterdam, The Netherlands
[2] Department of Obstetrics and Gynecology, Maxima Medical Center,
Veldhoven, The Netherlands
[3] Department of Midwifery Science, AVAG and the EMGO Institute for
Health and Care Research, VU University Medical Center, Amsterdam,
The Netherlands
[*] These authors contributed equally.

Address for correspondence Bart Jan Voskamp, MD, Department of
Obstetrics and Gynecology, Academic Medical Center, Room H4-232,
Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands
(e-mail: b.voskamp@amc.nl).

Abstract

Objective   To assess differences in mode of delivery and pregnancy outcome between prenatally detected and nonprenatally detected small for gestational age (SGA) neonates born at term.

Study Design   We performed a retrospective multicenter cohort study. All singleton infants, born SGA in cephalic position between $36^{0/7}$ and $41^{0/7}$ weeks gestation, were classified as either prenatally detected SGA or nonprenatally detected SGA. With propensity score matching we created groups with comparable baseline characteristics. We compared these groups for composite adverse perinatal outcome, labor induction, and cesarean section rates.

Results   We included 718 SGA infants, of whom 555 (77%) were not prenatally detected. Composite adverse neonatal outcome did not differ statistically significant between the matched prenatally detected and the nonprenatally detected group (5.5 vs. 7.4%, odds ratio [OR] 0.74, 95% confidence interval [CI]: 0.30–1.8). However, perinatal mortality only occurred in the nonprenatally detected group (1.8% [3/163] in the matched cohort, 1.3% [7/555] in the complete cohort). In the propensity matched prenatally detected SGA group both induction of labor (57 vs. 9%, OR 14.0, 95% CI: 7.4–26.2) and cesarean sections (20 vs. 8%, OR 2.9, 95% CI: 1.5–5.8) were more often performed compared with the nonprenatally detected SGA group.

Conclusion   Prenatal SGA detection at term allows timely induction of labor and cesarean sections thus potentially preventing stillbirth.

Keywords

► small for gestation
► prenatal detection
► neonatal death
► fetal death
► induction of labor
► propensity score

Small for gestational age (SGA) neonates are defined as neonates with a birth weight below the 10th percentile of the growth curve for a given gestational age (GA). In the Netherlands annually approximately 18,000 neonates are born SGA, the majority at term.[1,2] This is a heterogeneous group comprising neonates that are constitutionally small and neonates that have failed to achieve their growth potential because they are growth restricted based on uteroplacental insufficiency, congenital anomalies, or infection.[3–5] SGA is associated with an increased risk of adverse pregnancy

outcome and adverse events in the postpartum phase.[6–9] The more pronounced the SGA, the higher the risk of antepartum death.[10]

It is assumed that early detection of SGA could improve fetal outcome by close fetal monitoring and the possibility for timely induction of labor or instrumental delivery when fetal condition seems compromised.[11] At present, no effective intervention is available to improve the outcome of SGA infants at term.[6] Ohel et al and Verlijsdonk et al both assessed differences in management of labor and perinatal outcome between prenatally detected SGA and nonprenatally detected SGA at term.[12,13] While both studies showed more labor inductions and cesarean sections in the prenatally detected SGA group, pregnancy outcome differed between the two studies. Ohel et al showed a higher rate of adverse neonatal outcome in prenatally detected SGA,[12] whereas Verlijsdonk et al concluded that prenatal suspected SGA was associated with lower rates of adverse neonatal outcome compared with nonprenatally detected SGA (3.8 vs. 9.0%, $p = 0.056$).[13] Both studies were small and likely biased by confounding, as the results were not adjusted for severity of growth restriction. This resulted in comparison of more severely growth restricted—prenatally detected—infants to generally milder SGA infants that were detected after birth. The actual impact of the prenatal detection of SGA remains uncertain.

The aim of this study was to assess, in groups with a comparable possibility of prenatal SGA detection, whether prenatal SGA detection in term infants improves perinatal outcome and whether this detection influences the timing and mode of delivery.

Materials and Methods

Study Design
We conducted a retrospective cohort study of women with a singleton SGA child born at home or in the hospital between $36^{0/7}$ and $41^{0/7}$ weeks of gestation, between 1 April, 2005 and 31 December, 2008. We classified infants as being prenatally detected SGA and nonprenatally detected SGA. Classification of prenatal SGA detection was based on ultrasonographically measured abdominal circumference < p10, estimated fetal weight < p10, flattening of the growth curve in the third trimester (as judged by a clinician), or the presence of all three factors. Subsequently, we created comparable groups of prenatally detected SGA and nonprenatally detected SGA infants by propensity score matching and compared pregnancy outcome and mode of delivery between these two groups.

Inclusion and Exclusion Criteria
We included pregnant women ≥ 18 years with a singleton pregnancy who gave birth to SGA neonates at a GA between $36^{0/7}$ and $41^{0/7}$ weeks in the catchment area of one of the following two hospitals and seven midwifery practices: the Academic Medical Center in Amsterdam, the Maxima Medical Center in Veldhoven, or one of the seven independent midwifery practices referring to these two medical centers.

To warrant comparability of pregnancies with an prenatally detected SGA and nonprenatally detected SGA infants, we excluded women with a breech presentation at birth, women with a child with fetal structural or chromosomal anomalies, women with a previous cesarean section, and women with pregnancies with uncertainty about duration of pregnancy.

SGA was defined as a birth weight below the 10th percentile for GA.[14] The Dutch reference curves for birth weight by GA stratified for parity, sex, and ethnic background were used to calculate birth weight percentiles on a continuous scale for all infants.[14] Pregnancy dating was performed by last menstrual period, or ultrasound measurements before 20 weeks of gestation (crown-rump-length or head-circumference measurement).

Data Collection
We searched the perinatal databases from the two participating hospitals and seven midwifery practices, to identify pregnancies with an prenatally detected SGA infant. Prenatally detected SGA infants had previously been eligible for inclusion in the DIGITAT (the Disproportionate Intrauterine Growth Intervention Trial At Term) study, an randomized equivalence trail that was performed to compare the effect of induction of labor with a policy of expectant monitoring for intrauterine growth restriction near term.[6] We used the same GA criteria as in the DIGITAT study to avoid loss of cases through a cut off at term ($37^{0/7}$ weeks gestation) instead of $36^{0/7}$ weeks gestation. To ensure inclusion of all nonprenatally detected SGA infants in the study period, we used the Netherlands Perinatal Registry (PRN), to complement data that could not be retrieved from the medical files. The PRN is a national database that contains linked maternal and neonatal data entered by midwives, gynecologists, and pediatricians.[15] It contains information on 96% of all pregnancies, home and hospital births, and readmissions until 28 days after birth. It does not contain information on whether SGA is detected prenatally.[16] We collected information on maternal characteristics: body mass index (BMI), smoking, parity, gestational hypertension; delivery characteristics: start of labor, mode of delivery, GA at delivery; and neonatal characteristics: Apgar score, birth weight, sex, neonatal complications, intrauterine fetal death, and neonatal death.

Outcome Measures
Outcomes of this study were adverse perinatal outcomes, intrauterine fetal death, neonatal death, neonatal complications, and a composite of these adverse outcomes. We also assessed whether there were differences in induction of labor and instrumental delivery rates between both the groups.

Intrauterine fetal death was defined as spontaneous fetal demise between $36^{0/7}$ and $41^{0/7}$ weeks gestation and neonatal death was defined as a live birth resulting in infant death within 28 days of life. Neonatal complications were defined as 5 minute Apgar score < 7, asphyxia, infant respiratory distress syndrome, meconium aspiration, pneumothorax or pneumomediastinum, necrotizing enterocolitis, convulsions, sepsis, and meningitis. Instrumental delivery was divided

into primary cesarean section, cesarean section in labor, and instrumental vaginal delivery.

Analysis

We used propensity score matched-pairs analyses to determine the association between prenatal SGA detection and the primary and secondary outcomes, while balancing potentially important confounders between both the groups. The rationale and methods underlying the use of propensity scores for proposed causal exposure variables have been previously described.[17,18]

The propensity scores were generated by logistic regression, based on all covariates that were known to be associated with perinatal outcome and that existed before the start of labor. We considered the continuous covariates maternal age, maternal BMI, and birth weight percentile as well as the dichotomous covariates primiparity, maternal smoking, gestational hypertension, birth weight $<$ p2.3, and birth weight $<$ p5. Since propensity scores cannot be calculated if one of the variables is missing, single imputation was used to replace missing values.[19]

The standardized difference was used to assess the balance of the covariates, as unlike significance testing it is not dependent of the size of the sample.[17] A standardized difference greater than 10% points was used to indicate that the samples were meaningfully different.[20] After generation of propensity scores, pairs were matched on their propensity score, using one-to-one nearest neighbor matching without replacement.[21–25] We matched nonprenatally detected SGA infants to the smallest group (prenatally detected SGA), to ensure that as many matches as possible could be made.

To compare baseline and pregnancy characteristics of prenatally and not prenatally detected SGA pregnancies we used Student's t-tests, χ^2 tests, and Fisher exact tests. In the matched cohort, the standardized difference was used to assess the balance of the covariates of the propensity scores. We also assessed the baseline characteristics of both groups to ensure that matching increased comparability.

To compare outcomes between prenatally detected and not detected SGA infants, odds ratios (ORs) with 95% confidence intervals (CIs) were computed for all dichotomous outcomes using logistic regression. Mean differences and 95% CIs were calculated for continuous variables with the independent t-test for normally distributed data. The Mann-Whitney U test was used to assess differences in continuous variables that were not normally distributed.

In a sensitivity analysis, we performed the analyses on a propensity score matched cohort of the original (nonimputed) dataset. We also used multivariable logistic regression analysis in the original dataset to determine the adjusted association of prenatally detected SGA with adverse outcome and mode of delivery in the entire sample.

Statistical analyses were conducted with SPSS version 19.0 for Windows (IBM Corporation, Armonk, NY). Propensity score calculation and matching were performed in R with the SPSS R-plugin.[25] The following R packages were invoked: MatchIt,[23,24] RItools,[22] and CEM.[26] A two-tailed p-value $<$ 0.05 was considered statistically significant.

Results

In the study period, 11,142 women delivered in one of the selected centers. ►Fig. 1 displays all our exclusions to arrive at a final cohort of 718 SGA infants. ►Table 1 shows the baseline pregnancy characteristics for prenatally detected SGA and nonprenatally detected SGA pregnancies in the unmatched cohort. The majority of SGA infants, 77% (555/718) remained undetected until after birth.
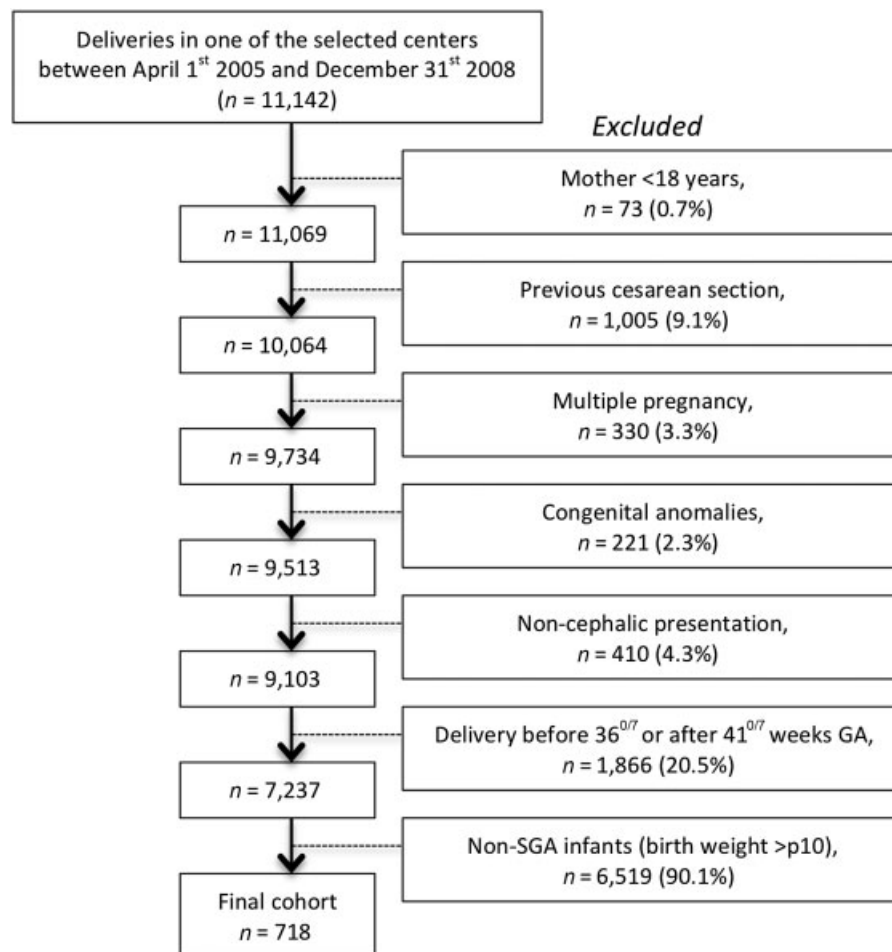
Characteristics of the prenatally detected SGA group and the nonprenatally detected SGA group are presented in ►Table 1. In the prenatally detected SGA group 51% of infants were $<$ p2.3 versus 21% in the nonprenatally detected SGA group. Smoking and primiparity were more prevalent in the prenatally detected SGA group (OR 2.8, 95% CI 1.9–4.2 and OR 1.6, 95% CI 1.2–2.4, respectively).

We know of 234 women in the cohort that they were referred for ultrasound growth assessment in the third trimester of pregnancy. A total of 19% (45/234) of these women were reassured about fetal growth but gave birth to a SGA infant. Because SGA was no longer suspected after the ultrasound, infants delivered by these women are classified as nonprenatally detected SGA infants.

Five of the six predefined baseline variables used for propensity score matching contained no missing data because these variables were required fields that caregivers are used to register. The sixth variable BMI lacked in 44% (316/718) of the women. Distribution plots of propensity scores in the two groups are shown in ►Fig. 2. Overall, as a function of baseline characteristics, the prenatally detected SGA group had a higher probability of prenatal SGA detection, as indicated by a higher mean propensity score (0.315 ± 0.156 vs. 0.201 ± 0.127; $p <$ 0.001). The initial difference in the two groups was further supported by the standardized difference criterion, which revealed that six of the eight baseline covariates (75%) had a standardized difference of $>$ 10% and therefore were imbalanced by this criterion. The identified differences and the inherent selection bias they represent, supported the need for further adjustment with propensity matching. This matching process resulted in the creation of 163 matched prenatally detected SGA and nonprenatally detected SGA pairs.

►Fig. 2 displays the distributions of the two matched groups' propensity scores. In contrast to the distributions of the unmatched groups, it reveals a high degree of overlap and similarity of shape between the two groups. This improved covariate balance was also reflected as the reduced difference in the means of the propensity scores reduced from 0.114 before matching, to 0.004 after matching (0.315 ± 0.156 in the prenatally detected SGA group and 0.311 ± 0.150 in the nonprenatally detected SGA group; $p =$ 0.80). The standardized difference criterion analysis confirmed the groups' similarity, as the highest standardized difference was 7.9%, where $<$ 10% is deemed acceptable (►Table 2).

The distribution of the outcomes in the matched pairs of prenatally detected SGA and nonprenatally detected SGA is presented in ►Table 3. Composite adverse neonatal outcome occurred in 5.5% (9/163) of infants in the prenatally detected SGA group and 7.4% (12/163) in the nonprenatally detected

```
┌─────────────────────────────────────────────┐
│  Deliveries in one of the selected centers    │
│ between April 1ˢᵗ 2005 and December 31ˢᵗ 2008 │
│               (n = 11,142)                     │
└─────────────────────────────────────────────┘
```

Fig. 1 Eligibility and exclusion of pregnancies.

SGA group (OR 0.74, 95% CI 0.30–1.8). Perinatal death occurred in none of the 163 prenatally detected SGA neonates and in 3 (1.8%) of the 163 nonprenatally detected SGA neonates (OR not calculable, $p = 0.996$). Birth weights of these three infants were below the first percentile. The cohort was too small to detect differences in subcategories of adverse neonatal outcome, but no obvious differences between the two groups were observed.

Table 1 Baseline characteristics of the total cohort in prenatally detected and nondetected SGA infants

Characteristic	SGA detected prenatally (n = 163)	SGA not detected prenatally (n = 555)	p-Value	Standardized difference (%)
Maternal characteristics[a]				
Age (y, mean ± SD)	29.3 ± 5.7	30.1 ± 5.8	0.11	−14.6
BMI (median, range)[a]	22.5 (16–44)	22.2 (16–43)	0.69	−2.6
Primiparity (%)	90 (55)	237 (43)	< 0.01	25.1
Smoking (%)	53 (33)	82 (15)	< 0.01	37.8
Gestational hypertension (%)	9 (5.5)	26 (4.7)	0.67	3.7
Fetal characteristics				
Birth weight percentile (median, range)	1.6 (0–9.9)	5.1 (0–10)	< 0.01	−63.9
Birth weight < p2.3 (%)	83 (51)	118 (21)	< 0.01	59.1
Birth weight < p5 (%)	124 (76)	270 (49)	< 0.01	64.1

Abbreviations: BMI, body mass index; SD, standard deviation; SGA, small for gestational age (birth weight < p10).
[a]The body mass index is the weight in kilograms divided by the square of the height in meters.
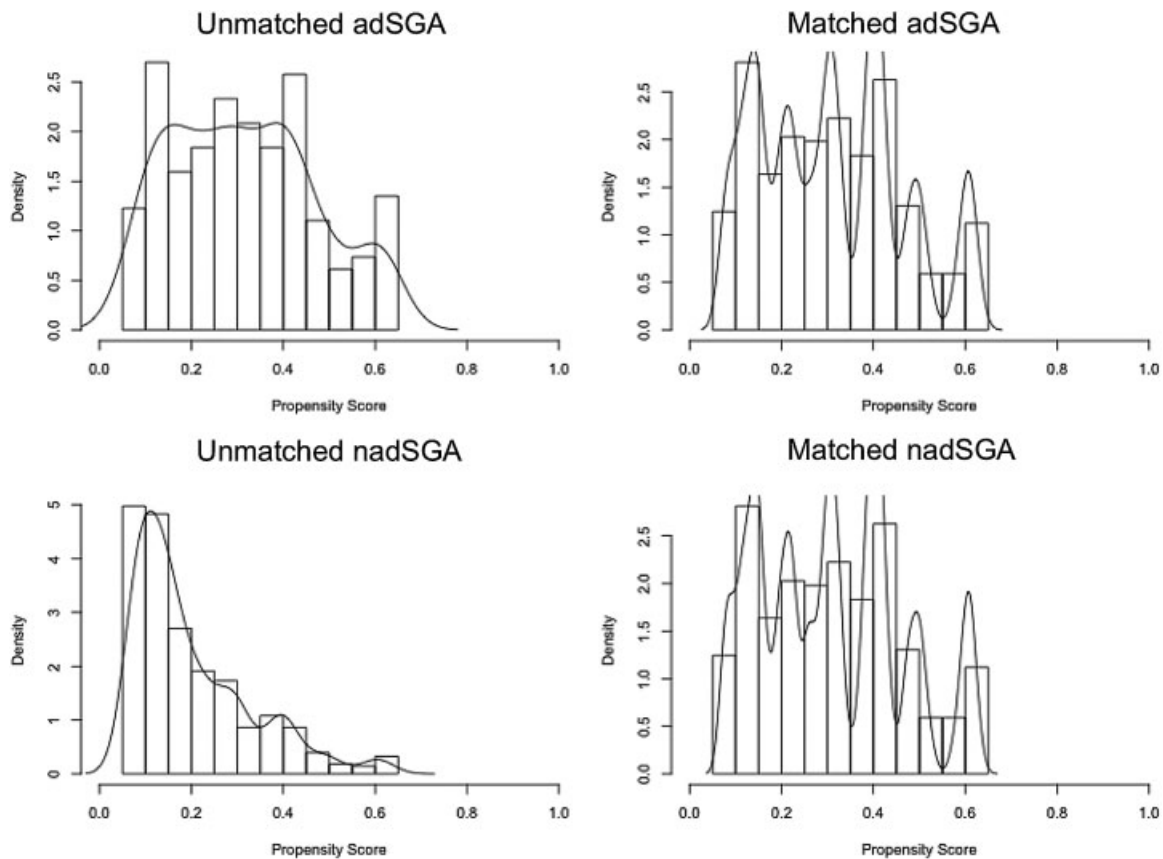
Fig. 2 Distribution of propensity scores before and after matching. Pregnancies were stratified by prenatal small for gestational age detection.

In the complete (unmatched) cohort ($n = 718$), perinatal mortality did not occur among 163 prenatally detected SGA infants and in 1.3% (7/555) of the nonprenatally detected SGA infants. These comprised six fetal deaths before the onset of labor (detected at $37^{4/7}$, $38^{1/7}$, $39^{1/7}$, $39^{2/7}$, $40^{2/7}$, and $40^{5/7}$ weeks GA), and one fetal death during labor ($40^{6/7}$ weeks GA).

To show a statistical significant difference (with α 0.05) in composite adverse neonatal outcome with 80% power, a sample size of 2,727 per group is needed. To show a statistical significant difference in perinatal death, a sample size of 422 per group is needed.

Labor was more often induced if SGA was detected prenatally (57 vs. 9% of women, OR 14, 95% CI 7.4–26). There were

Table 2 Baseline characteristics of matched cohort in prenatally detected and nondetected SGA infants

Characteristic	SGA detected prenatally ($n = 163$)	SGA not detected prenatally ($n = 163$)	*p*-Value	Standardized difference (%)
Maternal characteristics[a]				
Age (y, mean ± SD)	29.3 ± 5.7	29.3 ± 6.1	0.99	−0.1
BMI (median, range)[a]	22.9 (16–44)	22.8 (16–43)	0.45	7.9
Primiparity (%)	90 (55)	91 (56)	0.91	−1.2
Smoking (%)	53 (33)	52 (32)	0.91	1.3
Gestational hypertension (%)	9 (5.5)	9 (5.5)	1.00	0.0
Neonatal outcome				
Birth weight percentile (median, range)	1.6 (0–9.9)	1.9 (0–10)	0.68	−4.5
Birth weight < p2.3 (%)	83 (51)	83 (51)	1.00	0.0
Birth weight < p5 (%)	124 (76)	122 (75)	0.80	2.9

Abbreviations: BMI, body mass index; SD, standard deviation; SGA, small for gestational age (birth weight < p10).
[a]The body mass index is the weight in kilograms divided by the square of the height in meters.

Table 3 Propensity score matching (1:1) of the association between prenatal SGA (birth weight < p 10) detection and adverse pregnancy outcome and perinatal interventions

Outcome	SGA detected prenatally (n = 163)	SGA not detected prenatally (n = 163)	p-Value	OR	95% CI
Delivery, n (%)					
Induction of labor	93 (57)	14 (9)	< 0.001	13.95	7.43–26.19
Instrumental delivery	43 (26)	32 (20)	0.15	1.46	0.87–2.5
Cesarean section	33 (20)	13 (8)	0.02	2.93	1.48–5.80
Primary cesarean section	7 (4.3)	1 (0.6)	0.07	7.27	0.88–59.77
Cesarean section in labor	26 (16)	12 (7.4)	0.02	2.39	1.16–4.92
Failure to progress	0 (0)	1 (0.6)	0.996	not calculable	
Fetal distress	26 (16)	11 (6.7)	0.03	2.27	1.07–4.83
Vaginal instrumental delivery	10 (6.1)	19 (12)	0.09	0.50	0.22–1.10
Failure to progress	2 (1.2)	7 (4.3)	0.11	0.28	0.06–1.35
Fetal distress	8 (4.9)	12 (7.4)	0.36	0.65	0.26–1.63
Neonatal outcome, n (%)					
Mean gestational age at birth (d, mean ± SD)	270.6 ± 9.6	278.0 ± 6.6	< 0.001	−7.4[a]	−9.3 to −5.6
Birth weight (g; median, IQR)	2,410 (1,420–3,080)	2,640 (1,765–3,250)	< 0.001	−223[a]	−293 to −153
Composite adverse neonatal outcome	9 (5.5)	12 (7.4)	0.50	0.74	0.30–1.80
Perinatal mortality	0 (0.0)	3 (1.8)	0.996	not calculable	
5 min Apgar score < 7	4 (2.5)	4 (2.5)	1.00	1.00	0.25–4.07
Neonatal complications[b]	3 (1.8)	4 (2.5)	0.70	0.75	0.16–3.38
Asphyxia	1 (0.6)	0 (0.0)		n.c.	
IRDS	0 (0.0)	0 (0.0)		n.c.	
Meconium aspiration	0 (0.0)	2 (1.2)		n.c.	
Pneumothorax	1 (0.6)	3 (1.9)		n.c.	
NEC	1 (0.6)	0 (0.0)		n.c.	
Convulsions	1 (0.6)	0 (0.0)		n.c.	
Sepsis	0 (0.0)	0 (0.0)		n.c.	
Meningitis	0 (0.0)	0 (0.0)		n.c.	

Abbreviations: CI, confidence interval; IQR, interquartile ratio; IRDS, infant respiratory distress syndrome; n.c., not calculated; NEC, necrotizing enterocolitis; OR, odds ratio; SD, standard deviation; SGA, small for gestational age (birth weight < 10th percentile).
[a]Mean difference and 95% CI.
[b]Number of infants with neonatal complications, some infants have more than one complication.

more cesarean sections performed in the prenatally detected SGA group (20%) than in the nonprenatally detected SGA group (8%), OR 2.9, 95% CI 1.5–5.8, mostly all for suspected fetal distress. Failure to progress was never the indication for a cesarean section in labor in the prenatally detected SGA group and once (0.6%) in the nonprenatally detected SGA group.

The rate of vaginal instrumental delivery in the prenatally detected SGA group (6%) was lower than in the nonprenatally detected SGA group (12%) although this did not reach statistical significance (OR 0.50, 95% CI 0.22–1.1, p = 0.09). There were no significant differences in indication for vaginal instrumental delivery between the prenatally detected SGA and nonprenatally detected SGA group (►Table 3). On aver-

age, prenatally detected SGA neonates were born 7.4 days earlier than nonprenatally detected SGA neonates (38[5/7] vs. 39[5/7] weeks, 95% CI −9.3 to −5.6, p < 0.001), and prenatally detected SGA neonates weighed on average 223 g less than nonprenatally detected SGA neonates (median 2,410 vs. 2,640 g, 95% CI −293 to −153, p < 0.001).

Sensitivity Analyses
Both the percentages and p values of the multivariable logistic regression analysis on the complete cohort (n = 718) (**Appendix 1**), and the analyses on the nonimputed cohort after propensity score matching (**Appendices 2** and **3**), were comparable to the results of the propensity score analysis.

Discussion

This study confirms that in a system without routine third trimester growth screening ultrasounds, the large majority of women with a term-SGA pregnancy remain undetected until birth. However, severe SGA was more likely to be detected prenatally than mild SGA, although even in women with a child below the 2.3rd percentile the diagnosis fetal growth restriction was missed in 60%.[12,13] Obviously, prenatal SGA detection is associated with induction of labor and cesarean section.[12,13] Women with prenatally detected SGA gave birth more than a week earlier. Birth weight of prenatally detected SGA infants was more than 200 g lower. In the whole nonprenatally detected SGA group there were seven fetal deaths (of which three in the propensity score matched group), while none of the prenatally detected SGA infants died. The composite poor neonatal outcome occurred less often in the prenatally detected SGA group, although the difference was not statistically significant.

Strengths

Our study has several strengths. First, we assessed outcomes of prenatally detected SGA and nonprenatally detected SGA pregnancies balanced for propensity score, and therefore balanced for the covariates used to estimate the propensity score. These balanced covariates will no longer confound the relation between prenatal SGA detection and the outcome. Therefore, in contrast to two previous studies on the same subject where propensity score matching has not been performed to create comparable groups, the estimation will be theoretically unbiased, or at least less biasedness will have occurred.[12,13]

Second, we incorporated severity of SGA into the model as a continuous variable (birth weight percentile), instead of adjusting for birth weight and GA at delivery. Failure to do so in other studies might have biased the association between SGA detection and perinatal outcome.[12,13] The reliability of our results is further supported by the completeness and accuracy of prenatal and postnatal data of mother and child. Complete data were available for all pregnancies because data were extracted from the original patient files and complemented with use of the PRN registry if needed.

Limitations

The first limitation of this study is its sample size combined with the low incidence of adverse pregnancy outcome, specifically regarding perinatal mortality. Given the low incidence of adverse pregnancy outcome at term, a very large sample is required to show a difference. Although, this study does not have enough power to detect a statistically significant difference in rare adverse neonatal outcomes, to our knowledge this is the largest study that compared outcome of prenatally detected SGA with nonprenatally detected SGA infants. The precision of the results is quite limited due to the sample size, but the propensity score matching has resolved most of the bias that would be present in larger samples that are unmatched for relevant baseline variables.

Second, there is a possible a priori risk selection of prenatally detected SGA pregnancies. Women with an increased a priori risk of adverse pregnancy outcome receive regular ultrasound growth assessment. As a result, SGA infants in this high-risk population will likely be detected prenatally, whereas SGA is more likely to remain undetected until birth in low-risk pregnancies. We expect to have minimized this effect by the propensity score matching of the potential confounding maternal characteristics that were available and severity of SGA. However, we cannot fully exclude the possibility of residual confounding.

Third, in case of fetal death, there was no certainty about the moment of demise. This might have led to an overestimation of SGA severity in these infants. We expect this overestimation to be limited because—according to the Dutch protocol—all pregnant women undergo weekly checkups including Doppler auscultation of the fetal heart rate.

Fourth, unfortunately we did not know for all pregnancies if third trimester growth ultrasound had been performed. Therefore, we cannot report sensitivity and specificity of growth ultrasounds. The false reassurance about fetal growth in 19% of women that were referred for suspicion of SGA makes us suspect that the sensitivity of prenatal ultrasound especially in high-risk pregnancies can be improved. Due to propensity score matching we could not take the majority of nonprenatally detected SGA infants into account in the analyses. We first performed one-to-two matching to limit the data loss but the matching process did not yield comparable groups, mainly due to considerable difference in SGA severity between prenatally detected SGA and nonprenatally detected SGA infants. We have chosen one-to-one matching to warrant the optimal comparability of prenatally detected SGA and nonprenatally detected SGA infants, and also to obtain more reliable results. Additional sensitivity analyses on the entire imputed sample of 718 SGA neonates and on the original, nonimputed, cohort after propensity score matching showed results similar to the propensity-score analysis.

Considerations about Results

This study confirms the low prenatal detection rates of SGA.[12,13,27] The majority of pregnancies with an SGA infant remained undetected until birth, severe SGA is more likely to be detected prenatally. This is in concordance with literature which showed that the results of ultrasounds are unreliable to estimate the fetal weight $< $ p10 correctly.[28,29] Previous studies have also shown high-false positive rates (30%) of prenatal SGA detection.[6] Since we only assessed infants with a birth weight below the 10th percentile for GA, we could not rule on specificity of prenatal growth ultrasound.

Maternal smoking was more prevalent in the prenatally detected SGA group. This might be caused by awareness of caregivers for the potentially adverse effect of maternal smoking on fetal growth.[6,30–33] The statistically significant lower GA at birth and lower birth weight of the infants in the prenatally detected SGA group can be explained by the higher incidence of obstetrical interventions in this group.[12,13,28,33]

The study by Verlijsdonk et al concluded that suspicion of SGA was associated with a more active management of labor

and delivery, resulting in a better neonatal outcome at birth.[13] We observed a similar trend as Verlijsdonk et al that prenatally detected SGA fetuses have a better perinatal outcome. Combining the cohort of Verlijsdonk et al with our matched cohort results in 0.6% (2/321) perinatal deaths among prenatally detected SGA infants and 2.3% (10/435) perinatal deaths among nonprenatally detected SGA infants (OR 0.27, 95% CI 0.06–1.22, $p = 0.09$). Suggesting improved perinatal outcome of prenatally detected SGA infants compared with nonprenatally detected SGA infants.

Our study also confirms the more active management of labor among prenatally detected SGA infants. Increased induction of labor in the prenatally detected SGA group did not lead to higher rates of cesarean sections for failure to progress (stages I and II), but it led to more cesarean sections for suspected fetal distress, and less vaginal instrumental deliveries. The increased rate of cesarean sections and decreased rate of vaginal instrumental deliveries in prenatally detected SGA pregnancies might be caused by earlier intervention in case of suspected fetal distress—in view of the suspected SGA—or possible preference of the caregiver not to perform vaginal instrumental delivery if severe SGA is suspected. This assumption is supported by the trend toward more vaginal instrumental deliveries in the nonprenatally detected SGA group.

Choosing the 10th percentile as SGA cut off causes inclusion of a relatively large group of low-risk constitutionally small infants into the study population, by definition 10% of the population. Previous research has shown an association between the severity of SGA and perinatal outcome.[34] A study by Unterschneider et al showed that an estimated fetal weight < p3 is strongly and consistently associated with adverse perinatal outcome.[35] Our population consisted of a heterogeneous group of SGA infants. However, after propensity score matching, the median birth weight percentiles were 1.6 and 1.9 among prenatally detected SGA and nonprenatally detected SGA infants, indicating selection of mainly infants who are severely SGA.

Although, this study was underpowered to show a difference in the incidence of perinatal mortality between non-prenatally detected SGA and prenatally detected SGA infants, there were no perinatal deaths among prenatally detected SGA infants and seven among nonprenatally detected SGA infants in the complete cohort of 718 infants.

In the propensity-matched cohort, these numbers were zero and three, respectively. Six out of seven fetal deaths occurred before the onset of labor, versus one fetal death during labor. Considering the fact that death only occurred in SGA infants that were not detected prenatally, in which SGA was relatively milder than in the prenatally detected SGA group, it is not unlikely that death could have been avoided with fetal monitoring and induced labor if SGA had been detected before birth. However, we are not sure how prenatal SGA detection can be improved.

The low prenatal SGA detection rate in our study has several potential causes. First, the absence of third trimester ultrasound growth assessment as part of standard pregnancy care might play a role. Although, it seems logical that third trimester ultrasound as part of standard pregnancy care improves SGA detection rates, this has to our knowledge not been proven. Unfortunately our data do not allow quantification of how many women underwent third trimester ultrasound growth assessment.

Second, inaccuracy of ultrasound growth assessment in the third trimester might play a role. Prenatal ultrasound growth assessment is usually performed prior to $36^{0/7}$ weeks gestation because diagnostic accuracy decreases with advancing GA.[36]

A third possibility is that growth impairment starts after a reassuring third trimester growth ultrasound has been performed. We do not know if severe SGA always originates gradually and that poor detection is caused by inaccurate ultrasound measurements, or that growth of properly grown infants slows and comes to a halt after a—proper—ultrasound measurement in the third trimester.

Implications for Clinical Practice

This study shows that in the Dutch care system term SGA often remains undetected until birth and that prenatal SGA detection might prevent neonatal deaths. Caregivers and especially ultrasonographers should be aware of this to avoid as much as possible false reassurance of fetal growth.

If in any doubt about fetal growth, women should be followed up with umbilical artery measurements.[37] This allows for intervention if fetal condition is compromised and might prevent unnecessary interventions on constitutionally small infants that are not growth restricted. Also, it is rational to choose induction after $38^{0/7}$ weeks GA in case of suspected SGA to prevent possible neonatal morbidity and stillbirth.[6,38]

Women should be informed that—in case of suspected SGA at term—the risk of adverse pregnancy outcome is very small, but follow-up might be beneficial for them. The potential benefit for mother and child clearly outweighs the relatively light burden of follow-up ultrasounds.

Confirmation of SGA suspicion allows intervention, but caregivers should realize that intervention does not always improve outcome and does always bear risks for mother and child. Therefore, potential harm to mother and child in case of intervention should be weighed against the potential risk of expectant management on the other hand.

Note

The registry data are anonymous; therefore no ethical approval was needed. The Netherlands Perinatal Registry has given permission for the analysis of its data, approval number 11.42.

References

1 Stichting Perinatale Registratie Nederland. Perinatale Zorg in Nederland 2008. Available at: http://www.perinatreg.nl/uploads/150/122/Jaarboek_Zorg_in_Nederland_2008.PDF. Accessed July 01, 2013

2 Clausson B, Cnattingius S, Axelsson O. Preterm and term births of small for gestational age infants: a population-based study of risk factors among nulliparous women. Br J Obstet Gynaecol 1998; 105(9):1011–1017

3 Brodszki J, Marsál K. Management of pregnancies with suspected intrauterine growth retardation in Sweden. Results of a questionnaire. Acta Obstet Gynecol Scand 2000;79(9):723–728

4 Iams JD. Small for gestational age (SGA) and fetal growth restriction (FGR). Am J Obstet Gynecol 2010;202(6):513

5 Sheridan RM, Stanek J, Khoury J, Handwerger S. Abnormal expression of transcription factor activator protein-2α in pathologic placentas. Hum Pathol 2012;43(11):1866–1874

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

7 Clausson B, Gardosi J, Francis A, Cnattingius S. Perinatal outcome in SGA births defined by customised versus population-based birthweight standards. BJOG 2001;108(8):830–834

8 McIntire DD, Bloom SL, Casey BM, Leveno KJ. Birth weight in relation to morbidity and mortality among newborn infants. N Engl J Med 1999;340(16):1234–1238

9 Stoknes M, Andersen GL, Dahlseng MO, et al. Cerebral palsy and neonatal death in term singletons born small for gestational age. Pediatrics 2012;130(6):e1629–e1635

10 Pilliod RA, Cheng YW, Snowden JM, Doss AE, Caughey AB. The risk of intrauterine fetal death in the small-for-gestational-age fetus. Am J Obstet Gynecol 2012;207(4):e1–e6

11 Alfirevic Z, Neilson JP. Doppler ultrasonography in high-risk pregnancies: systematic review with meta-analysis. Am J Obstet Gynecol 1995;172(5):1379–1387

12 Ohel G, Ruach M. Perinatal outcome of idiopathic small for gestational age pregnancies at term: the effect of antenatal diagnosis. Int J Gynaecol Obstet 1996;55(1):29–32

13 Verlijsdonk JW, Winkens B, Boers K, Scherjon S, Roumen F. Suspected versus non-suspected small-for-gestational age fetuses at term: perinatal outcomes. J Matern Fetal Neonatal Med 2012; 25(7):938–943

14 Visser GHA, Eilers PHC, Elferink-Stinkens PM, Merkus HMWM, Wit JM. New Dutch reference curves for birthweight by gestational age. Early Hum Dev 2009;85(12):737–744

15 Méray N, Reitsma JB, Ravelli ACJ, Bonsel GJ. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. J Clin Epidemiol 2007; 60(9):883–891

16 Schaaf JM, Mol BWJ, Abu-Hanna A, Ravelli ACJ. Trends in preterm birth: singleton and multiple pregnancies in the Netherlands, 2000-2007. BJOG 2011;118(10):1196–1204

17 Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. Stat Med 2006;25(12):2084–2106

18 Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. Stat Med 2007;26(1):20–36

19 Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;59(10):1087–1091

20 Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. J Clin Epidemiol 2001;54(4):387–398

21 Hansen BB. Full matching in an observational study of coaching for the SAT. J Am Stat Assoc 2004;99(467):609–618

22 Hansen BB, Bowers J. Covariate balance in simple, stratified and clustered comparative studies. Stat Sci 2008;23(2):219–236

23 Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit Anal 2007;15(3):199–236

24 Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. J Stat Softw 2011; 42(8):1–28

25 Thoemmes F. Propensity Score Matching in SPSS. Tübingen, Germany: University of Tübingen; 2010. Available at: http://arxiv.org/ftp/arxiv/papers/1201/1201.6385.pdf. Accessed April 14, 2013

26 Iacus S, King G, Porro G. CEM: Software for Coarsened Exact Matching. J Stat Softw 2009;30(9):1–27

27 Chauhan SP, Beydoun H, Chang E, et al. Prenatal detection of fetal growth restriction in newborns classified as small for gestational age: correlates and risk of neonatal morbidity [published online ahead of print April 16, 2013]. Am J Perinatol. doi:10.1055/s-0033-1343771

28 Jahn A, Razum O, Berle P. Routine screening for intrauterine growth retardation in Germany: low sensitivity and questionable benefit for diagnosed cases. Acta Obstet Gynecol Scand 1998; 77(6):643–648

29 Jahn A, Razum O, Berle P. Routine screening for intrauterine growth retardation in Germany: low sensitivity and questionable benefit for diagnosed cases. Acta Obstet Gynaecol Scand 1998; 77:643–6482

30 Adams EK, Miller VP, Ernst C, Nishimura BK, Melvin C, Merritt R. Neonatal health care costs related to smoking during pregnancy. Health Econ 2002;11(3):193–206

31 Adams EK, Melvin CL, Raskind-Hood C, Joski PJ, Galactionova E. Infant delivery costs related to maternal smoking: an update. Nicotine Tob Res 2011;13(8):627–637

32 Castles A, Adams EK, Melvin CL, Kelsch C, Boulton ML. Effects of smoking during pregnancy. Five meta-analyses. Am J Prev Med 1999;16(3):208–215

33 DiFranza JR, Lew RA. Effect of maternal cigarette smoking on pregnancy complications and sudden infant death syndrome. J Fam Pract 1995;40(4):385–394

34 Campbell MK, Cartier S, Xie B, Kouniakis G, Huang WY, Han V. Determinants of small for gestational age birth at term. Paediatr Perinat Epidemiol 2012;26(6):525–533

35 Unterscheider J, Daly S, Geary MP, et al. Optimizing the definition of intrauterine growth restriction: the multicenter prospective PORTO Study. Am J Obstet Gynecol 2013;208(4):e1–e6

36 Mongelli M, Ek S, Tambyrajia R. Screening for fetal growth restriction: a mathematical model of the effect of time interval and ultrasound error. Obstet Gynecol 1998;92(6): 908–912

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

38 Spong CY, Mercer BM, D'alton M, Kilpatrick S, Blackwell S, Saade G. Timing of indicated late-preterm and early-term birth. Obstet Gynecol 2011;118(2 Pt 1):323–333

Appendix 1 Multivariable association between prenatal SGA (birth weight < 10th percentile) detection and adverse pregnancy outcome and perinatal interventions in the original dataset

Outcome	SGA detected prenatally (n = 163)	SGA not detected prenatally (n = 555)	p-Value	OR	95% CI
Delivery, n (%)					
Induction of labor	93 (57)	50 (9)	0.00	16.61	10.15–27.17
Instrumental delivery	43 (26)	82 (15)	0.07	1.56	0.97–2.50
Cesarean section	33 (20)	40 (7.3)	0.00	2.68	1.54–4.64
Primary cesarean section	7 (4.3)	4 (0.7)	0.00	8.34	2.07–33.66
Cesarean section in labor	26 (16)	36 (6.5)	0.02	2.06	1.14–3.74
Vaginal instrumental delivery	10 (6.1)	42 (7.6)	0.12	0.54	0.25–1.17
Neonatal outcome, n (%)					
Mean gestational age at birth (d, mean ± SD)	271 (9.6)	277 (6.0)	< 0.001	−6.5[a]	−7.9 to −5.2
Birth weight (g; median, IQR)	2,410 (1,420–3,080)	2,770 (1,765–3,250)	< 0.001	−343[a]	−392 to −294
Composite adverse neonatal outcome	9 (5.5)	27 (4.9)	0.36	0.67	0.29–1.57
Perinatal mortality	0 (0)	7 (1.3)	0.995	not calculable	
5 min Apgar score < 7	4 (2.5)	9 (1.6)	0.96	1.04	0.29–3.74
Neonatal complications[b]	3 (1.8)	10 (1.8)	0.21	0.40	0.09–1.67
Asphyxia	1 (0.6)	3 (0.5)		n.c.	
IRDS	0 (0)	1 (0.2)		n.c.	
Meconium aspiration	0 (0)	4 (0.7)		n.c.	
Pneumothorax	1 (0.6)	4 (0.7)		n.c.	
NEC	1 (0.6)	0 (0)		n.c.	
Convulsions	1 (0.6)	1 (0.2)		n.c.	
Sepsis	0 (0)	0 (0)		n.c.	
Meningitis	0 (0)	0 (0)		n.c.	

Abbreviations: CI, confidence interval; IRDS, infant respiratory distress syndrome; n.c., not calculated; NEC, necrotizing enterocolitis; OR, odds ratio; SD, standard deviation; SGA, small for gestational age (birth weight < 10th percentile).
[a]Mean difference and 95% CI.
[b]Number of infants with neonatal complications, some infants have more than one complication.

Appendix 2 Baseline characteristics of the matched original cohort. Pregnancies are stratified by prenatal SGA detection

Characteristic	SGA detected prenatally (n = 163)	SGA not detected prenatally (n = 163)	p-Value	Standardized difference (%)
Maternal characteristics[a]				
Age (y, mean ± SD)	29.3 ± 5.7	29.0 ± 5.9	0.71	4.2
Maternal BMI (median, range)[a]	22.5 (16–44)	22.4 (16–40)	0.51	9.5
Primiparity (%)	90 (55)	90 (55)	1.00	0.0
Smoking (%)	53 (32)	50 (31)	0.72	3.9
Gestational hypertension (%)	9 (5.5)	8 (4.9)	0.80	2.7
Neonatal outcome				
Birth weight percentile (median, range)	1.6 (0–9.9)	1.9 (0–9.9)	0.82	−2.5
Birth weight < p2.3 (%)	83 (51)	84 (52)	0.91	−1.2
Birth weight < p5 (%)	124 (76)	123 (76)	0.90	1.5

Abbreviations: BMI, body mass index; SD, standard deviation; SGA, small for gestational age (birth weight < 10th percentile).
[a]The body mass index is the weight in kilograms divided by the square of the height in meters.

Appendix 3 Propensity score analyses of prenatal SGA (birth weight $<$ p10) detection as predictor of adverse pregnancy outcome and perinatal interventions after 1:1 propensity score matching in the original dataset

Outcome	SGA detected prenatally ($n = 163$)	SGA not detected prenatally ($n = 163$)	p-Value	OR	95% CI
Delivery, n (%)					
Induction of labor	93 (58)	16 (9.9)	$<$ 0.001	12.04	6.59–21.99
Instrumental delivery	43 (26)	28 (17)	0.045	1.73	1.01–2.95
Cesarean section	33 (20)	11 (6.7)	0.002	3.51	1.71–7.22
Primary cesarean section	7 (4.3)	0 (0.0)		inf.	
Cesarean section in labor	26 (16)	11 (6.7)	0.01	2.62	1.25–5.51
Vaginal instrumental delivery	10 (6.1)	17 (10)	0.16	0.56	0.25–1.27
Neonatal outcome, n (%)					
Mean gestational age at birth (d, mean $\pm$ SD)	270.6 $\pm$ 9.6	278.1 $\pm$ 6.8	$<$ 0.001	−7.5[a]	−9.3 to −5.7
Birth weight (g; median, range)	2,410 (1,420–3,080)	2,650 (1,765–3,090)	$<$ 0.001	−225[a]	−295 to −156
Composite adverse neonatal outcome	9 (5.5)	13 (8.0)	0.38	0.67	0.28–1.62
Death	0 (0.0)	3 (1.8)	0.996	not calculable	
5 min Apgar score $<$ 7	4 (2.5)	4 (2.5)	1.00	1.00	0.25–4.07
Neonatal complications[b]	3 (1.8)	5 (3.1)	0.48	0.59	0.14–2.52
Asphyxia	1 (0.6)	0 (0.0)		n.c.	
IRDS	0 (0.0)	0 (0.0)		n.c.	
Meconium aspiration	0 (0.0)	3 (1.9)		n.c.	
Pneumothorax	1 (0.6)	4 (2.5)		n.c.	
NEC	1 (0.6)	0 (0.0)		n.c.	
Convulsions	1 (0.6)	0 (0.0)		n.c.	
Sepsis	0 (0.0)	0 (0.0)		n.c.	
Meningitis	0 (0.0)	0 (0.0)		n.c.	

Abbreviations: CI, confidence interval; OR, odds ratio; inf., infinite; IRDS, infant respiratory distress syndrome; n.c., not calculated; NEC, necrotizing enterocolitis; SD, standard deviation.
[a]Mean difference and 95% CI.
[b]Number of infants with neonatal complications, some infants have more than one complication.