Chapter 13

Summary

This thesis describes the results of the DIGITAT trial (Disproportionate Intrauterine Growth Intervention Trial At Term) and concentrates on strategies in intrauterine growth restriction (IUGR) at term.

Chapter 1

Chapter 1 forms the introduction of this thesis and gives a set-up to the DIGITAT trial. Around 9% to 11% of children is born with a birth weight below the 10th percentile and defined as small-for-gestational age (SGA). A significant part of these children are born in the term period. Pregnancies complicated by intrauterine growth restriction (IUGR) and children born small-for-gestational-age are known to have higher risk of perinatal mortality and neonatal morbidity, not only immediately after birth but also on the longer term. The terms intrauterine growth restriction (IUGR) and small-for-gestational-age (SGA) have been used interchangeably, creating confusion on the topic. In this way studies looking at associations between smallness and adverse neonatal outcomes have been blurred. In the introduction of this thesis definitions of SGA and IUGR are discussed. The methods to improve screening and identification of IUGR are discussed, as well as an insight in the latest opinions about the importance of Doppler-recordings of middle cerebral artery (MCA). To deal with at term pregnancies complicated by suspicion of IUGR there are two management strategies to approach the fetus at risk for morbidity and mortality. Would the fetus fare better by further growing and maturing in a possible undernourished environment, and thus postponing delivery with an expectant management? On the other hand induction of labour might pre-empt possible morbidity and stillbirth, may be at the cost of an increase in operative deliveries, and complications of relative (iatrogenic) prematurity. Results from prospective randomised trials as well as consensus among Dutch gynaecologists and residents about management and timing of delivery in at term IUGR were lacking. To deal with these questions about the controversial management of IUGR at term the DIGITAT trial (Disproportionate Intrauterine Growth Intervention Trial At Term) was designed. Embedded in the structure of the Dutch Obstetrical Consortium more than 50 hospitals, academic and non-academic, participated to this randomised controlled trial to enrol 650

women whose pregnancy was complicated by suspected IUGR at term. The aim of the study was to compare the effect of induction of labour with an expectant monitoring policy for suspected intrauterine growth restriction at term in singleton pregnancies in cephalic presentation beyond 36 weeks gestation on neonatal and obstetrical outcomes.

Chapter 2

This chapter presents the results of a retrospective cohort of SGA children in the Netherlands. Data of the National Dutch Perinatal Registry (PRN) were used of all nulliparae between 2000 and 2005 with a singleton in cephalic presentation beyond 36+0 weeks, with a birth weight below the 10th percentile. We analysed two groups of pregnancies: (I) with isolated SGA and (II) with both SGA and hypertensive disorders. Onset of labour was related to route of delivery and neonatal outcome. Induction was associated with a higher risk of emergency caesarean section (CS), without improvement in neonatal outcome. For women with isolated SGA the relative risk of emergency CS after induction was 2.3 (95% Confidence Interval [CI] 2.1 to 2.5) and for women with both SGA and hypertensive disorders the relative risk was 2.7 (95% CI 2.3 to 3.1). In concord with other retrospective studies it was concluded from this retrospective cohort that induction in pregnancies complicated by SGA at term was associated with a higher risk of instrumental deliveries without improvement of neonatal outcome.

Chapter 3

The full study protocol of the DIGITAT trial is described in this chapter. All women with a singleton pregnancy, with a child in cephalic presentation, with suspicion of IUGR (Fetal Abdominal Circumference < 10th centile, Estimated Fetal Weight < 10th percentile as defined by local protocols, or decreased relative growth though still > 10th centile) and a gestational age between 36+0 weeks and 41+0 weeks were eligible. Women with a history of caesarean section, serious congenital defects, ruptured membranes, renal diseases, diabetes mellitus, or positive HIV se-

rology were excluded. After written consent, pregnant women suspected of IUGR were randomised by means of a web-based application. Stratification was applied for previous vaginal birth (nullipara versus multipara) and for centre. Patients that withheld consent for randomisation were asked permission for data collection on pregnancy outcome and date were collected in the same prospective way. Before randomisation baseline demographics, past obstetric and medical history were collected for all women and cervical length was measured.
The study was staffed by obstetricians, research nurses, and midwives associated with the Dutch Obstetric Consortium. They counselled and recruited participants, monitored compliance with allocated treatment protocols, and collected outcome data.

Participants were either allocated to an induction of labour group, where induction had to take place within 48 hours of randomisation or to an expectant monitoring group. In the expectant group women had to be monitored until the onset of spontaneous labour with daily fetal movement counts and twice weekly fetal heart rate tracings, ultrasound examination, maternal blood pressure measurement, assessment of proteinuria, laboratory tests of liver and kidney function, and full blood count. Women were monitored as either an outpatient or an inpatient, according to local protocol. In the expectant monitoring group, induction of labour or planned caesarean section was performed for obstetrical indications—such as suboptimal fetal heart rate tracings, prolonged rupture of membranes, or postmaturity between T+7 and T+14 days—at the obstetrician's discretion. Women that were registered as non-participants were treated according to the opinion of their attending doctor with either an induction of labour or an expectant management policy.

The primary outcome was a composite measure of adverse neonatal outcome. This was defined as death before hospital discharge, five minute Apgar score of less than 7, umbilical artery pH of less than 7·05, or admission to neonatal intensive care. Secondary outcomes were delivery by caesarean section, instrumental vagi-

nal delivery, length of stay in the neonatal intensive care or neonatal ward, maternal length of stay in the hospital, and maternal morbidity. The latter was defined as post-partum haemorrhage of more than 1000 mL, development of gestational hypertension or pre-eclampsia (according to International Society for the Study of Hypertension in Pregnancy criteria), eclampsia, pulmonary oedema, thromboembolism, or any other serious adverse event.

Other secondary outcomes were a maternal health-related quality of life study and follow up of children's behavioural-, and (neuro)development by administering postal enquiries: the Child Behaviour Checklist-CBCL and Ages and Stages Questionnaire- ASQ after 2 years. Under assumption of comparable adverse outcomes a cost-minimisation analysis, in which only the costs of both strategies would be compared, was performed.

The trial was designed as an equivalence trial in which the null hypothesis was that the difference in the risk of the composite outcome between the two treatment groups was greater than 5.5% (absolute percentage). Assuming that the rate in the control group was 6% (on the basis of data from the National Dutch Perinatal Registry), this meant that we would exclude the null hypothesis and conclude that the two treatments were equivalent if the boundaries of the confidence interval of the observed risk difference were between -5.5% and 5.5%. With a 0.05 risk of type I error (α) and 80% (1-β) power, we calculated that we would require 650 participants (325 per group).

Data were analysed according to the intention to treat principle. Equivalence of the primary outcome measure was tested by checking if the 95% CI of the risk difference lay within the equivalence margins. Treatment effects were presented as differences in means or percentages with 95% confidence intervals (CI).

Chapter 4

The results of the DIGITAT trial are described in chapter 3. In this multicentre ran-

domised equivalence trial (the Disproportionate Intrauterine Growth Intervention Trial At Term (DIGITAT)) eight academic and 44 non-academic hospitals in the Netherlands participated to included eligible women between November 2004 and November 2008. A total of 1116 women with a singleton pregnancy in cephalic presentation, with a pregnancy suspected of IUGR beyond 36 weeks gestation were identified. Of these women, 466 declined randomisation, of whom 452 gave authorisation for use of their medical data.

321 women were randomly assigned to induction and 329 to expectant monitoring. Compared with the induction group, women in the expectant monitoring group were more likely to have a Bishop score of less than or equal to 6 and gestational hypertension, but otherwise the two randomised arms were comparable. Women who declined randomisation were older, had a higher education level, were less likely to smoke, had a lower body mass index (BMI), and were less likely to have a fetal abdominal circumference below the 10th centile. Most women who were randomised met either the fetal abdominal circumference below 10th centile inclusion criterion or the estimated fetal weight below the 10th centile criterion. Only 13 women in the induction group and 10 women in the expectant monitoring group were included because of flattening of the growth curve in isolation.

Induction was performed in 306 (95.6%) of the women in the induction group, resulting in a median time from randomisation to onset of labour of 0.9 days (IQR 0.7 to 1.7) in the induction group and 10.4 days (IQR 5.6 to 16.0) in the expectant monitoring group. In the expectant monitoring group labour was induced in 166 (50.6%) women.

Caesarean sections were performed on 45 (14.0%) mothers in the induction group and 45 (13.7%) in the expectant monitoring group (difference 0.3%, 95% CI -5.0% to 5.6%).

One (0.3%) woman allocated to induction of labour died at home 10 days after delivery. She had delivered a healthy child vaginally at 38+4 weeks of gestation after spontaneous onset of labour. No cause for her death was found at postmortem and it was classified as a serious unrelated adverse event. No women in the expectant monitoring group died during the study. All other maternal outcomes were comparable between the two groups.

There were no stillbirths or perinatal deaths. A total of 17 (5.3%) neonates in the induction arm and 20 (6.1%) neonates in the expectant monitoring arm had the primary composite adverse neonatal outcome (difference 0.8%, 95% CI - 4.3% to 2.8%). No differences between groups in any of the components of the composite adverse neonatal outcome were found. Median birth weight was lower in the induction group than in the expectant monitoring group (2420 g v 2550 g; difference 130 g, 95% CI 188 g to 71 g; P<0.001). Despite this difference, more fetuses in the expectant monitoring arm had a birth weight below the third percentile (100 (31%) v 40 (13%); difference 18.1%, 95% CI 24.3% to 12.0%; P<0.001).
More neonates in the induction group were admitted to a ward providing an intermediate level of neonatal care (155 (48.4%) v 118 (36.3%); difference 12.1%, 95% CI 4.6% to 19.7%; P<0.05).

In conclusion, we found equivalent fetal and maternal outcomes for induction and expectant monitoring in women with suspected intrauterine growth restriction at term, indicating that both approaches are acceptable. However, it is rational to choose induction to prevent possible neonatal morbidity and stillbirth on the grounds that we showed no increase in operative and instrumental delivery rates. By inducing labour in cases of intrauterine growth restriction, infants that will not grow any further can be released from their undernourished environment.

Chapter 5
In this chapter we describe details of a sub-analysis, by reporting neonatal morbidity between the two strategies based on a validated morbidity assessment index for newborns (MAIN). This score was developed to provide a numeric index of early neonatal outcomes of prenatal care and adverse prenatal exposures in babies delivered beyond 28 weeks of gestation. This sub-analysis was done mainly because we had found a significant difference in neonatal admissions to an intermediate type of care (48% v. 36%; difference 12%, 95% CI: 5% to 20%, P<0.05) in the DIGITAT trial. Complications of late prematurity might have explained this difference, since children in the induction group were born on average ten days earlier than in the

expectant group (266 days v 277 days; difference -9.9 days, 95% CI: -11 to -9). However, the difference may simply reflect policies for admission to intermediate levels of care related to prematurity rather than clinically relevant morbidity. In addition, more children were severely growth restricted in the expectant group, defined as a birth weight below the third percentile (13% v 31%; difference -18%, 95% CI -24% to -12%) and therefore had a possible higher risk of neonatal morbidity.

The MAIN score was assessed in 308 induction group babies and in 315 expectant management group babies. The categories of the MAIN scores (no/minimal, mild, moderate and severe morbidity) did not differ between the induction and expectant group. Morbidity in at term IUGR was relatively mild, and comparable for both induction and expectant management. When we looked at components of the MAIN score, more children suffered from hyperbilirubinemia >220 mmol/L or the need for phototherapy after induction of labour (n=32 (10.4%) for induction v n=18 (5.7%) for expectant management; difference 4.7%, 95% CI 0.4% to 8.9%, $p<0.05$).

For the outcomes neonatal admissions, a positive MAIN score and composite adverse outcome, we compared induction to an expectant management in women randomised before 38 weeks, from 38 till 40 weeks and after 40 weeks. The only difference was a higher percentage of neonatal admissions after induction before 38 weeks gestational age; 125 (61%) admissions v 92 (44%) after expectant management; difference 16%, 95% CI 6.7% to 26%, p=0.001).

We concluded that the apparent excess of neonatal care admission in the induction arm of the DIGITAT trial was probably a benign side effect of late prematurity and neonatal admission policies, rather than a marker of serious neonatal morbidity. If a policy of induction for near term growth restriction is to be followed, deferring induction until 38 weeks if feasible, while strictly monitoring mother and child, may prevent complications of late prematurity. Late effects of these policies need further study.

Chapter 6

Chapter 6 describes the results of the non-randomised women. All pregnant women who had a singleton pregnancy beyond 36+0 weeks' gestation with suspected intrauterine growth restriction who declined randomisation in the DIGITAT trial, but who gave authorisation for the use of their medical data were registered as non-participants. Identical data were collected prospectively.

The same primary outcome, a composite measure of adverse neonatal outcome (neonatal death before hospital discharge, a 5-minute Apgar score < 7, an umbilical artery pH <7·05 or admission to the neonatal intensive care unit) was used as well as operative delivery. Comparisons were between participants and non-participants, regardless of the group they were randomised to or treatment received.

In addition to 650 randomised women, 452 women consented for use of their medical data. Non-participants were older, had a lower body mass index (BMI), smoked less frequently and had a higher level of education.

A total of 37 (6%) infants of participants experienced the composite adverse neonatal outcome, compared with 32 (8%) in the non-participants (adjusted difference -2.0%, 95% CI -5.2% to 1.1%). In the non-participants group 3 (0.7%) deaths (2 stillbirths, 1 neonatal death) occurred, whereas no perinatal deaths occurred in the randomised group of women (difference -0.7%, 95% CI -1.4% to 0.1%, p=0.06). Caesarean sections were performed on 90 (14%) participants and on 71 (16%) non-participants (adjusted difference -2.8%, 95% CI -7.5% to 1.8%). In almost all comparisons, we found a tendency towards a more favourable neonatal outcome in women who were randomised. After adjustment for baseline imbalances in maternal age, smoking, BMI, education level and hypertensive disorders the adjusted difference and (95% CI) for perinatal death after participation in the trial was -0.5% (-1.4% to 0.4%, p=0.27).

We found a tendency towards more favourable outcomes in women randomised to the DIGITAT trial than in women who refused to participate, even after adjusting for baseline characteristics. We concluded that participation in a randomised clinical trial on growth restriction did not increase the risk of bad outcome. This information can be used when counselling women for trials.

Chapter 7

In this chapter Maternal health-related quality of life after induction of labour or expectant monitoring in pregnancy complicated by intrauterine growth retardation at term is described.

Both randomised and non-randomised women were asked to participate in the health-related quality of life (HR-QoL) study. Women were asked to fill out written validated questionnaires, covering background characteristics, condition-specific issues and the Short Form (SF-36), European Quality of Life (EuroQoL 6D3L), Hospital Anxiety and Depression scale (HADS), and Symptom Check List (SCL-90) at baseline (before and after randomisation), 6 weeks postpartum and 6 months postpartum. We compared the difference scores of all summary measures between the two management strategies by ANOVA. A repeated measures multivariate mixed model was defined to assess the effect of the management strategies on the physical (PCS) and mental (MCS) components of the SF-36. The analysis was by intention to treat. 361 randomised and 198 non-randomised patients were analysed. There were no clinically relevant differences between the treatments (induction or expectant management) at 6 weeks or 6 months postpartum on any summary measures; e.g., on the SF-36 (PhysicalComponentScore (PCS): P = 0.09; MentalComponentScore (MCS): P = 0.48). The PCS and the MCS were below norm values at inclusion. The PCS improved over time but stayed below norm values at 6 months, while the MCS did not improve. Main conclusion was that in pregnancies complicated by IUGR beyond 36 weeks, induction of labour does not affect the long-term maternal quality of life compared to expectant management

Chapter 8

After showing comparable medical outcomes and QoL, the economic impact of the two strategies is analysed. We used a health care perspective, in which only medical costs are included, with a time horizon from randomisation until hospital discharge. Thereby, by documenting details on utilisation of health care resources, we provided insight in the clinical origins of costs associated with management of these high-risk pregnancies. As both strategies were comparable in terms of health

outcomes, we performed a cost-minimisation analysis in which only the costs of both strategies were compared. We differentiated three phases of the clinical process in which costs arise: ante partum costs (from the moment of randomisation until childbirth), costs related to the delivery, and postpartum costs (from the moment of childbirth until hospital discharge). Resource use during the admission period was documented in the Case Report Form (CRF). The following resource items were collected: maternal and neonatal admissions, method of delivery, outpatient visits, medication, maternal laboratory tests, cardiotocograms (CTGs) and fetal ultrasounds. Maternal admissions were differentiated into three levels of care (intensive, medium, or ward). Neonatal admissions were divided into four levels of care (intensive, high, medium, or ward). Ante partum expectant monitoring generated more costs, mainly due to longer ante partum maternal stays in hospital. For the durante partu and postpartum stage, induction generated more direct medical costs, due to longer stay in the labour room and longer duration of neonatal high care/medium care admissions. From a health care perspective, both strategies generated comparable costs: on average € 7,106 per patient for the induction group (N=321) and € 6,995 for the expectant management group (N=329) with a cost difference of € 111 (95%CI: - € 1,296 to € 1,641). We can conclude that in women with pregnancies complicated by IUGR at term, induction of labour generates identical health care costs as compared to expectant management.

Chapter 9

In a secondary analysis we studied long term outcomes looking at the effects on (neuro)developmental and behavioural outcome at 2 years of age of induced labour compared with expectant management in intrauterine growth restricted infants. Parents of 2-year old children included in the DIGITAT-trial were asked to answer the Ages and Stages Questionnaire (ASQ) and Child Behaviour Check List (CBCL). The Ages and Stages Questionnaire is a screening questionnaire designed to detect developmental delay in children. The Child Behavior Checklist consists of 100 items concerning behavioural problems, on the basis of which a Total Problem score can be computed. It also informs on 7 narrow band syndrome scales (emo-

tionally reactive, anxious/depressed, somatic complaints, withdrawn, sleep problems, attention problems and aggressive behaviour), and two broad-band scales (internalizing and externalizing behaviour). We approached 582 (89.5%) of 650 parents. The response rate was 50%. Of these children, 27% had an abnormal score on the ASQ and 13 % on the CBCL. Results of the ASQ and the CBCL for the two policies were comparable. Low birth weight, positive morbidity assessment index (MAIN score) and admission to intermediate care, increased the risk of an abnormal outcome of the ASQ. This effect was not seen for the CBCL. With this secondary analysis we showed that in women with IUGR at term, both a policy of induction of labour and expectant management do not affect developmental and behavioural outcome when compared to expectant management.

Chapter 10

Whereas medical outcomes, maternal health-related QoL, costs and also long-term (neuro)development at 2 years of age of children born after IUGR are comparable between induction and expectant management women's preferences for one of the two strategies become even more interesting. To gain insight into how women value different obstetrical outcome scenarios, we compared induction of labour and expectant monitoring in intrauterine growth restriction at term through integration of trial outcomes and patient preferences. We used case scenarios ('vignettes'), involving five important factors ('attributes') that were evaluated by 24 trial participants using a discrete choice experiment (DCE) and by visual analogue scale (VAS). We combined these outcome valuations with outcome distributions of the RCT, and calculated a mean outcome for the strategies induction of labour and expectant management, respectively. These mean values were compared between the treatment groups using t-test for the total group and for subgroups, which were defined according to parity and gestational age. Using the DCE there was no overall treatment preference for the total group or for any of the subgroups. The VAS, however, did indicate preference towards expectant management for the total group as well as for subgroups. Based on the theoretical superiority of the DCE over the VAS method, the DCE results were leading. Therefore patient's prefer-

ences for expectant monitoring and induction of labour in case of IUGR at term were equal. These results reflected the outcomes of the DIGITAT trial.

Chapter 11

Chapter 11 presents the results from a retrospective cohort study among all term singleton neonates with a birth weight <10th percentile born in the Parkstad region (Heerlen) between 01-01-2006 and 03-31-2008. The aim of the study was to compare perinatal outcomes of suspected versus non-suspected small-for-gestational age fetuses (SGA) at term. The subjects were assigned to a prenatally Suspected or Non-Suspected SGA group. SGA was considered suspected when this was described unambiguously in the mothers' pregnancy chart. The clinical surveillance protocol used in case of prenatally suspected SGA comprised of fetal ultrasounds with fetal-placental Doppler velocimetry weekly, and cardiotocography twice weekly or more frequently depending on the severity of the growth restriction. Primary outcome was adverse neonatal outcome at birth, defined as a composite of intrauterine fetal death, Apgar <7 at 5 minutes, or pH umbilical artery <7.05. Secondary outcome included neonatal medium care unit (NMCU) admission ≥7 days. A total of 430 subjects were included in the study; 36.7% was suspected of SGA. In the Suspected SGA group mean gestational age at birth and birth weight were significantly lower, whereas maternal morbidity was significantly higher. The incidence of labour induction and elective caesarean section were also significantly higher in the Suspected SGA group. Total perinatal mortality was 2.1%. The crude odds ratio of adverse neonatal outcome at birth when comparing Suspected with Non-Suspected SGA at term, was 0.40 (95% CI 0.16-1.02, p=0.056). After correction for birth weight and hypertensive disorders, it was found that identification and subsequent labour and delivery management led to a significant decrease of adverse neonatal outcome at birth (OR 0.28, 95%CI 0.10-0.79, p= 0.016).
Identification of SGA and subsequent management led to a significant decrease of adverse neonatal outcome at birth, but did not lead to a significant decrease in NMCU admissions longer than 7 days. In conclusion suspicion of SGA was associated with a more active management of labour and delivery, resulting in a better neonatal outcome at birth.

Chapter 12

General discussion - principle findings:

- In a retrospective Dutch cohort of children born small for gestational age (SGA) induction of labour after 36 weeks gestation was associated with a higher risk of emergency caesarean section (CS), without improvement in neonatal outcome.
- The DIGITAT trial, basis of the thesis, concluded based on equivalence of the primary outcome of the trial, a composite adverse outcome of neonatal morbidity, that both induction of labour as well as an expectant management policy are safe strategies in at term IUGR.
- Induction of labour did not lead to higher rates of vaginal operative deliveries or an increase of emergency caesarean sections in the DIGITAT study.
- Even though both policies are safe, it is not unreasonable to induce labour to pre-empt the most devastating outcome in IUGR, stillbirth.
- Significantly more babies were admitted to intermediate type of neonatal care (high care and medium care) after a policy of induction of labour.
- More children get severely growth restricted after a policy of expectant management (<P 2.3).
- The MAIN-score was comparable for both induction group babies as well as for expectant management group babies. More children had a positive MAIN-score when born before 38 weeks gestational age, as compared to children born beyond 38 weeks gestation. Therefore, for as long as neonatal and maternal condition is reassuring, it is feasible to defer delivery beyond 38 weeks gestational age in at term IUGR.
- We showed that participating in a RCT on IUGR did not increase the risk of bad outcome, and this information can be used when counselling women for participation in a RCT.
- Induction of labour in at term IUGR does not affect the long-term maternal quality of life.
- From a health-care perspective, induction and expectant management generate comparable costs.

- Patient's preferences for expectant monitoring and induction of labour in case of IUGR at term are equal, and reflects the equivalence of the primary outcomes on top of medical outcomes, costs and QoL.
- We found no significant differences in developmental or behavioural outcomes at 2-years of age in children born at term with a clinical suspicion of growth restriction under a policy of induction of labour v expectant management. Severe growth restriction (P<2.3) and neonatal admission were found to be the most important predictive factors for (neuro)developmental problems at 2 years of age in children born after suspected IUGR at term.
- Suspicion of IUGR compared to cases where IUGR is not identified as such led to a more active management of labour and delivery, resulting in better neonatal outcomes at birth.

In conclusion, induction of labour and expectant management, while strictly monitoring mother and child both are safe strategies in at term growth restriction. Concerning obstetrical and neonatal outcomes - not only immediately after birth, but also on the long-term, health costs, maternal quality of life and maternal preferences, both strategies are comparable. To pre-empt the devastating outcome of stillbirth it is reasonable to induce labour after 38 weeks of gestation.
Hypothetically we could prevent 1 neonatal admission due to complications of relative prematurity, by delaying induction in 10 pregnancies suspected of IUGR beyond 38 weeks. Further delaying delivery to later gestational ages will increase the proportion of severely growth restricted children (<P2.3) which is not desirable. To determine genuine growth restriction and to detect the fetuses at highest risk for adverse outcome remains a great challenge. Customised growth, development of diagnostic risk scores and integration of UA-, and MCA-Doppler recordings are entries for future studies in at term IUGR. By development of treatment selection markers we can evaluate if tailor-made treatment for the individual women whose pregnancy is complicated by growth restriction at term is possible; to induce labour or to await spontaneous delivery with expectant management.