# Cost-effectiveness analysis of induction of labour at 41 weeks and expectant management until 42 weeks in low risk women (INDEX trial)

Aafke Bruinsma [a,b,*], Judit KJ Keulen [a,c], Rik van Eekelen [d], Madelon van Wely [d], Joep C Kortekaas [e], Jeroen van Dillen [f], Joris AM van de Post [a], Ben W Mol [g,h], Esteriek de Miranda [a]

[a] Amsterdam UMC, University of Amsterdam, Department of Obstetrics and Gynaecology, Amsterdam Reproduction & Development Research Institute, Meibergdreef 9, Amsterdam, the Netherlands
[b] School of Health Care Studies of Rotterdam University of Applied Sciences, Midwifery Academy Rotterdam, Rotterdam, the Netherlands
[c] Zuyd University, Research Center for Midwifery Science, Faculty Midwifery Education & Studies Maastricht, the Netherlands
[d] Amsterdam UMC, University of Amsterdam, Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Meibergdreef 9, Amsterdam, the Netherlands
[e] Elkerliek Medical Center, Department of Obstetrics & Gynaecology, Helmond, the Netherlands
[f] Radboud University Medical Center, Nijmegen, Department of Obstetrics & Gynaecology, Nijmegen, the Netherlands
[g] Monash University, Department of Obstetrics and Gynaecology, Clayton, Victoria, Australia
[h] University of Aberdeen, Aberdeen Centre for Women's Health Research, School of Medicine, Medical Sciences and Nutrition, Aberdeen, UK

## ARTICLE INFO

## ABSTRACT

*Objective:* To assess the cost-effectiveness of elective induction of labour (IOL) at 41 weeks and expectant management (EM) until 42 weeks.

*Design:* Cost-effectiveness analysis from a healthcare perspective alongside a randomised controlled trial (INDEX).

*Setting:* 123 primary care midwifery practices and 45 obstetric departments of hospitals in the Netherlands.

*Population:* We studied 1801 low-risk women with late-term pregnancy, randomised to IOL at 41 weeks (N = 900) or EM until 42 weeks (N = 901).

*Methods:* The incremental cost-effectiveness ratio (ICER) was expressed as the ratio of the difference in costs and the difference in main perinatal outcomes. A Cost-Effectiveness Acceptability Curve (CEAC) was constructed to assess whether induction is cost-effective for a range of monetary values as thresholds. We performed subgroup analysis for parity.

*Main outcome measures:* Direct medical costs, composite adverse perinatal outcome (CAPO) (perinatal mortality, NICU admission, Apgar 5 min < 7, plexus brachialis injury and/or meconium aspiration syndrome) and composite severe adverse perinatal outcome (SAPO) (including Apgar 5 min < 4 instead of < 7).

*Results:* The average costs were €3858 in the induction group and €3723 in the expectant group (mean difference €135; 95 % CI −235 to 493). The ICERs of IOL compared to EM to prevent one additional CAPO and SAPO was €9436 and €14,994, respectively. The CEAC showed a 80 % chance of IOL being cost-effective with a willingness-to-pay of €22,000 for prevention of one CAPO and €50,000 for one SAPO. Subgroup analysis showed a willingness-to-pay to prevent one CAPO for nulliparous of €47,000 and for multiparous €190,000. To prevent one SAPO the willingness-to-pay is €62,000 for nulliparous and €970,000 for multiparous women.

*Conclusions:* Induction at 41 weeks has an 80 % chance of being cost-effective at a willingness-to-pay of €22,000 for prevention of one CAPO and €50,000 for prevention of one SAPO. Subgroup analysis suggests that induction could be cost-effective for nulliparous women while it is unlikely cost-effective for multiparous women. Cost-effectiveness in other settings will depend on baseline characteristics of the population and health system organisation and funding.

* Corresponding author at: Amsterdam UMC, University of Amsterdam, Department of Obstetrics and Gynaecology, Amsterdam Reproduction & Development Research Institute, Meibergdreef 9, Amsterdam, the Netherlands.
E-mail address: a.bruinsma@amsterdamumc.nl (A. Bruinsma).

## 1. Introduction

Induction of labour is a common intervention in obstetric care for several reasons, with late-term pregnancy — $41^{0/7}$ through $41^{6/7}$ weeks of gestation — being a frequently used indication worldwide [1–4].

In 2019, two trials were published — INDEX (N = 1801) and SWEPIS (N = 2760) — comparing induction at 41 weeks with expectant management (EM) until 42 weeks on adverse perinatal outcome [3,5]. The results have been examined in an individual participant data meta-analysis (IPD-MA). It showed that elective induction at 41 weeks reduced the risk of adverse perinatal outcomes compared to expectant management until 42 weeks (IOL 0.4 % versus EM 1.0 %; RR 0.43; 95 % CI 0.21–0.91; *p*-value 0.027). This lower risk applied to nulliparous women (RR 0.20; 95 % CI 0.07–0.60) but not to multiparous (RR 1.93; 95 % CI 0.48–7.72). Induction did not increase the caesarean section rate (IOL 10.5 % versus EM 10.7 %; RR 0.98; 95 % CI 0.83–1.16) [6].

The INDEX trial was a randomised controlled trial comparing elective induction at 41 weeks with expectant management until 42 weeks and showed a significant reduction of the risk of adverse perinatal outcomes in favour of induction (IOL 1.7 % versus EM 3.1 %; RR 0.54; 95 % CI 0.29–1.00; *p*-value 0.045). For composite *severe* adverse perinatal outcome risks were 0.4 % after induction versus 1.3 % after expectant management (RR 0.3; 95 % CI 0.11–1.03) [3]. The INDEX cohort study — alongside the INDEX trial — found also lower incidences of adverse perinatal outcomes (IOL 1.1 % versus EM 1.9 %; adjRR 0.56; 95 % CI 0.17–1.79)) and *severe* adverse perinatal outcome (IOL 0.3 % versus EM 1.0 %; adjRR 0.39; 95 % CI 0.05–2.88)) than in the trial [7].

Studies on costs of induction of labour showed that antepartum costs were lower and intrapartum costs were higher in the induction group. These studies only examined costs, without cost-effectiveness analyses [8,9]. The SWEPIS cost-effectiveness analysis in late-term pregnancy concluded that there is no difference in costs between the two groups while induction of labour at 41 weeks is more effective and therefore cost-effective compared to expectant management [10].

In addition to clinical outcomes, patient's preference and experience as well as cost are increasingly considered as relevant components in policy making [11,12]. We analysed the cost-effectiveness — from a healthcare perspective — of induction of labour at 41 weeks and expectant management until 42 weeks in a low-risk population alongside the INDEX RCT.

## 2. Materials and methods

### 2.1. Study design

This economic evaluation was performed alongside the INDEX trial. The study protocol and the clinical results of this trial have been published previously [3,13]. In the period 2012–2016 the trial was conducted in 123 primary care midwifery practices and 45 obstetric departments of hospitals in the Netherlands and was approved by the local ethics committee of the Academic Medical Centre, Amsterdam (No NL38455.018.11).

### 2.2. Study population

The INDEX trial included low-risk women with an uncomplicated pregnancy of a singleton in stable cephalic position and a certain gestational age between $40^{5/7}$ and $41^{1/7}$ weeks without a contraindication for expectant management until $42^{0/7}$ weeks at time of randomisation. Gestational age was based on ultrasonography before a gestational age of 16 weeks. Exclusion criteria were age younger than 18 years, ruptured membranes and/or in labour, non-reassuring foetal status (e.g., abnormal foetal heart rate, or no foetal movements and/or expected intrauterine growth restriction), known foetal abnormalities (including abnormal karyotype) that could influence perinatal outcome, contraindications to induction (including previous caesarean section),

or contraindications to expectant management (e.g., pregnancy induced hypertension). In the Netherlands, obstetric care is provided by primary care (midwives) for low-risk women and by secondary care (clinical midwives, residents and gynaecologists) for women at increased risk of adverse maternal or perinatal outcomes, or both.

Women were randomised for either elective induction of labour at 41 weeks or expectant management. Women randomised for induction at 41 weeks underwent an induction of labour at $41^{0/7}$ or $41^{1/7}$ weeks. Women randomised for expectant management awaited spontaneous onset of labour with subsequent induction if necessary at $42^{0/7}$ weeks. Both groups received obstetrical care according to local protocol. The baseline characteristics of the study population are showed in Table 1 of the trial publication [3].

### 2.3. Outcomes of interest

#### 2.3.1. Primary outcomes

The primary outcome was a composite of adverse perinatal outcomes (CAPO) and was defined as perinatal mortality (until 28 day postpartum) and/or Apgar score < 7 at 5 min and/or arterial umbilical cord pH < 7.05 and/or meconium aspiration syndrome (MAS) and/or plexus brachialis injury and/or intracranial haemorrhage and/or being admitted to a neonatal intensive care unit (NICU). The co-primary outcome composite severe adverse perinatal outcome (SAPO) was comparable with the primary CAPO except that Apgar score at 5 min < 7 was replaced by Apgar score 5 min < 4 [14].

#### 2.3.2. Secondary outcomes

Secondary outcomes included caesarean section and a composite of adverse maternal outcome (CAMO). Composite of adverse maternal outcome was defined as postpartum haemorrhage ≥ 1000 mL and/or manual removal of placenta (MRP) and/or third or fourth degree tears (obstetrical anal sphincter injuries) and/or intensive care admission and/or maternal death. Analyses are on individual level, though neonates and mothers could experience more than one adverse outcome.

Analyses of the trial were performed according to the intention-to-treat principle.

### 2.4. Economic evaluation

Outcome data were available for 1801 women, 900 in the induction of labour group and 901 in the expectant management group respectively. This economic analysis was designed as a cost-effectiveness analysis with composite adverse perinatal outcome as effectiveness measure. We actually used a hospital perspective rather than a social perspective. This means that only direct medical costs and effects up to discharge to home were included. Discounting of costs was unnecessary because all costs were calculated over one year (2021). In this study, the phrasing 'healthcare perspective' is more appropriate than 'hospital perspective', as both primary (out of hospital) and secondary obstetrician-led care were studied. This is the reason the term 'healthcare perspective' was used. This also contains the costs made in primary midwifery-led care, for example consultations in independent midwifery practice and home births.

#### 2.4.1. Resource use

Data was collected from the Case Report Form (CRF) and measured separately for antepartum, intrapartum and postpartum phases. For the antepartum phase we included the number of outpatient consultations (from obstetrics department and/or independent midwife) and the number of ultrasounds and CTGs performed. For the intrapartum phase we included cervical ripening material, oxytocin for induction and/or augmentation, medication use (antibiotics, tocolytics and analgesics), foetal blood sampling (FBS), mode of delivery, manual removal of placenta (MRP), repair of perineal tear in operating theatre and blood transfusion. Dates and times during labour of all participants were

**Table 1**

Cost-analysis: units of resource, unit costs, valuation method and volume source (2021 €).

| | Unit | Unit cost (€) | Valuation method |
|---|---|---|---|
| **Antenatal healthcare provider** | | | |
| Outpatient consultation (dep. obstetrics) * | Procedure | 101 | Dutch costing guideline [15] |
| Outpatient consultation (independent midwife) | Procedure | 38 | Dutch costing guideline [15] |
| CTG | Procedure | 37 | Probaat-2 study [19] |
| Ultrasound | Procedure | 40 | Dutch Healthcare Authority [16] |
| **Medication during labour** | | | |
| Cervical ripening | | | |
| Misoprostol (PGE1) | Gift | 8 | Dutch drug registry [17] |
| | Dose/day | 66 | Dutch drug registry [17] |
| Dinoprostone (PGE2) | Gift | 32 | Dutch drug registry [17] |
| Foley catheter | Unit | 18 | Top-down calculation |
| Cook catheter (double-balloon) | Unit | 37 | Liu et al. [18] |
| Oxytocin | Total costs labour | 3 | Dutch drug registry [17] |
| Antibiotics† | Treatment | 38 | Dutch drug registry [17] |
| Tocolytics† | Gift | 45 | Dutch drug registry [17] |
| Analgesics during labour | | | |
| Epidural | Procedure | 204 | Probaat-2 study [19] |
| Patient-controlled remifentanil | Procedure | 153 | Probaat-2 study [19] |
| Pethidine/Promethazine | Gift | 3 | Dutch drug registry [17] |
| **Delivery** | | | |
| Mode of delivery | | | |
| Spontaneously (Secondary Care) | Procedure | 1345 | Top-down calculation |
| Spontaneously (Primary Care) | Procedure | 651 | Dutch Healthcare Authority [16] |
| Operative vaginal delivery | Procedure | 1589 | Top-down calculation |
| Caesarean section | Procedure | 2372 | Top-down calculation |
| Hospital stay – ward before delivery * | Hour | 22 | Top-down calculation |
| Labour room * | Hour | 101 | Hypitat-II study [20] |
| Foetal blood sampling (FBS) | Procedure | 21 | Probaat-2 study [19] |
| Manual removal of placenta (MRP) | Procedure | 207 | Probaat-2 study [19] |
| Repair perineal tear in operating theatre | Procedure | 207 | Probaat-2 study [19] |
| Blood transfusion | Gift | | |
| Erythrocytes | Gift | 240 | Dutch costing guideline [15] |
| Fresh Frozen Plasma | Gift | 206 | Dutch costing guideline [15] |
| **Admission mother** | | | |
| Ward * | Day | 528 | Dutch costing guideline [15] |
| Medium care unit * | Day | 666 | Top-down calculation |
| Intensive care unit (ICU)*‡ | Day | 2235 | Dutch costing guideline [15] |
| Ambulance transport | Ride | 302 | Dutch costing guideline [15] |
| **Admission child** | | | |
| Maternal ward * | Day | 528 | Dutch costing guideline [15] |
| Medium care unit (MC)* | Day | 666 | Top-down calculation |
| Neonatal intensive care unit (NICU)*‡ | Day | 1847 | Top-down calculation |
| Ambulance transport | Ride | 302 | Dutch costing guideline [15] |

*Mean unit cost for a general and an academic hospital.

†The mean of several methods/medication is presented.

‡Including diagnostics, medication and specialist (intensivist, gynaecologist and/or neonatologist.)

structurally recorded in the CRF. The items noted were 'admission in hospital', 'admission to labour ward, 'start cervical ripening', 'start induction', 'rupture of membranes/amniotomy', 'start active phase (> 4 cm dilation)' and 'birth'. Time spent in the labour room was calculated as the interval between admission to the labour room and birth, plus two hours extra for recovery care. The costs of a spontaneous vaginal delivery, operative vaginal delivery and/or caesarean section were calculated. Postpartum phase included neonatal and maternal admission, both the ward (maternal ward, medium care, intensive care), length of stay (in days) and ambulance transport to another hospital. Only costs of mothers and neonates with a medical indication for admission were calculated.

### 2.4.2. Unit costs

Unit costs were estimated with different methods and sources, all according to recent guidelines on costing of healthcare services (Table 1) [15–20]. All costs were expressed in Euros (2021 value) and inflated where appropriate using the consumer pricing index [21].

Unit cost were provided by the hospital conducting the trial (Amsterdam UMC). These were collected in the research consortium context and made available to researchers. Mean costs from academic and non-academic hospitals were used. Medication prices — including

misoprostol, dinoprostone, oxytocin, antibiotics, tocolytics and pethidine/promethazine — were obtained from the Dutch drug registry [17]. For a number of variables — CTG, double-balloon catheter, epidural, patient-controlled remifentanil, labour room, FBS, MRP and perineal tear repair in operating theatre — mean costs from other Dutch studies were used and indexed, as these could not be extracted from the Dutch costing guideline. This allowed conduction of a detailed costing [18–20].

Obstetric procedures were counted and valued separately to allow differentiation in associated costs between both management groups. These included analgesics during labour (epidural, patient-controlled remifentanil, and pethidine/promethazine), mode of delivery (spontaneously in primary/secondary care, operative vaginal delivery, caesarean section), labour room in hours, FBS, MRP, repair perineal tear in operating theatre and blood transfusion. Potential costs for neonatal admission to the maternity ward, only because of a maternal medical indication for admission, were not counted.

### 2.5. Statistical analysis

All analyses were performed according to the intention-to-treat principle, following the allocation of treatment in the INDEX trial. The

incremental cost-effectiveness ratio (ICER) was calculated by dividing the difference in mean total costs of induction of labour and expectant management by the difference in chances of composite adverse perinatal outcomes or composite severe adverse perinatal outcomes. We boot-strapped the costs and effects for 5000 samples with replacement to obtain the 95 % confidence intervals (95 % CI) of costs and effects. Results from bootstrapping in the cost-effectiveness plane with the four quadrants representing the four possible conclusions: 1) induction is more expensive and more effective (north-west), 2) induction is less expensive and more effective (south-west), 3) induction is more expensive and less effective (north-east), and 4) induction is less expensive and less effective (south-east). As CAPO and SAPO are un-desirable outcomes, lower chances are better, as opposed to most cost-effectiveness planes.

Due to an overlap between several quadrants, a simple bootstrap 95 % CI for the ICER could not be interpreted. Therefore, we used the bootstrap samples to construct a Cost-Effectiveness Acceptability Curve (CEAC) with varying ICER cut-offs following Fenwick et al. [22]. In CEACs, instead of considering one value such as the ICER, a full range of monetary values from €0 to €40,000 is considered. The purpose of the CEAC was to show the proportion of bootstrap samples in which a certain ICER (or lower i.e. better) was found, thereby reflecting the uncertainty around cost-effectiveness. Thus, the ICER provides a point estimate for the costs to prevent one additional CAPO or SAPO and the CEAC visualizes uncertainty in the sense of the probability that, given a range of monetary thresholds, induction of labour is more cost-effective than expectant management.

The concept of 'willingness-to-pay' (WTP) is often used for trade-off in health policy making. The threshold is not a solid cut-off point but depends on how the value of health is considered and what good health may cost in this perspective. In health related analyses, a willingness-to-pay threshold is an estimate of what a healthcare consumer is willing to pay for the health benefit. The method of contingent valuation is widely used for valuing the willingness-to-pay. In this method, individuals are asked to compare different hypothetical situations about the intervention under study [23,24]. The National Institute for Health and Clinical

Excellence (NICE) has been using a cost-effectiveness threshold range between £ 20,000 and £ 30,000 ( ± €23.000 – €35.000) per QALY. In the Netherlands, thresholds are used based on The National Health Service (NHS)/NICE.

### 2.5.1. Scenario and subgroup analyses

We conducted a scenario and a subgroup analysis. For the scenario analysis, we took into account the additional costs for neonates who were admitted to the maternal ward. All other methods and statistics remained the same.

In addition, we conducted a subgroup analysis for parity, according to the higher incidences of CAPO and SAPO in nulliparous women after stratifying for parity in the trial.

Data formatting was performed using IBM SPSS Statistics for Windows (Version 26.0. Armonk, NY: IBM Corp. Released 2019). Data analysis and bootstrapping was performed using R version 3.6.0 and RStudio [25].

## 3. Results

### 3.1. Clinical results

In the induction group 15/900 (1.7 %) of the neonates had a CAPO compared to 28/901 (3.1 %) in expectant group (RR 0.54; 95 % CI 0.29–1.00; *p*-value 0.045; mean difference −1.4 % (95 % CI −2.8 to 0.0); numbers needed to treat (NNT) 69 (95 % CI 35 – 3059)). SAPO was observed in 0.4 % (4/900) of the induction group versus 1.3 % (12/901) of the expectant group (RR 0.33; 95 % CI 0.11–1.03; NNT 113 (95 % CI 57–4624)), with a mean difference of − 0.9 % (95 % CI −1.8 to 0.0).

Secondary outcomes showed no significant difference in both caesarean section rates, 10.8 % in each group (IOL 97/900; EM 97/901; RR 1.00; 95 % CI 0.77–1.31) and composite of adverse maternal outcome (IOL 122/900; EM 102/901; RR 1.20; 95 % CI 0.94–1.53). The median gestational age at time of delivery was $41^{0/7}$ weeks (inter-quartile range (IQR) $41^{0/7}$ to $41^{1/7}$ weeks) for the induction group and $41^{2/7}$ weeks (IQR $41^{0/7}$ to $41^{5/7}$ weeks) in the expectant management

**Table 2**
Mean cost per group (2021 €).

| Variable | Induction of labour (IOL) n = 900 Mean (€) | Expectant management (EM) n = 901 Mean (€) | Mean difference (€) (IOL – EM)* |
|---|---|---|---|
| *Antepartum costs* | | | |
| Antepartum outpatient consultation | 6 | 70 | -64 |
| CTG, ultrasound | 8 | 35 | -27 |
| **Total antepartum costs** | **14** | **105** | **-91** |
| *Intrapartum costs* | | | |
| Induction material | 10 | 4 | 6 |
| Hospital stay – ward before delivery | 82 | 32 | 50 |
| Admission labour ward | 1168 | 785 | 383 |
| Medication during labour (including FBS, excluding analgesics) | 7 | 6 | 1 |
| Analgesics | 82 | 74 | 8 |
| Delivery | 1367 | 1390 | -23 |
| Operative vaginal delivery (extra costs) | 25 | 29 | -4 |
| Caesarean section (extra costs) | 111 | 111 | 0 |
| MRP | 9 | 8 | 1 |
| Repair perineal tear in operating theatre | 6 | 7 | -1 |
| Blood transfusion | 19 | 16 | 3 |
| **Total intrapartum costs** | **2888** | **2462** | **426** |
| *Postpartum costs* | | | |
| Ambulance transfer (maternal/neonatal) | 2 | 2 | 0 |
| Maternal admission secondary care | 511 | 530 | -19 |
| Neonatal admission | 443 | 623 | -180 |
| **Total postpartum costs** | **956** | **1155** | **-199** |
| **Total costs (95 % CI)**[†] | **3858** | **3722** | **135 (−235 to 493)**[‡] |

*Cost of induction of labour group (IOL) minus cost of expectant management group (EM).

[†] CI confidence interval.

[‡] On a non-rounded average basis, so numbers of total costs do not sum up.

group [3].

### 3.2. Outcomes

#### 3.2.1. Direct medical costs

The average total costs were €3858 in the induction group and €3723 in the expectant management group, with a mean difference (MD) of €135 (95 % CI €−235 to €493). Table 2 shows more details on the costs per group.

The mean difference of antepartum care was -€91 (IOL €14; EM €105; 95 % CI €−100 to €−82; p-value < 0.001), caused by fewer antepartum outpatient consultations (including CTG and/or ultrasound) in the induction group compared to the expectant group. In the induction group 6.9 % of women (62/901) received an antepartum consultation. In the expectant group 62.6 % (564/901) received an antepartum consultation, of which 29.5 % (266/901) at the independent midwifery practice, 14.1 % (127/901) in a department of obstetrics and 19.0 % (171/901) at both caregivers. A total of 337 women (37.4 %) did not receive antepartum consultation. The majority of women (314/337) without an antepartum consultation after 41 weeks went to labour before consultation (usual $\leq 41^{3/7}$ weeks) could have taken place. In the induction group 6.9 % (62/900) had an antepartum consultation, of whom 6.1 % (55/900) $\leq 41^{3/7}$ weeks and 0.8 % (7/900) $\geq 41^{4/7}$ weeks. In the induction group 97.8 % (880/900) went into labour at or before $41^{3/7}$ weeks.

For 213 women in primary care time of birth was recorded in the
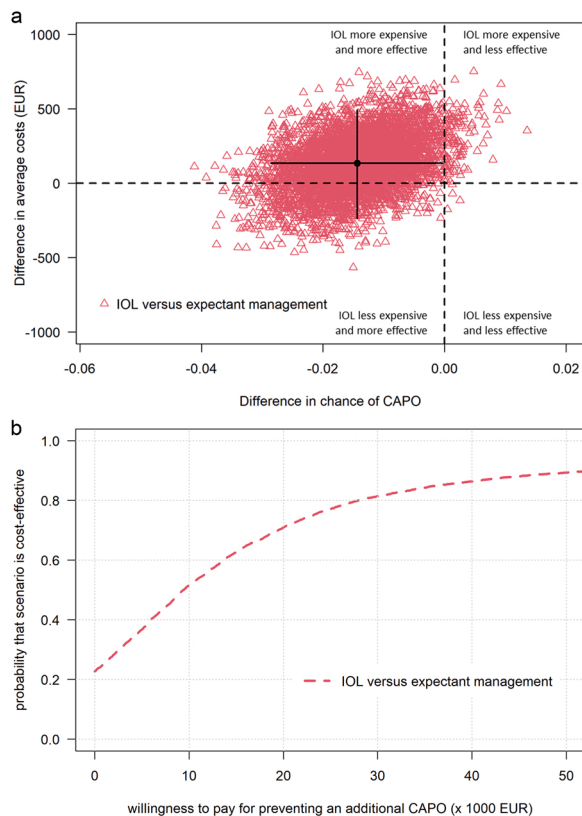
**Fig. 1. a** Cost-effectiveness plane composite adverse perinatal outcome (CAPO). Cost-effectiveness plane showing the costs and differences of bootstrap samples for **adverse perinatal outcome**. Each point in the cost-effectiveness plane represents the additional costs and health gain of induction of labour compared to expectant management. **b** Cost-effectiveness acceptability curve for composite adverse perinatal outcome (CAPO). Cost-effectiveness acceptability curve showing the proportion of bootstrap samples (y-axis) that were found cost-effective when compared to a range of threshold monetary values (x-axis) for **adverse perinatal outcome**.
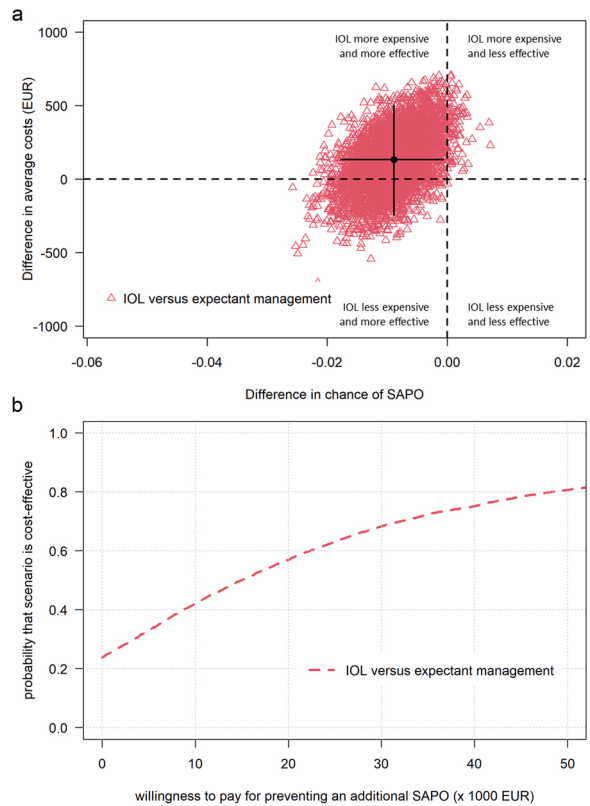
**Fig. 2. a** Cost-effectiveness plane composite severe adverse perinatal outcome (SAPO). Cost-effectiveness plane showing the costs and differences of bootstrap samples for **severe adverse perinatal outcome**. Each point in the cost-effectiveness plane represents the additional costs and health gain of induction of labour compared to expectant management. **b** Cost-effectiveness acceptability curve for composite severe adverse perinatal outcome (SAPO). Cost-effectiveness acceptability curve showing the proportion of bootstrap samples (y-axis) that were found cost-effective when compared to a range of threshold monetary values (x-axis) for **severe adverse perinatal outcome**.

CRF, but the woman concerned did not attend hospital. Which suggests that these women gave birth at home, case-level data confirms this.

During delivery the mean cost difference was €426 in favour of expectant management (IOL €2888; EM €2462; 95 % CI €294 to €557). This difference was mainly caused due to the length of hospital stay (in hours) for induction of labour (IOL 17.3 versus EM 11.6. hours), both for ward before delivery and admission to labour ward (MD €433).

The mean cost difference during the postpartum period was €199 in favour of induction of labour (IOL €956; EM €1155; 95 % CI €−510 to €112), mainly caused by the lower costs of neonatal admission in the induction group (MD €180; IOL €443; EM €623).

#### 3.2.2. Costs composite adverse perinatal outcome

For the composite adverse perinatal outcome (CAPO), the ICER was €−9436, meaning that by spending €9436 on induction of labour, we could prevent one additional CAPO as compared to expectant management. In Fig. 1a, we visualized the bootstrap samples in the cost-effectiveness plane. 75.1 % of samples appeared in the north-west (induction is more expensive and more effective), 22.6 % in the south-west (induction is less expensive and more effective), 2.2 % in the north-east (induction is more expensive and less effective) and 0.1 % in the south-east (induction is less expensive and less effective). Because samples appear in all four quadrants, the simple bootstrap 95 % CI of the ICER is uninterpretable [26].

For SAPO, the ICER was €− 14,994, meaning that by spending €14,994 on induction of labour, one additional SAPO could be prevented compared to expectant management. In Fig. 2a, we visualized the

**Table 3**
Scenario and subgroup analyses (€ 2021).

| Model | Description | IOL (€) | EM (€) | Mean difference (€) (IOL – EM) | 95 % CI |
|---|---|---|---|---|---|
| **Scenario analyses** | | | | | |
| 0 | Base Case | 3858 | 3723 | 135 | -234 to 505 |
| 1* | Additional costs for neonate on maternal ward | 4266 | 4130 | 136 | -285 to 555 |
| **Subgroup analyses** | | | | | |
| 2** | Nulliparous women (n = 968) | 4910 | 4700 | 210 | -367 to 787 |
| 3*** | Multiparous women (n = 833) | 2773 | 2442 | 331 | -48 to 710 |

*Scenario calculated in both arms (IOL and EM).

**Scenario base case calculated in both arms (IOL and EM), but only for nulliparous women.

***Scenario base case calculated in both arms (IOL and EM), but only for multiparous women.
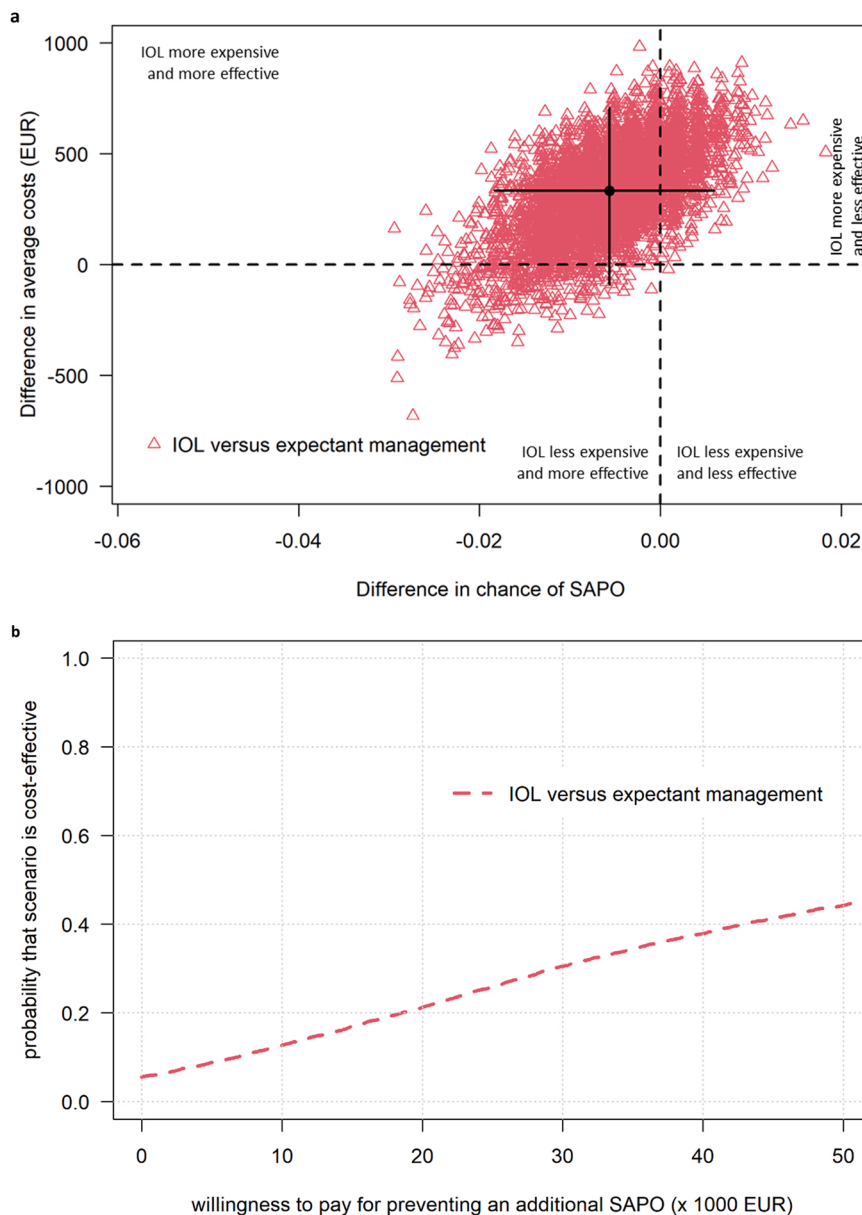
bootstrap samples in the cost-effectiveness plane. 74.5 % of samples appeared in the north-west (induction is more expensive and more effective), 23.8 % in the south-west (induction is less expensive and more effective), 1.7 % in the north-east (induction is more expensive and less effective) and 0.1 % in the south-east (induction is less expensive and less effective). Because samples also appear in all four quadrants,

the simple bootstrap 95 % CI of the ICER is uninterpretable.

Whether induction of labour at 41 weeks is considered cost-effective depends on the 'willingness-to-pay' for these health gains — less composite adverse and severe adverse perinatal outcome —. The CEAC in Fig. 1b shows the uncertainty surrounding the cost-effectiveness. We found that at a willingness-to-pay of approximately €22,000 for
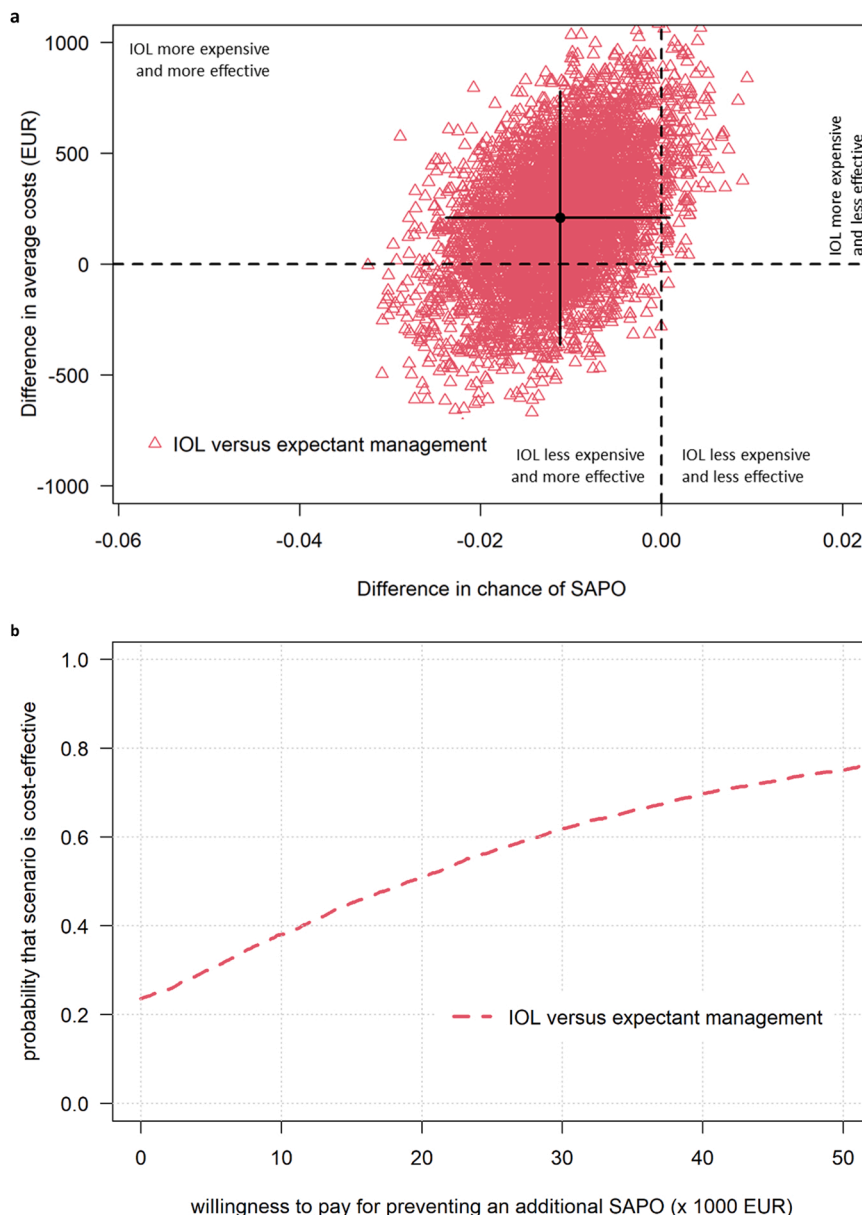


**Fig. 3.** **a** Cost-effectiveness plane composite severe adverse perinatal outcome (SAPO) in nulliparous women. Cost-effectiveness plane showing the costs and differences of bootstrap samples for **severe adverse perinatal outcome in nulliparous** women. Each point in the cost-effectiveness plane represents the additional costs and health gain of induction of labour compared to expectant management. **b** Cost-effectiveness acceptability curve for composite severe adverse perinatal outcome (SAPO) in nulliparous women. Cost-effectiveness acceptability curve showing the proportion of bootstrap samples (y-axis) that were found cost-effective when compared to a range of threshold monetary values (x-axis) for **severe adverse perinatal outcome in nulliparous** women.

**Fig. 4. a** Cost-effectiveness plane severe adverse perinatal outcome (SAPO) in multiparous women. Cost-effectiveness plane showing the costs and differences of bootstrap samples for **adverse perinatal outcome** in **multiparous** women. Each point in the cost-effectiveness plane represents the additional costs and health gain of induction of labour compared to expectant management. **b** Cost-effectiveness acceptability curve for severe adverse perinatal outcome (SAPO) in multiparous women. Cost-effectiveness acceptability curve showing the proportion of bootstrap samples (y-axis) that were found cost-effective when compared to a range of threshold monetary values (x-axis) for **severe adverse perinatal outcome** in **multiparous** women.

prevention of one CAPO, induction of labour has an 80 % chance of being cost-effective. For prevention of one SAPO a willingness-to-pay of €50,000 was found, also for 80 % chance of being cost-effective (Fig. 2b). Both CEACs visualize the increasing probability that induction is cost-effective for CAPO (Fig. 1b) and SAPO (Fig. 2b) when increasing the willingness-to-pay thresholds. Both CEACs never converge to 1.0 (100 %) because part of the bootstrap samples show that induction is 'never better', regardless of willingness-to-pay. This is because part of the bootstrap samples fall into the north-east (showing induction is more expensive and less effective; CAPO 2.2 %; SAPO 1.7 %) and south-east quadrants (showing induction is less expensive and less effective; CAPO 0.1 %; SAPO 0.1 %).

### 3.2.3. Scenario and subgroup analyses

Table 3 depicts the costs used in scenario analyses. For the scenario analysis, we accounted for additional costs for neonates admitted to the maternal ward. This analysis showed similar results as the primary analysis for both CAPO and composite severe adverse perinatal outcome (SAPO) (supplementary figures 5a, 5b, 6a and 6b).

For the subgroup analysis, we conducted a separate analysis of nulliparous (n = 968) and multiparous (n = 833) women. We found that induction of labour was more cost-effective for nulliparous women than for multiparous women: the ICER for one CAPO was €− 12,199 for nulliparous and €− 37,134 for multiparous women (supplementary figures 7a and 8a). For nulliparous women a willingness-to-pay of approximately €47,000 for prevention of one CAPO (supplementary figure 7b), induction has an 80 % chance of being cost-effective. For multiparous women this amount was €190,000 (supplementary figure 8b).

The ICER for one SAPO was €− 18,844 for nulliparous and €− 59,037 for multiparous women (Figs. 3a and 4a). For nulliparous women we found that at a willingness-to-pay of approximately €62,000 for prevention of one SAPO (Fig. 3b), induction has an 80 % chance of being cost-effective. For multiparous women this amount was €970,000 (Fig. 4b).

Note that these results were more uncertain due to the halving of sample size and that overall interpretations for both CAPO and SAPO were similar to the primary analysis.

## 4. Discussion

### 4.1. Main findings

This cost-effectiveness study alongside the INDEX-trial assessed the direct medical costs of induction of labour at 41 weeks compared with expectant management until 42 weeks. The costs in the induction group to prevent one CAPO was €9436 and to prevent one SAPO was €14,994. We found that at a willingness-to-pay of approximately €22,000 for prevention of one CAPO and approximately €50,000 for prevention of one SAPO, induction has an 80 % chance of being cost-effective. Induction of labour was more cost-effective for nulliparous women than for multiparous women.

The costs in the induction group to prevent one CAPO in nulliparous women was €12,199 and €37,134 for multiparous women. For nulliparous women a willingness-to-pay of approximately €47,000 for prevention of one CAPO, induction has an 80 % chance of being cost-effective. For multiparous women this amount was €190,000. To prevent one SAPO the ICERs for induction were €18,844 for nulliparous and €59,037 for multiparous women, with 80 % chance of being cost-effective at a willingness-to-pay of approximately €62,000 and €970,000 respectively.

After the start of the INDEX trial the American College of Obstetricians and Gynecologists and American Academy of Pediatrics (ACOG/AAP) committee released an update in 2015 on the use of the Apgar score. This update stated that the inappropriate use of the Apgar score ($< 7$ after 5 min) in outcome studies had led to an erroneous definition of asphyxia. Although it is incorrect to use Apgar score alone to diagnose birth asphyxia, an Apgar score $< 4$ after five minutes "may be considered a non-specific sign of illness, which may be one of the first indications of encephalopathy". It is not only associated with an increased risk of neonatal mortality but also with long-term adverse neonatal outcome [14]. Therefore a co-primary outcome measure (SAPO) was added in this study, including Apgar scores $< 4$ instead of $< 7$ after five minutes. In clinical practice and policy making we expect that the *severe* adverse perinatal outcome will be of more importance than the composite adverse perinatal outcome, considering its long term implications.

### 4.2. Strengths and limitations

Strengths of the study are the randomised design, its large sample size, the prospective registration of all resource use and the diversity and high number of different participating centres, both midwifery practices, non-academic hospitals and academic hospitals. No cases were lost to follow-up.

Only 70/1801 (3.9 %) women did not have a complete time-table. This allowed us to make a good and reliable cost estimate for the vast majority of women, i.e. 1731/1801 (96.1 %). The missing data in the 70 incomplete time-tables concerned mainly the starting time of cervical ripening and/or induction of labour, or time of arrival at hospital and/or labour room.

The rate of 11.8 % (213/1801) for home births in our trial is lower than the Dutch national figures from 2016 (end of trial). In 2016, 14.0 % (21,679/155,070) of women with a term (37–42 weeks) singleton pregnancy in The Netherlands gave birth at home [27]. A likely reason for this lower rate could be both that more women were induced because of their participation in the trial, and for the reason that the trial population consisted largely of women with a preference for IOL. Women participating in the INDEX trial were more anxious and had less preference for a 'natural' or home birth comparing to eligible women refusing participation in the trial, mainly because of preference for expectant management. [11].

We chose a healthcare perspective as insight in direct medical costs as this perspective is of importance for birth care professionals, policymakers as well as healthcare budget keepers and affects the budgeting of care. Value-Based Health Care (VBHC) focuses on maximising outcomes achieved per dollar spent and shows many parallels with cost-effectiveness analysis. Results of a CEA provide input into the process of providing VBHC [28]. The outcomes of this CEA are also components of the patient reported outcome measures (PROMs) of the concept of Value-Based Health Care.

We did not use quality-adjusted life year (QALY) as a summary measure of health outcome for our economic evaluation, thus combining the impact on both the quantity and quality of life [29]. Questionnaires covering quality of life, experience and satisfaction were not available from all included women in the trial.

### 4.3. Interpretation

Three other randomised controlled trials starting induction at 41 weeks versus expectant management until 42 weeks have been performed in the last two decades [5, 30, 31]. The recently updated Cochrane review *'Induction of labour at or beyond 37 weeks' gestation'* (July 2020) concluded that induction of labour (IOL) $\geq 40$ weeks will improve maternal and perinatal outcomes [32]. In this review however, gestational age in expectant management was allowed to go far beyond 42 weeks in a majority of the included trials comparing induction from 41 weeks onwards with a policy of expectant management [33].

The IPD-MA *'Induction of labour at 41 weeks or expectant management until 42 weeks'* (December 2020) focused on the 41–42 week timeframe and shows both a decrease of perinatal mortality (RR 0.21; 95 % CI 0.06–0.78) 0.019) and of composite severe adverse outcome (RR 0.43; 95 % CI 0.21–0.91; *p*-value 0.027) in favour of induction of labour [6]. Early stopping of the SWEPIS trial, one of the included trials, and differences in antenatal foetal monitoring between the expectant management groups of the two included trials hampers a clear interpretation of the risk reduction of perinatal mortality.

In the INDEX trial, one of the included trials in the IPD-MA, lower composite adverse perinatal outcome was observed in multiparous (IOL: 0.9 %; EM: 1.8 %) compared with nulliparous women (IOL: 2.4 %; EM: 4.1 %). This was noticed also in the IPD-MA. Nulliparous women had a significantly lower risk of severe adverse perinatal outcomes after induction at 41 weeks compared with expectant management up to 42 weeks (IOL 0.33 %; EM 1.58 %; RR 0.20; 95 % CI 0.07–0.60). These results imply that some children of low-risk nulliparous women reaching a gestational age of 41 weeks' could benefit from elective induction at 41 weeks, with a number needed to treat of 79 (95 % CI 49–201). For neonates of low-risk multiparous women the results show no benefit in respect of severe adverse perinatal outcome. (IOL 0.56 %; EM 0.30 %; RR 1.93; 95 % CI 0.48–7.72) [6].

The PROBAAT-2 study suggests, as potential alternative to decrease the higher costs of hospital stay in the induction group, outpatient induction with a Foley catheter in low-risk pregnancies and only induction among nulliparous women. In their scenario cost effectiveness analysis outpatient induction resulted in lower costs in the Foley group [19]. This scenario analysis could also be applicable to low-risk nulliparous women reaching a gestational age of 41 weeks. Studies are underway in the Netherlands to examine the effects of outpatient induction using a Foley catheter in a low risk population.

The recently published prospective cohort study alongside the INDEX trial shows less composite adverse as well as less *severe* adverse perinatal outcomes in both the induction of labour (CAPO 1.1 %; SAPO 0.3 %) and expectant management (CAPO 1.9 %; SAPO 1.0 %) groups compared to the trial [7]. This could indicate selection bias in trial, which is likely when we compare baseline characteristics with higher BMI and anxiety and lower SES and quality of life in the trial population. Furthermore, women participating in the trial and cohort study differed in their management preference; 79.6 % of the women who participated in the trial had a preference for IOL (215/288) and 89.0 % of the women in the prospective cohort study had a preference for EM (226/316). As the risk differences were lower in the cohort study (CAPO: trial 1.4 % versus cohort 0.8 %; SAPO: trial 0.9 % versus cohort 0.7 %), the

cost-effectiveness will decrease compared to the trial.

The conclusion of our study is mostly in line with the SWEPIS cost-effectiveness study, with the difference that we based our analyses on composite (severe) adverse outcomes and SWEPIS on perinatal mortality. Furthermore, SWEPIS used mean neonatal QALYs from general Swedish birth records. The QALYs for the neonates were calculated by adjusting the survival probabilities for each future life year by the mean (Swedish population) health-related quality of life score at each age [10]. For mothers they calculated a non-significant difference (QALY difference 0.0016; *p*-value 0.17) in health-related quality of life. A separate analysis in nulli- and multiparous women was not performed in the SWEPIS trial [10]. The disadvantage of not stratifying by parity is that possible differences between nulliparous and multiparous women will stay invisible. In daily clinical practice, women are always nulli- or multiparous with apparent differences in the labour and childbirth process. Therefore the results of the sensitivity analysis according to parity are more relevant than the general findings. We reported that with a willingness-to-pay for multiparous women of €190,000 and €970,000 for CAPO and SAPO, respectively, induction of labour at 41 weeks for multiparous women is unlikely to be cost-effective.

## 5. Conclusion

In our trial population of low-risk women with a gestational age of 41 weeks, induction of labour at 41 weeks has an 80 % chance of being cost-effective at a willingness-to-pay of €22,000 for prevention of one adverse perinatal outcome and €50,000 for prevention of one severe adverse perinatal outcome. According to the clinical results of the INDEX study, induction at 41 weeks resulted in less composite adverse perinatal outcome (CAPO) and non-significant less composite severe adverse perinatal outcome (SAPO) compared to a policy of expectant management until 42 weeks.

According to the ICERs, spending €9436 and €14,994 on induction of labour, respectively, could prevent one additional CAPO and one SAPO as compared to expectant management.

Subgroup analysis suggests that induction of labour could be cost-effective for nulliparous women while it is unlikely that induction is cost-effective for multiparous women.

To prevent one adverse perinatal outcome, induction of labour has an 80 % chance of being cost-effective at a willingness-to-pay for nulliparous of €47,000 and for multiparous €190,000. To prevent one SAPO the willingness-to-pay is €62,000 for nulliparous and €970,000 for multiparous women.

Cost effectiveness in other settings will depend on baseline characteristics of the population and health system organisation and funding.

### Linked article

Induction of labour at 41 weeks versus expectant management until 42 weeks (INDEX): multicentre, randomised non-inferiority trial. BMJ 2019;364:l344.

### Ethical approval

Both the INDEX trial and this economic analysis, were approved by the ethics committee of the Academic Medical Centre, Amsterdam (No NL38455.018.11). The board of directors of each of the participating centres approved local execution of the study.

### Funding

### CRediT authorship contributions statement

EdM and BWM initiated the INDEX study. RvE designed the economic analysis. AB and RvE conducted the statistical analyses and take responsibility for the integrity of the data and accuracy of the data analyses. AB wrote the first and subsequent drafts of the paper, with RvE writing the methods section of the first draft. MvW and RvE advised on statistical issues and interpretation of the results. JKJK, JCK, JvD, RvE, JvdP, BWM and EdM have provided feedback on all preliminary versions. All authors have approved the final version of this manuscript submitted for publication. AB, RvE and EdM are guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. The authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Disclosure of interests

### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.eurox.2023.100178.

### References

[1] Boers KE, Vijgen SM, Bijlenga D, van der Post JA, Bekedam DJ, Kwee A, et al. Induction versus expectant monitoring for intrauterine growth restriction at term: randomised equivalence trial (DIGITAT). BMJ 2010;341:c7087.

[2] Koopmans CM, Bijlenga D, Groen H, Vijgen SM, Aarnoudse JG, Bekedam DJ, et al. 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(9694):979–88.

[3] Keulen JK, Bruinsma A, Kortekaas JC, van Dillen J, Bossuyt PM, Oudijk MA, et al. Induction of labour at 41 weeks versus expectant management until 42 weeks (INDEX): multicentre, randomised non-inferiority trial. BMJ 2019;364:l344.

[4] ACOG. Committee opinion No 579: definition of term pregnancy. Obstet Gynecol 2013;122(5):1139–40.

[5] Wennerholm UB, Saltvedt S, Wessberg A, Alkmark M, Bergh C, Wendel SB, et al. Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (SWEdish Post-term Induction Study, SWEPIS): multicentre, open label, randomised, superiority trial. BMJ 2019;367:l6131.

[6] Alkmark M, Keulen JKJ, Kortekaas JC, Bergh C, van Dillen J, Duijnhoven RG, et al. Induction of labour at 41 weeks or expectant management until 42 weeks: a systematic review and an individual participant data meta-analysis of randomised trials. PLoS Med. 2020;17(12):e1003436.

[7] Bruinsma A, Keulen JK, Kortekaas JC, van Dillen J, Duijnhoven RG, Bossuyt PM, et al. Elective induction of labour and expectant management in late-term pregnancy: a prospective cohort study alongside the INDEX randomised controlled trial. Eur J Obstet Gynecol Reprod Biol X 2022;16:100165.

[8] Einerson BD, Nelson RE, Sandoval G, Esplin MS, Branch DW, Metz TD, et al. Cost of elective labor induction compared with expectant management in Nulliparous women. Obstet Gynecol 2020;136(1):19–25.

[9] Hersh AR, Greiner KS, Garg B, Skeith AE, Caughey AB. Maternal and neonatal hospitalization costs associated with elective induction of labor at term in California, 2007-2011. Obstet Gynecol 2020;136(1):8–18.

[10] Alkmark M, Wennerholm UB, Saltvedt S, Bergh C, Carlsson Y, Elden H, et al. Induction of labour at 41 weeks of gestation versus expectant management and induction of labour at 42 weeks of gestation: a cost-effectiveness analysis. BJOG; 2021.

[11] Keulen JKJ, Nieuwkerk PT, Kortekaas JC, van Dillen J, Mol BW, van der Post JAM, et al. What women want and why. Women's preferences for induction of labour or expectant management in late-term pregnancy. Women Birth 2021;34(3):250–6.

[12] Brazier JE, Dixon S, Ratcliffe J. The role of patient preferences in cost-effectiveness analysis: a conflict of values? Pharmacoeconomics 2009;27(9):705–12.

[13] Kortekaas JC, Bruinsma A, Keulen JK, van Dillen J, Oudijk MA, Zwart JJ, 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.

[14] Committee Opinion. No. 644: The Apgar Score. Obstet Gynecol 2015;126(4): e52–5.

[15] Hakkaart-van Roije L, van der Linden N, Bouwmans C, Kanters T, Tan SS. Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg [Cost manual: Cost research

methodology and reference prices for economic evaluations in health care]. Diemen: Zorginstituut Nederland,; 2015.

[16] Nederlandse Zorgautoriteit[Dutch Healthcare Autority tariff]; 2021. Available from: ⟨www.nza.nl⟩.

[17] Pharmacotherapeutic Compass: Zorginstituut Nederland [Health Insurance Board]; 2021. Available from: ⟨www.farmacotherapeutischkompas.nl⟩.

[18] Liu X, Wang Y, Zhang F, Zhong X, Ou R, Luo X, et al. Double- versus single-balloon catheters for labour induction and cervical ripening: a meta-analysis. BMC Pregnancy Childbirth 2019;19(1):358.

[19] Ten Eikelder M, van Baaren GJ, Oude Rengerink K, Jozwiak M, de Leeuw JW, Kleiverda G, et al. Comparing induction of labour with oral misoprostol or Foley catheter at term: cost-effectiveness analysis of a randomised controlled multi-centre non-inferiority trial. BJOG 2018;125(3):375–83.

[20] van Baaren GJ, Broekhuijsen K, van Pampus MG, Ganzevoort W, Sikkema JM, Woiski MD, et al. An economic analysis of immediate delivery and expectant monitoring in women with hypertensive disorders of pregnancy, between 34 and 37 weeks of gestation (HYPITAT-II). BJOG 2017;124(3):453–61.

[21] Statistics Consumer Pricing IndexNetherlands Den Haag/Heerlen: Centraal Bureau voor de Statistiek (CBS) [Central Statistical Office]; 2021. Available from: ⟨www.statline.nl⟩.

[22] Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves–facts, fallacies and frequently asked questions. Health Econ 2004;13(5):405–15.

[23] Willingness-to-Pay [online]. York; York Health Economics Consortium; 2016. [Available from: ⟨https://yhec.co.uk/glossary/willingness-to-pay/⟩.

[24] Bertram MY, Lauer JA, De Joncheere K, Edejer T, Hutubessy R, Kieny MP, et al. Cost-effectiveness thresholds: pros and cons. Bull World Health Organ 2016;94 (12):925–30.

[25] R Core Team. R: A language and environment for statistical computing. Austria: R Foundation for Statistical Computing; 2019. Available from: ⟨https://www.R-project.org/⟩.

[26] Andrew Briggs, Mark Sculpher, Karl Claxton. Decision modelling for health economic. In: Evaluation. Oxford University Press; 2006.

[27] Perined Utrecht; 2022 [07–06-2022]. Available from: ⟨www.peristat.nl⟩.

[28] Tsevat J, Moriates C. Value-Based Health Care Meets Cost-Effectiveness Analysis. Ann Intern Med 2018;169(5):329–32.

[29] Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull 2010;96:5–21.

[30] Gelisen O, Caliskan E, Dilbaz S, Ozdas E, Dilbaz B, Ozdas E, 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(2):164–9.

[31] Sahraoui W, Hajji S, Bibi M, Nouira M, Essaidi H, Khairi H. Management of pregnancies beyond forty-one week's gestation with an unfavorable cervix. J Gynecol Obstet Biol Reprod (Paris) 2005;34(5):454–62.

[32] Middleton P, Shepherd E, Morris J, Crowther CA, Gomersall JC. Induction of labour at or beyond 37 weeks' gestation. Cochrane Database Syst Rev 2020;7:CD004945.

[33] 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.