Induction of labour at 41 weeks versus expectant management until 42 weeks (INDEX): multicentre, randomised non-inferiority trial

Judit KJ Keulen,[1] Aafke Bruinsma,[1] Joep C Kortekaas,[2] Jeroen van Dillen,[2] Patrick MM Bossuyt,[3] Martijn A Oudijk,[1,4] Ruben G Duijnhoven,[1] Anton H van Kaam,[5] Frank PHA Vandenbussche,[2] Joris AM van der Post,[1] Ben Willem Mol,[6] Esteriek de Miranda[1]

[1]Department of Obstetrics and Gynaecology, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, Netherlands

[2]Department of Obstetrics and Gynaecology, Radboud University Medical Centre, Nijmegen, Netherlands

[3]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, Netherlands

[4]Department of Obstetrics and Gynaecology, University Medical Centre, Utrecht, Netherlands

[5]Department of Neonatology, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Netherlands

[6]Department of Obstetrics and Gynaecology, Monash University, Clayton, Victoria, Australia

Correspondence to:
E de Miranda
e.demiranda@amc.uva.nl
(ORCID 0000-0002-0392-4626)

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

ABSTRACT

OBJECTIVE
To compare induction of labour at 41 weeks with expectant management until 42 weeks in low risk women.

DESIGN
Open label, randomised controlled non-inferiority trial.

SETTING
123 primary care midwifery practices and 45 hospitals (secondary care) in the Netherlands, 2012-16.

PARTICIPANTS
1801 low risk women with an uncomplicated singleton pregnancy: randomised to induction (n=900) or to expectant management until 42 weeks (n=901).

INTERVENTIONS
Induction at 41 weeks or expectant management until 42 weeks with induction if necessary.

PRIMARY OUTCOME MEASURES
Primary outcome was a composite of perinatal mortality and neonatal morbidity (Apgar score ‹7 at five minutes, arterial pH ‹7.05, meconium aspiration syndrome, plexus brachialis injury, intracranial haemorrhage, and admission to a neonatal intensive care unit (NICU)). Secondary outcomes included maternal outcomes and mode of delivery. The null hypothesis that expectant management is inferior to induction was tested with a non-inferiority margin of 2%.

RESULTS
Median gestational age at delivery was 41 weeks+0 days (interquartile range 41 weeks+0 days-41 weeks+1 day) for the induction group and 41 weeks+2 days (41 weeks+0 days-41 weeks+5 days) for the expectant management group. The primary outcome was analysed for both the intention-to-treat population and the per protocol population. In the induction group, 15/900 (1.7%) women had an adverse perinatal outcome versus 28/901 (3.1%) in the expectant management group (absolute risk difference −1.4%, 95% confidence interval −2.9% to 0.0%, P=0.22 for non-inferiority). 11 (1.2%) infants in the induction group and 23 (2.6%) in the expectant management group had an Apgar score ‹7 at five minutes (relative risk (RR) 0.48, 95% CI 0.23 to 0.98). No infants in the induction group and three (0.3%) in the expectant management group had an Apgar score ‹4 at five minutes. One fetal death (0.1%) occurred in the induction group and two (0.2%) in the expectant management group. No neonatal deaths occurred. 3 (0.3%) neonates in the induction group versus 8 (0.9%) in the expectant management group were admitted to an NICU (RR 0.38, 95% CI 0.10 to 1.41). No significant difference was found in composite adverse maternal outcomes (induction n=122 (13.6%) v expectant management n=102 (11.3%)) or in caesarean section rate (both groups n=97 (10.8%)).

CONCLUSIONS
This study could not show non-inferiority of expectant management compared with induction of labour in women with uncomplicated pregnancies at 41 weeks; instead a significant difference of 1.4% was found for risk of adverse perinatal outcomes in favour of induction, although the chances of a good perinatal outcome were high with both strategies and the incidence of perinatal mortality, Apgar score ‹4 at five minutes, and NICU admission low.

TRIAL REGISTRATION
Netherlands Trial Register NTR3431.

WHAT IS ALREADY KNOWN ON THIS TOPIC

A policy of labour induction at or beyond term compared with expectant management is associated with fewer perinatal deaths and fewer caesarean sections; but more operative vaginal births (Cochrane review)

Aggregated results of trials need to be interpreted with caution because of trials heterogeneity caused by different outcome measures, protocols, and time frames of comparison

Evidence is lacking for the recommendation to induce labour at 41 weeks instead of 42 weeks for the improvement of perinatal outcome

WHAT THIS STUDY ADDS

Induction of labour at 41 weeks resulted in less overall adverse perinatal outcome than a policy of expectant management until 42 weeks, although the absolute risk of severe adverse outcome (perinatal mortality, NICU admission, Apgar score ‹4 at five minutes) was low in both groups

Introduction

Post-term pregnancy, defined as a pregnancy extended to or beyond 42 weeks, or 294 days or more, is associated with increased perinatal morbidity and mortality.[1-10] The World Health Organization and various guidelines throughout the world therefore recommend induction of labour after 42 weeks.[10-15] Although the overall probability of favourable perinatal outcomes between 40 and 42 weeks is good in high resource settings, the risk of adverse perinatal outcome increases gradually after 40 weeks.[16-19]

Several studies concluded that induction of labour from 41 weeks onwards improves perinatal outcomes,

and this has been confirmed in a meta-analysis.[16 17 19 20] These results need to be interpreted with caution, however, because of heterogeneity between trials as a result of different outcome measures, protocols, and time frames of comparison because several trials compared induction beyond 41 weeks or starting induction at 42 weeks with a policy of expectant management far beyond 42 weeks.[21]

The obstetric management of women with a pregnancy exceeding 41 weeks varies considerably between and within countries. Although induction at 41 weeks has now become an accepted policy in many countries, in some others no consensus exists on the timing of induction in late term pregnancy. In Sweden and the Netherlands, for example, expectant management until 42 weeks is considered standard of care in women with an uncomplicated pregnancy.[15 22] In Norway, induction is started no later than 42 weeks, and in Denmark delivery takes place before 42 weeks. Guidelines from the Royal College of Obstetricians and Gynaecologists/National Institute for Health and Care Excellence recommend that women should be offered induction between 41 and 42 weeks.[23]

We compared two strategies: induction of labour at 41 weeks (+0 days/+1 day) and expectant management until 42 weeks (+0 days) with subsequent induction if necessary. We anticipated that a policy of expectant management at 42 weeks, being the simpler strategy, would be acceptable for a low risk population if it did not lead to a substantially higher proportion of women with adverse perinatal outcomes compared with induction at 41 weeks.

Methods

Study design

Because induction of labour at 41 weeks as well as expectant management until 42 weeks are practised in the Netherlands, our study was designed to investigate non-inferiority of expectant management. We conducted a multicentre, open label, randomised controlled non-inferiority trial to investigate the effect of INDuction of labour at 41 weeks with a policy of EXpectant management until 42 weeks (INDEX trial) on adverse perinatal outcomes. Women were recruited at 123 primary care midwifery practices and 45 hospitals (secondary care) equally distributed across the Netherlands. Twenty six of these 45 hospitals actively recruited participants, and 19 supported the study by inducing labour in women who had been recruited in a primary care setting and were allocated to induction. In the Netherlands obstetric care is provided by primary care (midwives) for low risk women and secondary care (clinical midwives, residents, and obstetricians) for women with an increased risk of adverse maternal or perinatal outcome, or both. Low risk women in primary care can give birth at home or in an outpatient setting (birth centre or hospital), whereas women in secondary care give birth in hospital. For most low risk women, independent primary care midwives provide obstetric care. If risk factors are present during pregnancy, labour, or the postpartum period, women are referred to secondary care (obstetrician or gynaecologist). Secondary care may also be provided by clinical midwives or trainee obstetricians under the responsibility of an obstetrician.[24-30]
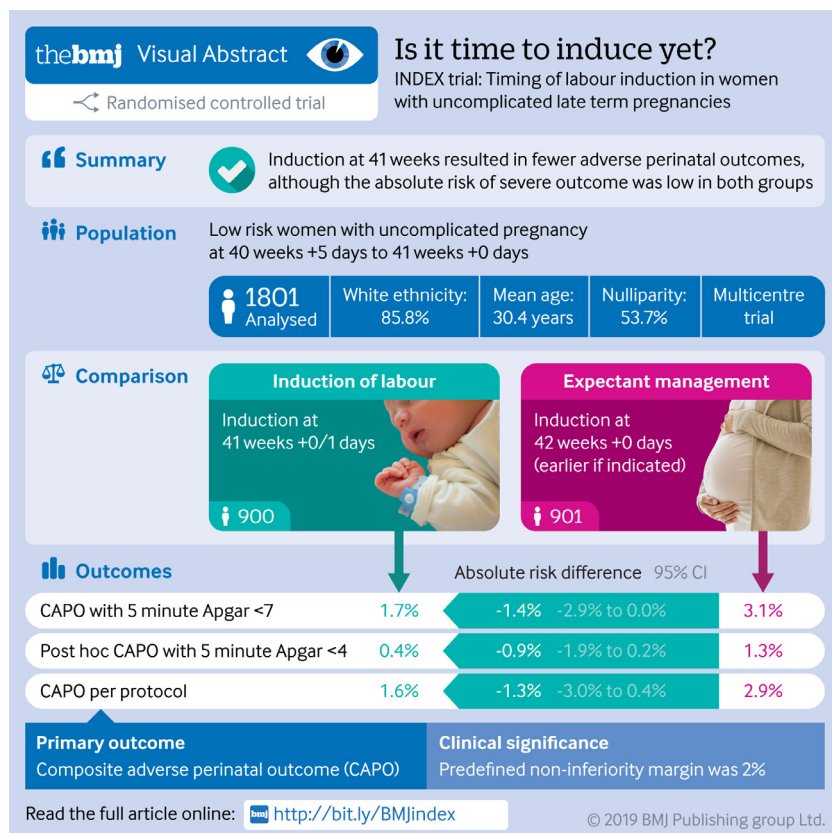
Our protocol has been published previously.[31] The study was performed within the Dutch Consortium for Healthcare Evaluation and Research in Obstetrics and Gynaecology in cooperation with the Midwifery Research Network Netherlands.

Participants

Women were eligible for the study if they had a low risk, uncomplicated singleton pregnancy with the child in a stable cephalic position at a certain gestational age of 40 weeks+5 days to 41 weeks+0 days and no contraindications to expectant management until 42 weeks. Gestational age had to be determined by ultrasonography before a gestational age of 16 weeks. Exclusion criteria for the study were age younger than 18 years, ruptured membranes or in labour, or both, non-reassuring fetal status (eg, no fetal movements, or abnormal fetal heart rate and/or expected intrauterine growth restriction), known fetal abnormalities (including abnormal karyotype) that could influence perinatal outcome, contraindications to induction (including previous caesarean section), or contraindications to expectant management (eg, pregnancy induced hypertension).

Randomisation and masking

Eligible women were informed about the study at the 40 week antenatal check. At their next visit (40 weeks+5 days to 41 weeks+0 days) the women were counselled

thebmj Visual Abstract — Is it time to induce yet?
INDEX trial: Timing of labour induction in women with uncomplicated late term pregnancies

Randomised controlled trial

Summary — Induction at 41 weeks resulted in fewer adverse perinatal outcomes, although the absolute risk of severe outcome was low in both groups

Population — Low risk women with uncomplicated pregnancy at 40 weeks +5 days to 41 weeks +0 days

| 1801 Analysed | White ethnicity: 85.8% | Mean age: 30.4 years | Nulliparity: 53.7% | Multicentre trial |

Comparison

Induction of labour	Expectant management
Induction at 41 weeks +0/1 days	Induction at 42 weeks +0 days (earlier if indicated)
900	901

Outcomes — Absolute risk difference 95% CI

	Induction of labour	Absolute risk difference	95% CI	Expectant management
CAPO with 5 minute Apgar <7	1.7%	-1.4%	-2.9% to 0.0%	3.1%
Post hoc CAPO with 5 minute Apgar <4	0.4%	-0.9%	-1.9% to 0.2%	1.3%
CAPO per protocol	1.6%	-1.3%	-3.0% to 0.4%	2.9%

Primary outcome — Composite adverse perinatal outcome (CAPO)

Clinical significance — Predefined non-inferiority margin was 2%

Read the full article online: http://bit.ly/BMJindex

© 2019 BMJ Publishing group Ltd.

by the community midwife, secondary obstetric caregiver, or research-nurse or research-midwife of the participating centres collaborating in the Dutch Obstetric Research Consortium. After written informed consent had been obtained, the study participants underwent digital vaginal examination to determine the Bishop score which is used to assess the ripeness of the cervix before planning of induction of labour. It rates position, consistency, and dilation of the cervix and engagement of the fetal head (station) in a single score. Sweeping of the membranes was optional. Participants were randomly allocated by a web based program (ALEA) using randomly permuted block sizes of 4 and 2, stratified by centre to induction of labour at 41 weeks+0 days-1 week or to expectant management with subsequent induction if necessary at 42 weeks+0 days. Owing to the nature of the intervention it was not possible to blind the women or caregivers to treatment allocation.

Procedures

Women allocated to induction were scheduled for the procedure at 41 weeks+0 days-41 weeks+1 day. All women were primed or induced, or both according to local protocols. Women with a Bishop score of less than 6 received cervical priming with prostaglandin E1 (misoprostol, oral or vaginal), prostaglandin E2 (dinoprostone), Foley catheter or double balloon catheter, or a combination of these until amniotomy could be performed. Amniotomy was followed by intravenous oxytocin if required.

Women in primary and secondary care who were allocated to expectant management awaited spontaneous onset of labour until 42 weeks+0 days in their initial care setting, with monitoring according to local protocol. Monitoring typically involved a combination of cardiotocography, and sonographic assessment of amniotic fluid in secondary care at 41-42 weeks. Women in the expectant management group with ongoing pregnancies were scheduled for induction at 42 weeks+0 days in secondary care, following a similar induction protocol to the intervention group.

In both groups, labour was induced if the maternal or fetal condition was no longer reassuring—for example, reduced fetal movements, non-optimal cardiotocography findings, or oligohydramnios. Labour was also induced if prelabour rupture of membranes had occurred more than 24 hours previously or meconium stained amniotic fluid was present.

The caregivers systematically collected information on perinatal and maternal condition, as well as protocol deviations and the reasons for these. Every case report form was checked on completion and inconsistency. Trained staff entered data in an online digital case report form (Oracle Clinical, version 4.6.6.4.1). Anonymised source documents were collected at the midwifery practice or hospital to check adverse perinatal and maternal outcomes. Serious adverse events were reported on a case by case basis to an independent Data Safety and Monitoring Board and to the Dutch national internet portal for the submission,

review, and disclosure of medical-scientific research with participants (www.toetsingonline.nl).

Outcomes

The primary outcome was a composite of perinatal mortality and neonatal morbidity.

Perinatal mortality was defined as fetal death, intrapartum death, and neonatal death until 28 days. Neonatal morbidity was defined as having an Apgar score <7 at five minutes and/or an arterial umbilical cord pH <7.05 and/or meconium aspiration syndrome and/or plexus brachialis injury and/or intracranial haemorrhage and/or or being admitted to a neonatal intensive care unit (NICU). Though a neonate could suffer from more than one adverse event, it is counted as one composite adverse perinatal outcome (neonatal level).

We defined meconium aspiration syndrome as respiratory distress after birth in the presence of meconium stained amniotic fluid. NICU admissions were reviewed to reveal final diagnosis and presence of congenital anomalies.

The cut-off for Apgar score <7 at five minutes was based on the committee opinion of the American College of Obstetricians and Gynecologists and American Academy of Pediatrics (ACOG/AAP), 2006. October 2015, after trial registration and during inclusion for this study, the ACOG/AAP committee released an update, which stated that the inappropriate use of the Apgar score in outcome studies had led to an erroneous definition of asphyxia.[32] Although it is incorrect to use Apgar score alone to diagnose birth asphyxia, an Apgar score <4 at five minutes "can be considered as a non-specific sign of illness." Because of this mid-trial change of cut-off value, we also planned an additional analysis of the primary outcome including Apgar scores <4 instead of <7 at five minutes.

Secondary perinatal outcomes consisted of maternal outcomes: instrumental delivery (instrumental vaginal delivery, caesarean section), pain treatment (epidural, remifentanyl, pethidine), postpartum haemorrhage, and severe perineal injury (third or fourth degree perineal tear (obstetrical anal sphincter injuries (OASIS)). Other neonatal outcomes included admission to medium care, congenital abnormality, hypoglycaemia, neonatal infection or sepsis, and small for gestational age (<10th centile) or large for gestational age (>90th centile). We also added a composite of adverse maternal outcome and other delivery outcomes.

The composite adverse maternal outcome included postpartum haemorrhage (≥1000 mL), manual removal of the placenta, third or fourth degree perineal tear (obstetrical anal sphincter injuries), and admission to an intensive care unit (ICU). Other delivery outcomes concerned onset of labour, pain treatment during labour, use of tocolytics, maternal intrapartum infection, meconium stained amniotic fluid, gestational age at delivery, mode of delivery, episiotomy, total postpartum blood loss, and blood transfusion. Though a woman could experience more

than one adverse event, it is counted as one composite adverse maternal outcome.

For both the perinatal and the maternal composite outcomes, we also compared the individual components.

Statistical analysis

Before the start of the trial, we formed an expert panel, consisting of midwives, gynaecologists, and paediatricians, and methodologists to conceive the design, content, and execution of the trial. Using data on adverse perinatal outcomes in the Netherlands from the Perined registry (www.perined.nl/), we expected an incidence of 3% for the primary composite adverse perinatal outcome with both strategies. The panel made a reasoned choice about the acceptable difference in adverse perinatal outcome and feasibility of the trial. As a result, the non-inferiority margin (Δ) was defined as a 2% risk difference in incidence of the composite outcome favouring induction to justify a possible change in management strategy of pregnancies reaching a gestational age of 41 weeks+0 days.

With a one sided α of 0.05, the study could achieve a power (β) of more than 0.80 if 900 women were recruited in each trial arm (1800 women in total). Non-inferiority would be concluded if the lower limit of the 95% confidence interval of the risk difference excluded a 2% higher proportion of women with an adverse perinatal outcome in the group allocated to expectant management. We established a Data Safety Monitoring Board to review the accumulating data of the trial. Interim analyses were conducted on safety after 517 and 1088 women had been recruited.

The statistician who performed the analyses was blinded to the allocation of the participants and performed the analysis according to a predefined analysis plan. The analysis of the primary outcome was done for both the intention-to-treat groups and the per protocol groups. For the per protocol analysis, we selected all randomised women with start of cervical ripening or spontaneous onset of labour at 41 weeks+0 days or more. Subsequently we defined the per protocol induction group as women allocated to induction who received induction before 41 weeks+2 days or who had a spontaneous onset of labour before induction could be started (<41 weeks+2 days). The per protocol expectant management group included women allocated to expectant management with spontaneous onset of labour until 42 weeks+0 days, women with a medical reason for induction before 42 weeks+0 days during expectant management, and women with induction at 42 weeks+0 days or more.

For all outcomes we estimated relative risks (RR) or median or mean differences, with 95% confidence intervals. As appropriate, we investigated significance using χ^2 test, Fisher's exact test, t test, or Mann-Whitney U test statistics. We plotted Kaplan-Meier curves for the time between randomisation and birth. The log-rank test statistic was used to evaluate the difference in time to birth. Birth centiles were determined using national reference data for the Netherlands on birthweight, ethnicity, parity, and gestational age by week and day. Analyses were performed using SAS software for Windows, version 9.4 (SAS Institute, Cary, NC).

Patient and public involvement

No patients were asked for input in the creation of this article. Patient representatives will be asked to join a multidisciplinary working group consisting of (representatives of) obstetric caregivers (primary and secondary care) and neonatologists to create a new nationwide guideline addressing the management of late term pregnancy. Patients will also be involved in writing patient information brochures and a patient decision aid on this topic.

Results

Between 14 May 2012 and 17 March 2016, 6088 eligible women were invited to participate in the INDEX trial, of whom 4273 declined owing to a maternal preference for induction of labour or expectant management, or refusal to let randomisation determine the management strategy. After randomisation but before analysis, one woman (induction group) withdrew her consent, and 13 women did not to meet the eligibility criteria (n=6 induction and n=7 expectant management). Of the remaining 1801 participants, 900 were randomly allocated to the induction group and 901 to the expectant management group (fig 1). Baseline characteristics were comparable between the groups, except for nulliparity: induction 50.8% (457/900) and expectant management 56.7% (511/901). (table 1).

In the induction group, 28.9% (260/900) of the women had a spontaneous onset of labour before the planned induction, and 71.1% (640/900) underwent induction, of whom 59.7% (382/640) had cervical ripening (fig 2). In the induction group, 4.8% (43/900) of the women were not induced at 41 weeks+0 days–41 weeks+1 days but at 41 weeks+2 days or later.

In the expectant management group, 73.7% (664/901) of the women had a spontaneous onset of labour and 26.3% (237/901) were induced (55.7% (132/237) underwent cervical ripening). In the expectant management group, 35.9% (85/237) underwent induction at 42 weeks for post-term pregnancy, and 27.4% (65/237) underwent induction before 42 weeks due to medical reasons (eg, fetal condition in 15.6% (37/237), maternal condition in 9.7% (23/237)), whereas 36.7% (87/237) in the expectant management group underwent induction on request. The median gestational age at time of delivery was 287 days (interquartile range 287-288 days) corresponding with 41 weeks+0 days (interquartile range 41 weeks+0 days-41 weeks+1 day) for the induction group and 289 days (interquartile range 287-292 days), corresponding with 41 weeks+2 days (41 weeks+0 days-41 weeks+5 days) in the expectant management group (table 2). In both groups three quarters of the women had a Bishop score <6 at study entry. Figure 3 shows the time to delivery for both groups.

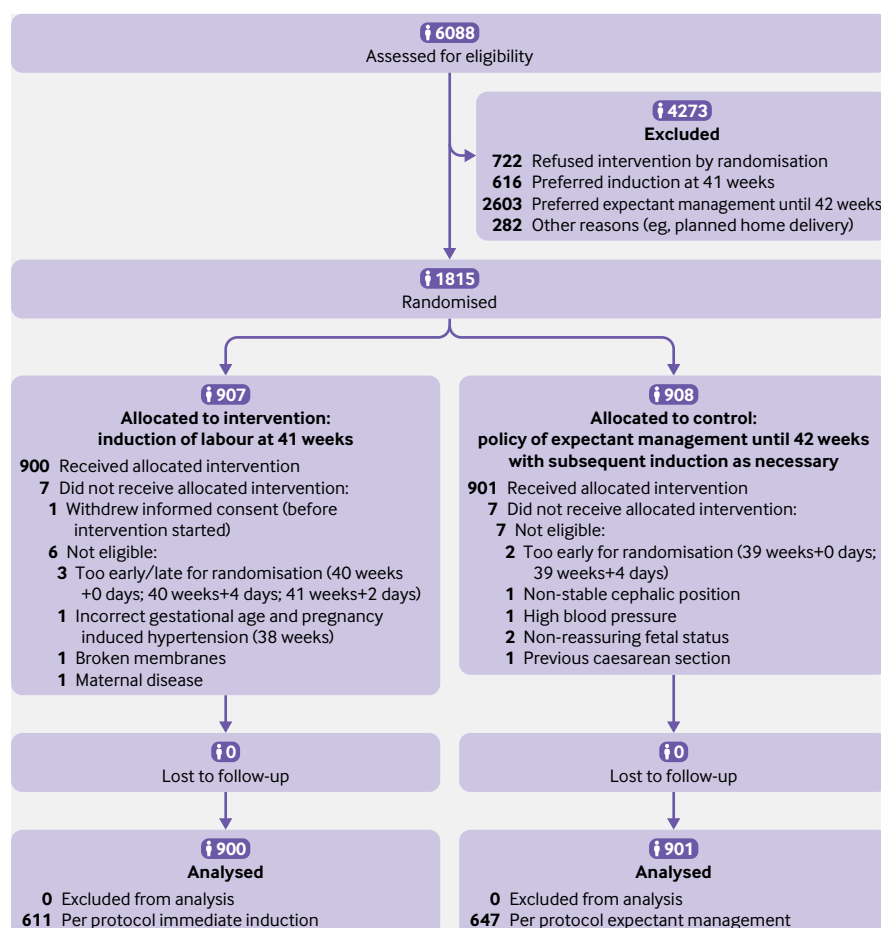In the per protocol induction group, 15.1% (92/611) of the women had spontaneous onset

Fig 1 | Flow of women through study

of labour before the planned induction. Of these women, 11.1% (67/611) had a spontaneous onset of labour at 41 weeks+0 days and 4.1% (25/611) at 41 weeks+1 day. In the per protocol induction group, 84.9% (519/611) of the women were induced: 62.5% (382/611) at 41 weeks+0 days and 22.4% (137/611) at 41 weeks+1 day. In the per protocol expectant management group, 80.9% (524/647) of the women had a spontaneous onset of labour at 41 weeks+0 days or later and 19.0% (123/647) were induced: 5.4% (35/647) because of concerns about fetal condition at 41 weeks+0 days or later, 3.4% (22/647) because of maternal condition 41 weeks+0 days or later, 0.6% (4/647) because of rupture of the membranes more than 24 hours previously at 41 weeks+0 days or later, and 9.6% (62/647) because of post-term pregnancy (≥42 weeks+0 days).

Primary outcome

Table 3 presents the perinatal outcomes in the intention-to-treat analysis. Fifteen women in the induction group (1.7%) and 28 in the expectant management group (3.1%) had a composite adverse perinatal outcome (absolute risk difference −1.4%, 95% CI −2.9% to 0.0%; number needed to treat (NNT) 69, 95% CI 35 to 3059). The P value for non-inferiority was 0.22, indicating that we could not exclude that

expectant management leads to 2% or more adverse perinatal outcomes compared with induction. All neonates in the expectant management group with a composite adverse perinatal outcome were born in secondary care. In these cases, women either had labour started in secondary care or were transferred during labour from primary to secondary care.

The per protocol analysis showed a 1.6% risk of an adverse perinatal outcome (10/611) in the induction group compared with 2.9% (19/647) in the expectant management group (risk difference −1.3%, 95% CI −3.0% to 0.4%, P=0.21 for non-inferiority; see supplementary appendix).

Additional analysis of the composite primary outcome including Apgar score <4 at five minutes instead of <7 resulted in 0.4% (4/900) adverse perinatal outcomes in the induction group and 1.3% (12/901) in the expectant management group (absolute risk difference −0.9%, −1.9% to 0.2%; NNT 113, 57 to 4624, P=0.02 for non-inferiority).

The additional per protocol analysis of the composite primary outcome including Apgar score <4 at five minutes showed a 0.5% risk (3/611) of an adverse perinatal outcome in the induction group versus 1.2% (8/647) in the expectant management group (risk difference −0.7%, −2.0% to 0.5%, P=0.02 for non-inferiority; see supplementary appendix).

Table 1 | Baseline characteristics of participants by intervention group. Values are numbers (percentages) unless stated otherwise

Characteristics	Induction of labour (n=900)	Expectant management (n=901)
Mean (SD) maternal age (years):	30.6 (4.8)	30.2 (4.6)
18-34	728 (80.9)	758 (84.2)
35-39	148 (16.4)	132 (14.7)
≥40	24 (2.7)	11 (1.2)
Ethnicity:		
White	779 (86.6)	767 (85.1)
Other	121 (13.4)	134 (14.9)
Body mass index at start of pregnancy:		
<18.5	26 (2.9)	19 (2.1)
18.5-<25	532 (59.1)	523 (58.1)
25-<30	230 (25.6)	229 (25.4)
≥30	89 (9.9)	117 (13.0)
Missing	23 (2.6)	13 (1.4)
Highest level of education:		
Primary school	7 (0.8)	4 (0.4)
Secondary school	37 (4.1)	15 (1.7)
Lower/medium professional education	358 (39.8)	350 (38.8)
Higher professional education/university	286 (31.8)	322 (35.7)
Other/unknown	212 (23.4)	210 (23.3)
Social economic status:		
Low	219 (24.3)	251 (27.9)
Medium	401 (44.6)	365 (40.5)
High	225 (25.0)	233 (25.9)
Unknown	55 (6.1)	52 (5.8)
Parity:		
Nulliparous	457 (50.8)	511 (56.7)
Multiparous	443 (49.2)	390 (43.3)
Previous post-term pregnancy (≥294 days)*	51/443 (11.5)	34/390 (8.7)
Level of care at recruitment:		
Primary	851 (94.6)	850 (94.3)
Secondary	49 (5.4)	51 (5.7)
Bishop score at study entry		
Nulliparous women:		
≥6	47/457 (10.3)	71/511 (13.9)
<6	360/457 (78.8)	365/511 (71.4)
Missing	50/457 (10.9)	75/511 (14.7)
Multiparous women:		
≥6	71/443 (16.0)	46/390 (11.8)
<6	310/443 (70.0)	294/390 (75.4)
Missing	62/443 (14.0)	50/390 (12.8)
Membrane sweeping before randomisation	286/900 (31.8)	343/901 (38.1)

*Numerator: multiparous women.

Three perinatal deaths (stillbirths) occurred: one in the induction group and two in the expectant management group. There were no neonatal deaths. The stillbirth in the induction group was in a 30 year old multiparous woman who was randomised at 40 weeks+5 days and scheduled for induction at 41 weeks+1 day. She had reduced fetal movements at 40 weeks+6 days, and fetal death was diagnosed at consultation. She delivered a neonate weighing 3595 g (20th to 50th centiles). Investigations, including a postmortem examination, did not explain the stillbirth. In the expectant management group, stillbirth was diagnosed in a 36 year old nulliparous woman at 41 weeks+3 days, when she was admitted to hospital in labour. She delivered a neonate weighing 2945 g (5th to 10th centiles). Investigations, including placental examination, did not explain the stillbirth, and the parents declined a postmortem examination. The second stillbirth in the expectant management group

was diagnosed in a 32 year old multiparous woman at 41 weeks+4 days during a regular consultation in secondary care for impending post-term pregnancy. She delivered a neonate weighing 3715 g (20th to 50th centiles). No postmortem examination was performed, but the placenta showed signs of chorioamnionitis.

The main contributor to the composite adverse outcome was an Apgar score <7 at five minutes: 1.2% (11/900) of neonates in the induction group and 2.6% (23/901) in the expectant management group (RR 0.48, 95% CI 0.23 to 0.98). Three of these neonates, all in the expectant management group, had an Apgar score <4 at five minutes. The first neonate was born at 41 weeks+6 days after spontaneous onset of labour and an operative vaginal delivery (vacuum), because of fetal distress and failure to progress in second stage. The diagnosis was meconium aspiration syndrome, and the neonate was admitted to the NICU (table 4). Sepsis after spontaneous onset of labour at 40 weeks+6 days and rupture of membranes of more than 24 hours was diagnosed in the second neonate. The third neonate, weighing 4320 g, was born after cervical ripening that started at 41 weeks+6 days and failure to progress of second stage followed by a caesarean section at 42 weeks+2 days. The diagnosis in this neonate was airway obstruction caused by vernix caseosa. Both these neonates were admitted to a medium care unit for observation. All three neonates recovered without complications. Admission to an NICU was reported in 0.3% (3/899) of neonates in the induction group versus 0.9% (8/899) in the expectant management group (RR 0.38, 95% CI 0.10 to 1.41). Of the 11 children admitted to the NICU, six (three in each group) had a diagnosis of severe congenital disorder. Meconium aspiration syndrome was diagnosed twice, but only in the expectant management group, and both neonates recovered fully. No plexus brachialis lesions and no intracranial haemorrhage were diagnosed in the study population. In two admissions because of a (suspected) infection, one neonate had group B streptococcus and the other had a negative culture result. One neonate was admitted because of a pneumothorax.

Arterial pH measurements were not recorded systematically and therefore could not be included in the analysis. Imputing was not possible owing to many missing data (62.0% induction *v* 70.0% expectant management). However, when we analysed data including the available pH measurements, the composite adverse perinatal outcome was 27/900 (3.0%) in the induction group versus 37/901 (4.1%) in the expectant management group (risk difference −1.11%, 95% CI −2.84% to 0.63%, P=0.16 for non-inferiority). For the per protocol analysis, the composite adverse perinatal outcome including the available pH measurement was 3.1% (19/611) in the induction group versus 4.0% (26/647) in the expectant management group (risk difference −0.91%, −2.98 to 0.01%, P=0.15 for non-inferiority).

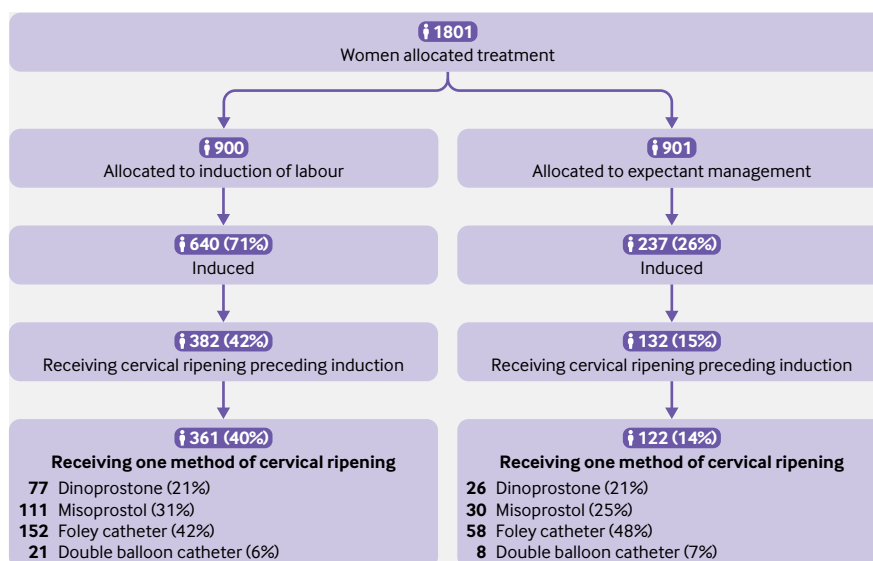When stratifying by parity, we observed 2.4% (11/457) nulliparous women with a composite adverse

```
                          ( 1801
                    Women allocated treatment

      ( 900                               ( 901
  Allocated to induction of labour    Allocated to expectant management

      ( 640 (71%)                         ( 237 (26%)
      Induced                             Induced

      ( 382 (42%)                         ( 132 (15%)
  Receiving cervical ripening         Receiving cervical ripening
  preceding induction                 preceding induction

      ( 361 (40%)                         ( 122 (14%)
  Receiving one method of             Receiving one method of
  cervical ripening                   cervical ripening
   77  Dinoprostone (21%)              26  Dinoprostone (21%)
  111  Misoprostol (31%)               30  Misoprostol (25%)
  152  Foley catheter (42%)            58  Foley catheter (48%)
   21  Double balloon catheter (6%)     8  Double balloon catheter (7%)
```

Fig 2 | Cervical ripening during study

perinatal outcome in the induction group and 4.1% (21/511) in the expectant management group (RR 0.59, 95% CI 0.29 to 1.20). In multiparous women the incidence of adverse perinatal outcome was lower in both groups compared with nulliparous women: 0.9% (4/443) in the induction group and 1.8% (7/390) in the expectant management group (RR 0.50, 95% CI 0.15 to 1.71). In logistic regression analysis, no interaction was found between parity and induction or expectant management.

Secondary outcomes

Table 3 shows the secondary perinatal outcomes in the intention-to-treat groups. No difference was found in medium care admissions, 6.6% and 6.7% (induction 59/899 v expectant management 60/899). Small for gestational age (<10th centile), according to Dutch birthweight centiles, was similar between the groups: 6.8% (61/900) in the induction group versus 6.9% (62/901) in the expectant management group. Overall, 9.6% (86/900) of infants in the induction group were large for gestation age (>90th centile) versus 11.0% (99/901) in the expectant management group. The incidence of congenital abnormalities was similar between groups: 1.8% in the induction group (16/900) versus 2.1% in the expectant management group (19/901).

Table 2 summarises the characteristics of labour and mode of delivery. Oxytocin was given significantly more often in the induction group than in the expectant management group (59.2% (533/900) and 39.4% (355/901) (RR 1.50, 95% CI 1.36 to 1.66)). Meconium stained amniotic fluid occurred significantly less often in the induction group compared with expectant management group (16.3% (147/900) and 22.8% (205/901) (RR 0.72, 95% CI 0.59 to 0.87). Ninety seven women in each group (10.8%) had a caesarean section (RR 1.00, 95% CI 0.77 to 1.31), mainly for non-progressive labour at the first stage of labour (table 2).

Table 5 shows the results of adverse maternal outcomes in the intention-to-treat groups. The composite adverse maternal outcome occurred in 13.6% (122/900) of the women in the induction group versus 11.3% (102/901) in the expectant management group (RR 1.20, 95% CI 0.94 to 1.53). Postpartum haemorrhage ≥1000 mL was the main contributor to the composite adverse maternal outcome and occurred in 9.1% (82/900) of women in the induction group versus 8.0% (72/901) in the expectant management group (RR 1.14, 95% CI 0.84 to 1.54). Manual removal of the placenta occurred in 5.1% (41/803) in the induction group versus 4.1% (33/804) in the expectant management group (RR 1.24, 95% CI 0.79 to 1.95). Obstetrical anal sphincter injuries were diagnosed in 3.5% (28/803) of women in the induction group versus 3.9% (31/804) in the expectant management group (RR 0.90, 95% CI 0.55 to 1.49). Three mothers (0.3%) in the induction group and two (0.2%) in the expectant management group were admitted to an ICU post partum (RR 1.50, 95% CI 0.25 to 8.97), all after postpartum haemorrhage. Blood loss in these women was 3000 mL, 5100 mL, and 7000 mL in the induction group and 3390 mL and 5000 mL in the expectant management group. No maternal deaths occurred. During labour, 29.4% (265/900) of the women in the induction group received epidural anaesthesia compared with 25.6% (231/901) in the expectant management group (RR 1.15, 95% CI 0.99 to 1.33).

Discussion

This randomised controlled trial compared the effect of induction of labour at 41 weeks with expectant management until 42 weeks with subsequent induction if necessary on perinatal and maternal outcomes in women with an uncomplicated pregnancy. A policy of induction resulted in a median reduction in gestational age at delivery of two days. We found a 1.4% difference

Table 2 | Delivery outcomes in intention-to-treat population. Values are numbers (percentages) unless stated otherwise

Outcomes	Induction of labour (n=900)	Expectant management (n=901)	Relative risk (95% CI)	P value
Median (interquartile range) gestational age delivery (days)	287 (287-288)	289 (287-292)	−2.1 (−2.3 to −1.9)*	<0.001†
Mean (SD) time from randomisation to delivery (days)	2.1 (1.6)	4.2 (3.0)	−2.2 (−2.5 to −2.0)*	<0.001†
Level of care at onset of labour:				
Primary	255 (28.3)	619 (68.7)	NC	-
Secondary	645 (71.7)	282 (31.3)	NC	-
Onset of labour:				
Spontaneous (reference)	260 (28.9)	664 (73.7)	1.00	-
Induction	640 (71.1)	237 (26.3)	2.70 (2.41 to 3.04)	<0.001
Mode of induction:	n=640	n=237		
Cervical ripening (catheter/prostaglandins)	382 (59.7)	132 (55.7)	1.07 (0.94 to 1.22)	0.30
Amniotomy without oxytocin	87 (13.6)	34 (14.8)	0.95 (0.66 to 1.37)	0.77
Amniotomy with oxytocin	156 (24.4)	59 (24.9)	0.98 (0.76 to 1.27)	0.87
Indication for induction:				
Randomisation	634 (99.1)	0 (0.0)	NC	-
Post-term pregnancy	0 (0.0)	85 (35.9)	NC	-
Fetal condition	5 (0.8)	37 (15.6)	NC	-
Maternal condition	0 (0.0)	23 (9.7)	NC	-
Elective or maternal request	1 (0.2)	87 (36.7)	NC	-
Membranes ruptured >24 h	0 (0.0)	4 (1.7)	NC	-
Other	0 (0.0)	1 (0.4)	NC	-
Use of oxytocin	533 (59.2)	355 (39.4)	1.50 (1.36 to 1.66)	<0.001
Use of tocolytics	28 (3.1)	16 (1.8)	1.75 (0.95 to 3.22)	0.07
Maternal intrapartum infection:	n=900	n=901		
Fever during labour (≥38°C)	50 (5.6)	46 (5.1)	1.09 (0.74 to 1.61)	0.67
Use of antibiotics	48 (5.3)	35 (3.9)	1.37 (0.90 to 2.10)	0.14
Meconium stained amniotic fluid	147 (16.3)	205 (22.8)	0.72 (0.59 to 0.87)	0.001
Level of care at time of birth:				
Primary	129 (14.3)	309 (34.3)	NC	-
Secondary	771 (85.7)	592 (65.7)	NC	-
Mode of delivery:				
Spontaneous vaginal	710 (78.9)	696 (77.2)	1.02 (0.97 to 1.07)	0.40
Operative vaginal	93 (10.3)	108 (12.0)	0.86 (0.66 to 1.12)	0.27
(Secondary) caesarean section	97 (10.8)	97 (10.8)	1.00 (0.77 to 1.31)	0.99
Indication successful operative vaginal delivery:	n=93	n=108		
Failure to progress at second stage	39 (41.9)	49 (45.4)	0.92 (0.67 to 1.27)	0.63
Suspected fetal distress	43 (46.2)	37 (34.3)	1.35 (0.96 to 1.90)	0.08
Suspected fetal distress and failure to progress	10 (10.8)	22 (20.4)	0.53 (0.26 to 1.06)	0.07
Maternal complication or other	1 (1.1)	0 (0.0)	NA	-
Indication for secondary caesarean section:	n=97	n=97		
Failure to progress at first stage	29 (29.9)	21 (21.6)	1.38 (0.85 to 2.25)	0.19
Failure to progress at second stage	12 (12.4)	18 (18.6)	0.67 (0.34 to 1.31)	0.24
Failed operative vaginal delivery	6 (6.2)	12 (12.4)	0.50 (0.20 to 1.28)	0.22‡
Suspected fetal distress	24 (24.7)	21 (21.6)	1.14 (0.68 to 1.91)	0.61
Suspected fetal distress and failure to progress at first stage	7 (7.2)	8 (8.3)	0.75 (0.17 to 3.26)	1.00‡
Suspected fetal distress and failure to progress at second stage	4 (4.1)	3 (3.1)	1.00 (0.26 to 3.88)	1.00‡
Maternal complication or other	15 (15.5)	14 (14.4)	0.93 (0.48 to 1.83)	0.84

NC=not calculable; NA=not applicable.
*Mean (95% confidence interval) difference between groups.
†Mann-Whitney U test.
‡Fisher's exact test.

in composite adverse perinatal outcome favouring induction, although the absolute risk of severe adverse perinatal outcome (perinatal mortality, Apgar score <4 at five minutes, admission to a neonatal intensive care unit (NICU) without severe congenital anomalies) was low in both groups.

Most of our primary composite outcomes can be attributed solely to the component Apgar score <7 at five minutes 73.3% (11/15) in the induction group *v* 64.3% (18/28) in the expectant management group), which means that these neonates did not have any other adverse outcome besides the Apgar score being <7 at five minutes. We performed a post hoc analysis of the composite outcome including Apgar scores <4 instead of <7 at five minutes owing to the American

College of Obsetricians and Gynecologists/American Academy of Pediatrics (ACOG/AAP) mid-trial change in recommended cut-off value for Apgar score at five minutes indicating a non-specific sign of illness. A considerable lower incidence of adverse perinatal outcome was found in both groups (0.4% induction and 1.3% expectant management), with an absolute risk difference of −0.9% (95% confidence interval −1.9% to 0.2%) favouring induction, showing non-inferiority of expectant management with respect to the predefined margin of 2% (P=0.02 for non-inferiority).

Comparison with other studies
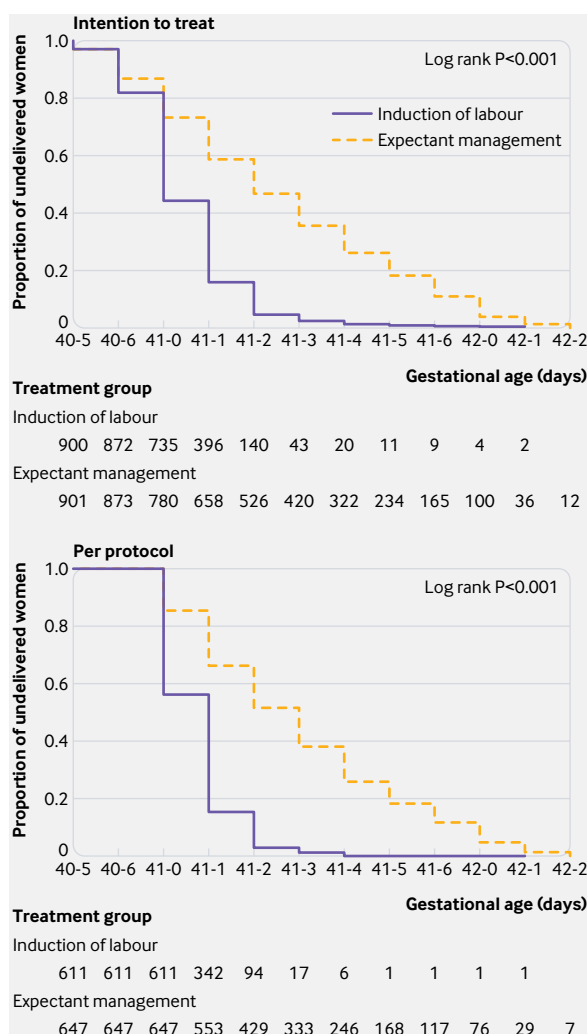The incidence of perinatal death in our study was one after induction compared with two after expectant

Fig 3 | Time to delivery

−0.9%, 95% CI −1.6% to 0.2%; P=0.01 for non-inferiority; P=0.02 for Fisher's exact test; and NNT of 112 (95% CI 63 to 491)).

Since in our trial all women in the 41 week induction group received obstetrician led intrapartum secondary care whereas in the expectant management group until 42 weeks 68.7% of the women received midwifery led primary care at start of labour and 34.3% at time of birth, it could be suggested that our study is prone to performance bias (different care) and measurement bias (different assessment of neonates). Several studies, however, showed that Apgar scoring does not differ significantly between midwives and obstetricians.[34] [35] [36] Furthermore, in our trial all neonates in the expectant management group with an adverse outcome were born in secondary care—the women had started labour in secondary care or were referred from primary to secondary care during labour. Various studies have shown that it is safe for low risk women in the Netherlands to deliver in midwifery led care, and the level of care does not seem to influence delivery outcome for these women.[24-28] Although this study could be considered as a comparison between obstetrician led care with labour induction and midwife led care with a policy of expectant management, we cannot adjudicate whether the difference in the composite adverse perinatal outcome is due to the level of care (performance bias) or to a possible difference in Apgar scoring (measurement bias). We do not, however, expect bias to be a major factor.

In our study, meconium aspiration syndrome occurred in two neonates in the expectant management group. In a randomised controlled trial with a comparable time frame, Gelisen et al reported meconium aspiration syndrome in 16/600 neonates of whom 12/300 were in the expectant management group. We found a 10 and 20 times lower rate of meconium aspiration syndrome (0.0% and 0.2% versus 1.3 and 4%) in the induction and expectant management groups compared with the study by Gelisen et al. Since these authors did not specify meconium aspiration syndrome, the difference in magnitude could be attributed to a difference in definition. Despite this, Gelisen et al found no difference in NICU admissions (4.3% induction v 5.0% expectant management), which is expected to be associated with meconium aspiration syndrome. We found a lower rate of NICU admissions compared with the Cochrane systematic review on induction of labour at more than 41 weeks: 0.3% induction and 0.9% expectant management (INDEX trial) v 11% induction and 12% expectant management (systematic review). The systematic review lacked details on NICU admission, such as diagnosis, potential association with gestational age, or presence of congenital anomalies, which hampers a clear comparison.[16] [33]

We did not find differences in caesarean section or operative vaginal delivery rates, which is consistent with other large studies on induction of labour.[37] [38] In the only study that compared the same timeframes as our study, the risk ratio for caesarean section was

management. The corresponding risk ratio for perinatal death (RR 0.50, 95% CI 0.05 to 5.51) is comparable with that of the four studies (n=998) starting induction at 41 weeks (n=501) versus expectant management with varying upper limits of gestational age (n=497) included in a Cochrane systematic review (RR 0.33, 95% CI 0.03 to 3.17).[19] [21] [33]

Congenital anomaly accounted for a substantial part of the NICU admissions in our trial, although it was an exclusion criterion at study entry. It is unknown if the outcome for these children would have been better if they had been born earlier, although it is unlikely (table 4). For these reasons we also analysed the primary composite outcome using an Apgar score <4 at five minutes and NICU admission without severe congenital anomalies. With these adapted adverse outcomes (perinatal mortality, Apgar score <4 instead of <7 at five minutes, and NICU admission without severe congenital anomalies), the absolute risk on the composite adverse perinatal outcome was substantially lower in both groups, with a still significant difference in favour of induction (0.1% (1/897)) versus expectant management (1.0% (9/898)): absolute risk difference

Table 3 | Perinatal outcomes in intention-to-treat groups

Outcomes	Induction of labour (n=900)	Expectant management (n=901)	Relative risk (95% CI)	P value
Composite adverse perinatal outcome*	15 (1.7)	28 (3.1)	0.54 (0.29 to 1.00)	0.045†
with 5 min Apgar score <4 instead of <7	4 (0.4)	12 (1.3)	0.33 (0.11 to 1.03)	0.06†
including arterial pH <7.05	27 (3.0)	37 (4.1)	0.72 (0.44 to 1.20)	0.16†
Stillbirth	1 (0.1)	2 (0.2)	0.50 (0.05 to 5.51)	1.00†
Neonatal death post partum	0 (0.0)	0 (0.0)	NA	-
Apgar score 5 mins post partum‡:				
<7	11 (1.2)	23 (2.6)	0.48 (0.23 to 0.98)	0.038
<4	0 (0.0)	3 (0.3)	NA	-
Neonate admitted to:				
NICU	3/899 (0.3)	8/899 (0.9)	0.38 (0.10 to 1.41)	0.23†
Medium care	59 (6.6)	60 (6.7)	0.98 (0.69 to 1.39)	0.90
Meconium aspiration syndrome§	0 (0.0)	2 (0.2)	NA	-
Plexus brachialis injury	0 (0.0)	0 (0.0)	NA	-
Intracranial haemorrhage¶	0 (0.0)	0 (0.0)	NA	-
Umbilical cord pH (arterial):				
<7.05	16 (1.8)	12 (1.3)	1.06 (0.51 to 2.20)	0.88
Missing	557 (62,0)	629 (70,0)	NA	-
Congenital abnormality	16 (1.8)	19 (2.1)	0.84 (0.44 to 1.63)	0.61
Hypoglycaemia**	3 (0.3)	6 (0.7)	0.50 (0.13 to 2.00)	0.51†
Neonatal infection/sepsis††	37 (4.1)	37 (4.1)	1.00 (0.64 to 1.56)	1.00
Female	453 (50.3)	463 (51.4)	0.98 (0.89 to 1.07)	0.65
Mean (SD) birthweight (g)	3685 (417.4)	3741 (430.0)	−56.6 (−95.8 to −17.4)‡‡	0.005
Small for gestational age:				
<2.3rd centile	13 (1.4)	11 (1.2)	1.18 (0.53 to 2.62)	0.68
<10th centilethcentile	61 (6.8)	62 (6.9)	0.99 (0.70 to 1.39)	0.93
Large for gestational age:				
>90th centile	86 (9.6)	99 (11.0)	0.87 (0.66 to 1.14)	0.32
>97th centile	15 (1.7)	27 (3.0)	0.56 (0.30 to 1.04)	0.07

NA=not applicable; NICU=neonatal intensive care unit.
*Composite outcome defined as perinatal mortality (fetal death, intrapartum death, and neonatal death until 28 days) or perinatal morbidity (a 5 minute Apgar score <7, and/or meconium aspiration syndrome, and/or plexus brachialis injury, and/or intracranial haemorrhage, and/or NICU admission). In the induction group all live births with a composite adverse perinatal outcome (CAPO) had either a 5 minute Apgar score <7 or a NICU admission. In the expectant management group, live births with a CAPO: two neonates had meconium aspiration syndrome, 5 minute Apgar score <7, and NICU admission; three neonates were admitted to NICU and also had a 5 minute Apgar score <7; three neonates were admitted to NICU but had no 5 minute Apgar score <7; 18 neonates had a 5 minute Apgar score <7 but no NICU admission.
†Fisher's exact test.
‡Apgar score of live births.
§Defined as respiratory distress after birth in presence of meconium stained amniotic fluid.
¶Defined as clinical signs of intracranial haemorrhage.
**Defined as glucose concentration <1.9 mmol/L and need for intravenous glucose.
††Defined as clinical suspected findings or proved positive blood culture result.
‡‡Mean (95% confidence interval) difference between groups.

comparable for both groups, although the absolute risk was twice as high compared with that of our study. This could be due to other inclusion criteria (Gelisen et al, Bishop score <5) or differences in policy during labour, as reflected by differences in national overall caesarean rates in Turkey (53% v 16% in the Netherlands).[35][39] The Cochrane systematic review concluded that induction at or beyond 41 weeks is associated with

lower caesarean section rates. The largest contribution to this outcome was from a randomised controlled trial in which women in the control group were induced only with oxytocin according to study protocol, whereas prostaglandin use was allowed in women with low Bishop scores in the induction group.[16] Two other systematic reviews including the same trial concluded that the difference in caesarean section

Table 4 | Admission to neonatal intensive care unit (NICU) by intervention

Allocation	NICU admission: diagnosis	Congenital anomaly	Gestational age at birth
Induction	Long QT syndrome	Yes	40 weeks+6 days
Induction	Mild mitralis insufficiency, persistent ductus arteriosus	Yes	41 weeks+0 days
Induction	Interstitial lung disorder	Yes	41 weeks+1 days
Expectant management	Diaphragm herniation, atrial septal defect, ventricular septal defect	Yes	40 weeks+6 days
Expectant management	Muscular ventricular septal defect	Yes	41 weeks+5 days
Expectant management	Vocal cord paresis, dysmorphic features	Yes	41 weeks+5 days
Expectant management	Infection (suspected, but culture was sterile)	No	41 weeks+2 days
Expectant management	Infection, Group B Streptococcus positive	No	41 weeks+3 days
Expectant management	Pneumothorax	No	41 weeks+4 days
Expectant management	Meconium aspiration syndrome	No	41 weeks+2 days
Expectant management	Meconium aspiration syndrome	No	41 weeks+6 days

rate is possibly due to the influence of this study with incomparable study arms.[17 20 40] Population based cohort studies showed conflicting results on the effect of induction on caesarean section rates.[41-43] In the recently published ARRIVE (A Randomized Trial of Induction Versus Expectant Management) trial, low risk nulliparous women were randomised in the 39th week of pregnancy to be induced at 39 weeks+0 days to 39 weeks+4 days or to expectant management until 41 weeks. No statistically significant difference was found in perinatal outcome (RR 0.80, 95% CI 0.64 to 1.00), although fewer caesarean sections took place in the induction group (18.6% v 22.2%; RR 0.84, 95% CI 0.76 to 0.93). Our study comprised more white women (86% v 44%), with a higher median age (30 v 24 years) and a lower percentage of body mass index ≥30 (12% v 52%), whereas 46% of the participants in our study were multiparous women and ARRIVE included only nulliparous woman. Caesarean section rates in our nulliparous low risk women were comparable between the groups: 18.6% in the induction group and 18.0% in the expectant group. This could be due to the

differences in gestational age, baseline characteristics, indication for induction of labour, or indication for a caesarean section (suspected fetal distress or failure to progress).[44]

Our trial had some notable results besides those for the main outcomes. Around 85% of participating women were of white ethnicity. The risk of perinatal mortality beyond term has been shown to be higher in women of South Asian, African, and Mediterranean origin compared with white women.[45] In our study, we were not able to assess the effect of induction in women of non-white ethnicity owing to the low number of women of other ethnic origin. Also, we were unable to assess the effect of age on adverse perinatal outcome because of the low number of participating older mothers (>35 years).

As in other studies on pregnancies at or beyond 41 weeks, most women in our study had an unfavourable cervix, with a Bishop score of <6 at randomisation. Although induction was planned one or two days after randomisation, 28.9% of the women in the induction group had a spontaneous onset of labour

Table 5 | Adverse maternal outcomes in intention-to-treat population. Values are numbers (percentages) unless stated otherwise

Adverse outcomes	Induction of labour (n=900)	Expectant management (n=901)	Relative risk (95% CI)	P value
Composite adverse maternal outcome*	122 (13.6)	102 (11.3)	1.20 (0.94 to 1.53)	0.15
Maternal death	0 (0.0)	0 (0.0)	NA	-
Postpartum blood loss:				
<1000 mL (reference)	818 (90.9)	829 (92.0)	1.00	-
≥1000 mL	82 (9.1)	72 (8.0)	1.14 (0.84 to 1.54)	0.40
1000-1499 mL	34 (3.8)	35 (3.9)	0.97 (0.61 to 1.54)	0.91
1500-1999 mL	21 (2.3)	14 (1.6)	1.50 (0.77 to 2.93)	0.23
≥2000 mL	27 (3.0)	23 (2.6)	1.18 (0.68 to 2.04)	0.56
Median (interquartile range) postpartum blood loss (mL)	300 (200-500)	300 (250-500)	-	0.18†
Transfusion (packed cells or plasma)	23 (2.6)	17 (1.9)	1.35 (0.73 to 2.52)	0.34
Manual removal placenta	41/803 (5.1)	33/804 (4.1)	1.24 (0.79 to 1.95)	0.34
Perineal tear:	n=803	n=804		
Episiotomy (without tear)	234 (29.3)	246 (30.6)	0.95 (0.82 to 1.11)	0.52
Obstetrical anal sphincter injuries	28 (3.5)	31 (3.9)	0.90 (0.55 to 1.49)	0.69
Third degree tear	15 (1.9)	19 (2.4)	0.79 (0.40 to 1.54)	0.49
Fourth degree tear	8 (1.0)	7 (0.9)	1.14 (0.41 to 3.14)	0.80‡
Episiotomy and third degree tear	4 (0.5)	2 (0.3)	2.00 (0.37 to 10.90)	0.45‡
Episiotomy and fourth degree tear	1 (0.1)	3 (0.4)	0.33 (0.03 to 3.20)	0.62‡
Maternal admission (highest level of care):				
Intensive care	3 (0.3)	2 (0.2)	1.50 (0.25 to 8.97)	0.66‡
Medium care	5 (0.6)	5 (0.6)	1.00 (0.29 to 3.45)	1.00‡
Ward	271 (30.1)	277 (30.7)	0.98 (0.85 to 1.13)	0.77
Indications for maternal admission:				
Thromboembolic complications	0 (0.0)	0 (0.0)	NA	-
Hypertensive disorders§	5 (0.6)	14 (1.6)	0.36 (0.13 to 0.99)	0.06‡
Postpartum blood loss	49 (5.5)	52 (5.8)	1.06 (0.72 to 1.54)	0.78
Post-caesarean section	97 (10.8)	97 (10.8)	1.00 (0.77 to 1.31)	0.99
Pain treatment during labour:	420 (46.7)	386 (42.8)	1.09 (0.98 to 1.20)	0.10
Remifentanil	128 (14.2)	129 (14.3)	0.99 (0.79 to 1.25)	0.95
Pethidine/promethazine/other opiates	60 (6.7)	51 (5.7)	1.18 (0.82 to 1.69)	0.38
Epidural anaesthesia	265 (29.4)	231 (25.6)	1.15 (0.99 to 1.33)	0.07
Other	1 (0.1)	5 (0.6)	0.20 (0.02 to 1.71)	0.22‡

NA=not applicable.
*Defined as postpartum haemorrhage ≥1000 mL, and/or manual removal of placenta, and/or third or fourth degree tears (obstetrical anal sphincter injuries), and/or intensive care admission, and/or maternal death. Denominator for perineal tear are vaginal deliveries only.
†Mann-Whitney U test.
‡Fisher's exact test.
§Including pre-eclampsia, and HELLP syndrome.

before induction started, compared with 73.7% in the expectant management group. Despite women with suspected or established intrauterine growth restriction being ineligible for inclusion in the study, the birthweight for 7% of the children was less than the 10th Dutch centile (61/900 induction and 62/901 expectant management), confirming the difficulty in diagnosing growth restricted babies at term. In the induction group, 2/61 infants had a birthweight less than the 10th centile and an adverse perinatal outcome: one neonate, weighing 3100 g (<10th centile), had an Apgar score of 6 at five minutes after operative vaginal delivery by forceps because of fetal distress. The other neonate, weighing 2595 g (<2.3rd centile), had an Apgar score of 6 at five minutes after caesarean section because of fetal distress, with an umbilical cord pH of 6.87, possibly due to hypotension of the mother after epidural analgesia for pain relief or multiple entanglement of the umbilical cord. In the expectant management group, 3/62 infants weighed less than the 10th centile at birth and had an adverse perinatal outcome: one (birthweight 2945 g) was a stillbirth, one (2980 g) was admitted to the NICU because of a pneumothorax, and one (3040 g) had an Apgar score of 6 at five minutes that was attributed to pethidine use in the mother.

Strengths and limitations of this study

A major strength of our study is that it concerns a nationwide multicentre randomised controlled trial of a well defined obstetrical population at low risk; the largest trial to date to compare induction of labour at 41 weeks with expectant management until 42 weeks.[46] No cases were lost to follow-up.

In the Netherlands, expectant management until 42 weeks is the standard of care in the low risk obstetrical population at 41-42 weeks according to the Dutch Obstetrical Indication List, although there is wide variation in practice because of women and caregiver preferences, which complicated inclusion.[15] Not all eligible women were invited, and not all women who were asked participated, because of a preference for induction or expectant management. Despite this selective participation, our trial offers the best possible representation of pregnant women reaching 41 weeks+0 days in the Netherlands.

We are aware of some potential limitations of our trial. We chose to use a composite adverse perinatal outcome instead of a single outcome like perinatal mortality. We considered any major adverse perinatal outcome in an otherwise uncomplicated pregnancy as undesirable. It is debatable if all the included adverse perinatal outcomes in our composite outcome measure are relevant to identify real severe adverse perinatal outcome with an effect on an infant's short term or long term health status. However, if we included an Apgar score of <4 instead of <7 at five minutes, according to the ACOG/AAP criteria, and excluded severe congenital abnormalities, induction of labour resulted in a statistically significant risk reduction of 0.9%, although with a substantially lower incidence of the composite adverse outcome in both groups.

We chose the non-inferiority design because we did not expect the Dutch standard policy of expectant management in our low risk obstetrical population to be inferior to a policy of induction of labour but acceptable or preferable if leading to comparable outcomes.[47] It is good practice to use a per protocol analysis in non-inferiority trials, as an intention-to-treat analysis carries a risk of falsely rejecting the null hypothesis of inferiority. Because we did not reject the null hypothesis and do not conclude non-inferiority, we presented the intention-to-treat analyses first, since such analyses are more common in reports of clinical trials. We also reported the per protocol outcome of the primary outcome (see supplementary appendix for the other per protocol analyses).

We did not stratify randomisation by parity, because we expected a balanced allocation in both groups owing to the large study population. However, it did result in an imbalance between groups: 50.8% of nulliparous women in the induction group compared with 56.7% in the expectant management group. After stratifying by parity in an additional analysis, we observed similar results. A higher incidence of the composite adverse perinatal outcome was seen in the nulliparity group in both the induction group (nulliparous 2.4% v multiparous 0.9%) and the expectant management group (nulliparous 4.1% v multiparous 1.8%), which is in concordance with other studies.[48] Furthermore, we saw no interaction between parity and induction of labour or expectant management in logistic regression analysis.

The measurement of arterial pH is not possible in primary care, and pH measurement is no standard policy for uncomplicated birth in most hospitals in the Netherlands. Because of the high number of missing pH measurements (60-70%) and the impossibility to impute, we could not include umbilical artery pH in the composite outcome, which could have led to selection bias. Including the available data on umbilical arterial pH in the analyses, however, did not alter the results.

The results of our study can be interpreted in different ways, which might have implications for standard practice. If the composite outcome is interpreted straightforwardly, there is a small benefit of induction at 41 weeks that could justify standard induction at 41 weeks. It could be argued, however, that a change of policy to earlier induction, concerning roughly one fifth of all women with a singleton pregnancy, is too rigorous in light of the relatively low incidence of perinatal mortality, gestational age associated NICU admission, and Apgar score <4 at five minutes as indicator for encephalopathy. This could justify expectant management if women want to avoid induction. On both sides of the spectrum, caregivers are challenged to provide neutral, evidence based counselling of low risk women in late term pregnancy on the pros and cons of induction. In a recent report by Walsh et al, women felt they were not offered a real choice when it came to management of their prolonged pregnancy, and this is confirmed by other studies; induction of

labour is often presented as an inevitable next step to women, without information provided on alternative management strategies.[49-51]

Conclusions and policy implications

Our large trial compared induction of labour at 41 weeks with expectant management until 42 weeks and subsequent induction if necessary. Substantial larger trials are needed to evaluate differences in rare outcomes, such as perinatal mortality and NICU admission. A systematic review or individual participant meta-analysis on the comparison between 41 weeks and 42 weeks could then be performed including findings from those studies as well as those of our own study. Future research could also focus on long term adverse perinatal outcome of both strategies, although this requires long term follow-up of children.[46] In addition, a more tailored approach will need identification of women who could maintain pregnancy until 42 weeks or are at increased risk of adverse perinatal outcomes (eg, relational model).

The incidence of late term pregnancy varies between countries because of different management strategies.[52] Women need to be counselled on the desired policy in late term pregnancy. In this trial, induction of labour at 41 weeks resulted in less overall adverse perinatal outcome than a policy of expectant management until 42 weeks, although the absolute risk of severe adverse outcome (perinatal mortality, NICU admission, Apgar score <4 at five minutes) was low in both groups. As with every intervention in the natural birth process, the decision to induce labour must be made with caution, as the expected benefits should outweigh possible adverse effects for both mother and child.[53] The results of our study should be used to inform women approaching a gestational age of 41 weeks, so they can weigh the respective outcomes and decide whether to be induced at 41 weeks or to continue pregnancy until 42 weeks.

1 WHO Collaborating Centers. International classification of diseases. *Epidemiol Bull* 1998;19:7-11.

2 Hilder L, Costeloe K, Thilaganathan B. Prolonged pregnancy: evaluating gestation-specific risks of fetal and infant mortality. *Br J Obstet Gynaecol* 1998;105:169-73. doi:10.1111/j.1471-0528.1998.tb10047.x

3 Ingemarsson I, Källén K. Stillbirths and rate of neonatal deaths in 76,761 postterm pregnancies in Sweden, 1982-1991: a register study. *Acta Obstet Gynecol Scand* 1997;76:658-62. doi:10.3109/00016349709024606

4 Nakling J, Backe B. Pregnancy risk increases from 41 weeks of gestation. *Acta Obstet Gynecol Scand* 2006;85:663-8. doi:10.1080/00016340500543733

5 Chantry AA, Lopez E. [Fetal and neonatal complications related to prolonged pregnancy]. *J Gynecol Obstet Biol Reprod (Paris)* 2011;40:717-25. doi:10.1016/j.jgyn.2011.09.007

6 Chantry AA. [Epidemiology of prolonged pregnancy: incidence and maternal morbidity]. *J Gynecol Obstet Biol Reprod (Paris)* 2011;40:709-16. doi:10.1016/j.jgyn.2011.09.006

7 De Los Santos-Garate AM, Villa-Guillen M, Villanueva-García D, Vallejos-Ruíz ML, Murguía-Peniche MTNEOSANO's Network. Perinatal morbidity and mortality in late-term and post-term pregnancy. NEOSANO perinatal network's experience in Mexico. *J Perinatol* 2011;31:789-93. doi:10.1038/jp.2011.43

8 Hermus MAA, Verhoeven CJM, Mol BW, de Wolf GS, Fiedeldeij CA. Comparison of induction of labour and expectant management in postterm pregnancy: a matched cohort study. *J Midwifery Womens Health* 2009;54:351-6. doi:10.1016/j.jmwh.2008.12.011

9 ACOG Committee Opinion No 579: Definition of term pregnancy. *Obstet Gynecol* 2013;122:1139-40. doi:10.1097/01. AOG.0000437385.88715.4a

10 Vayssière C, Haumonte JB, Chantry A, et al, French College of Gynecologists and Obstetricians (CNGOF). Prolonged and post-term pregnancies: guidelines for clinical practice from the French College of Gynecologists and Obstetricians (CNGOF). *Eur J Obstet Gynecol Reprod Biol* 2013;169:10-6. doi:10.1016/j.ejogrb.2013.01.026

11 American College of Obstetricians and Gynecologists. Practice bulletin no. 146: Management of late-term and postterm pregnancies. *Obstet Gynecol* 2014;124:390-6. doi:10.1097/01. AOG.0000452744.06088.48

12 Mandruzzato G, Alfirevic Z, Chervenak F, et al, World Association of Perinatal Medicine. Guidelines for the management of postterm pregnancy. *J Perinat Med* 2010;38:111-9. doi:10.1515/jpm.2010.057

13 Induction of labour: new NICE quality standard. *Midwives* 2014;17:8.

14 Workgroup. South Australian Perinatal Practice Guidelines (SAPPG). Clinical Guideline Prolonged Pregnancy, 2017.

15 Nederlandse Vereniging voor Obstetrie & Gynaecologie. *Guideline Prolonged Pregnancy.* Seroniteit, 2007.

16 Gülmezoglu AM, Crowther CA, Middleton P, Heatley E. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev* 2012;6:CD004945. doi:0.1002/14651858.CD004945.pub3

17 Hussain AA, Yakoob MY, Imdad A, Bhutta ZA. Elective induction for pregnancies at or beyond 41 weeks of gestation and its impact on stillbirths: a systematic review with meta-analysis. *BMC Public Health* 2011;11(Suppl 3):S5. doi:10.1186/1471-2458-11-S3-S5

18 Cheng YW, Nicholson JM, Nakagawa S, Bruckner TA, Washington AE, Caughey AB. Perinatal outcomes in low-risk term pregnancies: do they differ by week of gestation? *Am J Obstet Gynecol* 2008;199:370. e1-7.

19 Middleton P, Shepherd E, Crowther CA. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev* 2018;5:CD004945.

20 Wennerholm UB, Hagberg H, Brorsson B, Bergh C. Induction of labor versus expectant management for post-date pregnancy: is there sufficient evidence for a change in clinical practice?*Acta Obstet Gynecol Scand* 2009;88:6-17. doi:10.1080/00016340802555948

21 Keulen JKJ, Bruinsma A, Kortekaas JC, van Dillen J, van der Post JAM, de Miranda E. Timing induction of labour at 41 or 42 weeks? A closer look at time frames of comparison: A review. *Midwifery* 2018;66:111-8. doi:10.1016/j.midw.2018.07.011

22 Svensk Förening för Obstetric och Gynekologi. Induktion av förlossning. In: Riktlinje, ed., 2016.

23 National Institute for Health and Clinical Excellence. Guidance. Induction of Labour. London 2008. ISBN: 9781904752486.

24 de Jonge A, Geerts CC, van der Goes BY, Mol BW, Buitendijk SE, Nijhuis JG. Perinatal mortality and morbidity up to 28 days after birth among 743 070 low-risk planned home and hospital births: a cohort study based on three merged national perinatal databases. *BJOG* 2015;122:720-8. doi:10.1111/1471-0528.13084

25 de Jonge A, Mesman JA, Manniën J, et al. Severe Adverse Maternal Outcomes among Women in Midwife-Led versus Obstetrician-Led Care at the Onset of Labour in the Netherlands: A Nationwide Cohort Study. *PLoS One* 2015;10:e0126266. doi:10.1371/journal.pone.0126266

26 de Jonge A, Mesman JA, Manniën J, Zwart JJ, van Dillen J, van Roosmalen J. Severe adverse maternal outcomes among low risk women with planned home versus hospital births in the Netherlands: nationwide cohort study. *BMJ* 2013;346:f3263. doi:10.1136/bmj.f3263

27 de Jonge A, Peters L, Geerts CC, et al. Mode of birth and medical interventions among women at low risk of complications: A cross-national comparison of birth settings in England and the Netherlands. *PLoS One* 2017;12:e0180846. doi:10.1371/journal.pone.0180846

28 de Jonge A, van der Goes BY, Ravelli AC, et al. Perinatal mortality and morbidity in a nationwide cohort of 529,688 low-risk planned home and hospital births. *BJOG* 2009;116:1177-84. doi:10.1111/j.1471-0528.2009.02175.x

29 Amelink-Verburg MP, Verloove-Vanhorick SP, Hakkenberg RM, Veldhuijzen IM, Bennebroek Gravenhorst J, Buitendijk SE. Evaluation of 280,000 cases in Dutch midwifery practices: a descriptive study. *BJOG* 2008;115:570-8. doi:10.1111/j.1471-0528.2007.01580.x

30 Martijn LL, Jacobs AJ, Maassen II, Buitendijk SS, Wensing MM. Patient safety in midwifery-led care in the Netherlands. *Midwifery* 2013;29:60-6. doi:10.1016/j.midw.2011.10.013

31 Kortekaas JC, Bruinsma A, Keulen JK, et al. Effects of induction of labour versus expectant management in women with impending post-term pregnancies: the 41 week - 42 week dilemma. *BMC Pregnancy Childbirth* 2014;14:350. doi:10.1186/1471-2393-14-350

32 Committee on Obstetric Practice, ACOGAmerican Academy of PediatricsCommittee on Fetus and Newborn, ACOG. ACOG Committee Opinion. Number 644, Oct 2015 (replaces No. 333, May 2006): The Apgar score. *Obstet Gynecol* 2015;126(4):E52-E55.

33 Gülmezoglu AM, Crowther CA, Middleton P, Heatley E. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev* 2012;6:CD004945.

34 Bodner-Adler B, Bodner K, Kimberger O, Lozanov P, Husslein P, Mayerhofer K. Influence of the birth attendant on maternal and neonatal outcomes during normal vaginal delivery: a comparison between midwife and physician management. *Wien Klin Wochenschr* 2004;116:379-84. doi:10.1007/BF03040917

35 Bodner-Adler B, Kimberger O, Griebaum J, Husslein P, Bodner K. A ten-year study of midwife-led care at an Austrian tertiary care center: a retrospective analysis with special consideration of perineal trauma. *BMC Pregnancy Childbirth* 2017;17:357. doi:10.1186/s12884-017-1544-9

36 Grünebaum A, McCullough LB, Brent RL, Arabin B, Levene MI, Chervenak FA. Justified skepticism about Apgar scoring in out-of-hospital birth settings. *J Perinat Med* 2015;43:455-60. doi:10.1515/jpm-2014-0003

37 Ten Eikelder ML, Oude Rengerink K, Jozwiak M, et al. Induction of labour at term with oral misoprostol versus a Foley catheter (PROBAAT-II): a multicentre randomised controlled non-inferiority trial. *Lancet* 2016;387:1619-28. doi:10.1016/S0140-6736(16)00084-2

38 Koopmans CM, Bijlenga D, Groen H, et al, HYPITAT study group. Induction of labour versus expectant monitoring for gestational hypertension or mild pre-eclampsia after 36 weeks' gestation (HYPITAT): a multicentre, open-label randomised controlled trial. *Lancet* 2009;374:979-88. doi:10.1016/S0140-6736(09)60736-4

39 Gelisen O, Caliskan E, Dilbaz S, et al. Induction of labor with three different techniques at 41 weeks of gestation or spontaneous follow-up until 42 weeks in women with definitely unfavorable cervical scores. *Eur J Obstet Gynecol Reprod Biol* 2005;120:164-9. doi:10.1016/j.ejogrb.2004.08.013

40 Hannah ME, Hannah WJ, Hellmann J, Hewson S, Milner R, Willan AThe Canadian Multicenter Post-term Pregnancy Trial Group. Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. A randomized controlled trial. *N Engl J Med* 1992;326:1587-92. doi:10.1056/NEJM199206113262402

41 Burgos J, Rodríguez L, Otero B, et al. Induction at 41 weeks increases the risk of caesarean section in a hospital with a low rate of caesarean sections. *J Matern Fetal Neonatal Med* 2012;25:1716-8. doi:10.3109/14767058.2012.663018

42 Roos N, Sahlin L, Ekman-Ordeberg G, Kieler H, Stephansson O. Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. *Acta Obstet Gynecol Scand* 2010;89:1003-10. doi:10.3109/00016349.2010.500009

43 Mahomed K, Pungsornruk K, Gibbons K. Induction of labour for postdates in nulliparous women with uncomplicated pregnancy - is the caesarean section rate really lower?*J Obstet Gynaecol* 2016;36:916-20. doi:10.1080/01443615.2016.1174824

44 Grobman WA, Rice MM, Reddy UM, et al, Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal–Fetal Medicine Units Network. Labor Induction versus Expectant Management in Low-Risk Nulliparous Women. *N Engl J Med* 2018;379:513-23. doi:10.1056/NEJMoa1800566

45 Ravelli AC, Schaaf JM, Eskes M, Abu-Hanna A, de Miranda E, Mol BW. Ethnic disparities in perinatal mortality at 40 and 41 weeks of gestation. *J Perinat Med* 2013;41:381-8. doi:10.1515/jpm-2012-0228

46 Elden H, Hagberg H, Wessberg A, et al. Study protocol of SWEPIS a Swedish multicentre register based randomised controlled trial to compare induction of labour at 41 completed gestational weeks versus expectant management and induction at 42 completed gestational weeks. *BMC Pregnancy Childbirth* 2016;16:49. doi:10.1186/s12884-016-0836-9

47 *Perinatal Registration Netherlands Yearbook* , 2011.

48 Lindegren L, Stuart A, Herbst A, Källén K. Improved neonatal outcome after active management of prolonged pregnancies beyond 41^{+2} weeks in nulliparous, but not among multiparous women. *Acta Obstet Gynecol Scand* 2017;96:1467-74. doi:10.1111/aogs.13237

49 Walsh DRJ, Khan K, et al. *NHS Research Authority, University of Nottingham. Women's experiences of being 'overdue' in pregnancy.* NHS Research Authority, University of Nottingham, 2017.

50 Moore JE, Low LK, Titler MG, Dalton VK, Sampselle CM. Moving toward patient-centered care: Women's decisions, perceptions, and experiences of the induction of labor process. *Birth* 2014;41:138-46. doi:10.1111/birt.12080

51 Stevens G, Miller YD. Overdue choices: how information and role in decision-making influence women's preferences for induction for prolonged pregnancy. *Birth* 2012;39:248-57. doi:10.1111/j.1523-536X.2012.00554.x

52 Zeitlin J, Blondel B, Alexander S, Bréart GPERISTAT Group. Variation in rates of postterm birth in Europe: reality or artefact?*BJOG* 2007;114:1097-103. doi:10.1111/j.1471-0528.2007.01328.x

53 WHO recommendations: Intrapartum care for a positive childbirth experience: World Health Organization 2018.

Supplementary information: Perinatal, delivery, and adverse maternal outcomes in per protocol population