Challenges and Limitations of Clinical Trials on Labor Induction: A Review of the Literature

Oscar A. Viteri, MD1 Baha M. Sibai, MD2

1 Avera Medical Group Maternal Fetal Medicine, Avera McKennan Hospital and University Health Center, Sioux Falls, South Dakota 2 Department of Obstetrics, Gynecology and Reproductive Sciences, UTHealth McGovern Medical School, Houston, Texas


Abstract

Induction of labor is a common obstetric procedure performed in nearly a quarter of all deliveries in the United States. Pharmacological (prostaglandins, oxytocin) and/or mechanical methods (balloon catheters) are commonly used for labor induction; however, there is ongoing debate as to which method is the safest and most effective. This narrative review discusses key limitations of published trials on labor induction, including the lack of well-designed randomized controlled trials directly comparing specific methods of induction, heterogeneous trial populations, and wide variation in the protocols used and outcomes reported. Furthermore, the majority of published trials were underpowered to detect significant differences in the most clinically relevant efficacy and safety outcomes (e.g., cesarean delivery, neonatal mortality). By identifying the limitations of labor induction trials, we hope to highlight the importance of quality published data to better inform guidelines and drive evidence-based treatment decisions.

Keywords
► labor induction
► prostaglandins
► dinoprostone
► misoprostol
► oxytocin
► clinical trials

Induction of labor is a common obstetric procedure that is used in nearly a quarter (24.5%) of all deliveries in the United States (U.S.). Induction of labor is indicated when the risks to maternal or fetal health outweigh the benefits of continuing the pregnancy. One of the most common reasons for labor induction in the U.S. is postterm pregnancy (i.e., > 41 weeks of gestation). Other indications for labor induction include chorioamnionitis, gestational hypertension, pre-eclampsia, premature rupture of membranes (PROM), maternal conditions (e.g., diabetes mellitus, chronic hypertension), and fetal compromise (e.g., severe fetal growth restriction, oligohydramnios). Women may also choose to undergo elective labor induction (i.e., inductions performed without maternal or fetal medical indication) to shorten the duration of pregnancy or to schedule the day of delivery. In the U.S., elective inductions account for approximately 10% of all labor inductions.

The process of labor induction often begins with cervical ripening which involves softening and thinning the cervix in preparation for labor and delivery. Different methods are available for cervical ripening and labor induction, including pharmacological and/or mechanical methods. Oxytocin, a hormone that stimulates uterine contractions, is considered less effective for cervical ripening and is typically used alone for labor induction when the cervix is favorable (Bishop’s score ≥ 6) or after cervical ripening with prostaglandins or mechanical methods. The use of oxytocin alone for cervical ripening has been associated with higher rates of unsuccessful vaginal deliveries within 24 hours and increased rates of cesarean delivery compared with the use of prostaglandins. Dinoprostone (prostaglandin E2) is the only prostaglandin approved by the U.S. Food and Drug Administration (FDA) for cervical ripening and induction of labor. Misoprostol (prostaglandin E1 analog) is widely used for labor induction, although it is not approved by the FDA. Mechanical methods of labor induction (e.g., Foley’s balloon catheters, double-balloon catheters, hygroscopic and osmotic dilators) are effective alone but are also commonly combined with pharmacological methods. Oxytocin, both prostaglandin preparations, and balloon catheters are

received
June 15, 2018
accepted after revision
October 12, 2018

ISSN 2157-6998.

Copyright © 2018 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
Tel: +1(212) 584-4662.

License terms
recommended by the American College of Obstetricians and Gynecologists (ACOG) and World Health Organization (WHO) for labor induction.\(^2,^4\)

There are several potential maternal and neonatal complications with labor induction, although mechanical methods, such as the Foley catheter, are associated with the least number of adverse effects.\(^1^0\) Previous trials have suggested that women undergoing induction of labor with pharmacological and/or mechanical methods may be at a higher risk of cesarean section. However, recent analyses dispute the association between labor induction and increased risk of a cesarean delivery.\(^1^5^–^1^7\) For example, preliminary data from a recent large, randomized, controlled trial (RCT) in healthy nulliparous women comparing elective induction of labor at 39 weeks versus expectant management (A Randomized Trial of Induction Versus Expectant Management [ARRIVE]; \(n = 6,106\)) showed a reduction in cesarean delivery rates with labor induction versus expectant management (18.6 vs. 22.2%; relative risk [RR] = 0.84; 95% confidence interval [CI]: 0.76–0.93).\(^1^8\) Of note, rates of the composite primary outcome of adverse perinatal events were similar in women undergoing labor induction versus those in the expectant management group (4.3 vs. 5.4%; RR = 0.80; 95% CI: 0.64–1.00).\(^1^8\)

While ACOG recommends several methods for labor induction,\(^2\) the safest and most effective method has not yet been clearly established. Although effective, pharmacological methods of labor induction have been associated with potential risks, such as uterine tachysystole with or without fetal heart rate changes, uterine rupture, fetal distress, and maternal and perinatal morbidity and mortality.\(^2,^1^9^–^2^2\) The incidence of these adverse events is dependent on the agent used for induction, dose, and other factors. The appropriate method of induction, whether pharmacological and/or mechanical, may vary depending on maternal and neonatal factors, patient demographics, and other variables that can complicate clinical trial design and present a barrier to evidence-based recommendations. Clinical trials in labor induction are further limited by the lack of uniform protocols, inconsistent and incomplete reporting of outcomes, and lack of statistical power for clinically relevant efficacy and safety outcomes, such as cesarean delivery rates or severe neonatal morbidity and mortality. These limitations make it difficult to interpret trial data and draw meaningful conclusions about which method of labor induction is appropriate based on indication and other patient-related characteristics.

This narrative review discusses the limitations of published clinical trials on labor induction and highlights fundamental issues in trial design and methodology that can impact the interpretation of results. We provide recommendations for future trials on labor induction, including important components of a rigorous, well-designed, and properly conducted clinical trial. We hope that these recommendations will yield quality data to better inform guidelines and drive evidence-based practice to help physicians make better treatment decisions.

**Methods**

A comprehensive literature search was conducted for this narrative review using PubMed and the search terms “labor AND induction” or “labor AND delivery AND vaginal,” including both the U.S. and British spelling of “labor” and “labour.” The search was limited to English language RCTs, prospective trials, and observational studies published between January 2007 and October 2017.

A total of 548 titles and abstracts were manually screened for inclusion. Clinical trials of pharmacological (i.e., prostaglandins, oxytocin) and mechanical methods of labor induction were included in the review. Exclusion criteria included: retrospective studies; descriptions of trial protocols (i.e., no efficacy or safety data reported); secondary analyses; intra-dose-comparison trials; trials of alternative methods of labor induction (i.e., acupuncture, nipple stimulation, nitric oxide donors [isosorbide mononitrate, isonicotinic acid hydrazide], dexamethasone, propranolol); trials with outpatient agent administration; trials in women with PROM, preterm PROM (pPROM), or trial of labor after cesarean (TOLAC); trials of induction for abortion or fetal death; and any other off-topic trials that did not compare methods of labor induction (e.g., studies to identify predictors of successful labor induction, methods for pain management during induction, and pharmacokinetic studies).

**Summary of Literature Search Results**

A total of 103 clinical trials met the inclusion and exclusion criteria; 63 were RCTs, 21 were nonrandomized prospective trials, and 19 were observational studies. Of the 63 RCTs,\(^2^3^–^8^5\) most were single-center trials conducted outside the U.S. that enrolled a limited number of participants. Just 12 RCTs were conducted in the U.S. (\(\rightarrow\) Table 1). The inclusion criteria for most trials were nulliparous or multiparous women with single, cephalic pregnancies and obstetric or medical indications for induction of labor. The most common primary outcomes were time from the start of induction to delivery and vaginal delivery within 24 hours. The reporting of safety outcomes varied, yet all trials included at least some maternal and/or neonatal safety data.

The published clinical trials on labor induction, including trials conducted within and outside the U.S., have many limitations that are both inherent to the trial designs and a result of the therapeutic area and patient population. These limitations are discussed in detail in the following section and are summarized in \(\rightarrow\) Table 2.

**Limitations of Trials on Labor Induction**

**Trial Design, Sample Size, Randomization, and Blinding**

Various clinical trial designs were identified in the literature search, including observational trials, nonrandomized prospective trials, and RCTs. Of the 103 trials identified, just 63 were RCTs.\(^2^3^–^8^5\) The majority of RCTs were single-center trials conducted in academic hospitals, which can limit the
Table 1: Summary of published labor induction RCTs in the United States

<table>
<thead>
<tr>
<th>Clinical trial</th>
<th>Drug/device</th>
<th>Trial design</th>
<th>Trial population</th>
<th>Outcomes</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schoen et al (2017), intracervical Foley catheter with and without oxytocin for labor induction: a randomized controlled trial</td>
<td>Foley's catheter with concurrent oxytocin infusion vs. Foley's catheter followed by oxytocin infusion</td>
<td>Multicenter parallel RCT</td>
<td>n = 323</td>
<td>Nulliparous and multiparous women with a singleton pregnancy at ≥ 24 wk. Bishop’s score &lt; 6. Women with pPROM were not excluded.</td>
<td>Primary: Delivery ≤ 24 h Secondary: Time to Foley’s expulsion Change in Bishop’s score Need for additional ripening. Analgesia during Foley’s catheter use Time to second stage Delivery ≤ 12 h Total time to delivery Duration of oxytocin use Mode of delivery Tachysystole Chorioamnionitis Postpartum hemorrhage Severe maternal morbidity (uterine rupture, ICU admission, or maternal death) Neonatal outcomes (weight, 5-min Apgar’s score &lt; 7, NICU admission, NICU length of stay) More nulliparous (64 vs. 43%; p = 0.003) women who received Foley’s catheter with concurrent oxytocin delivered ≤ 24 h vs. Foley’s catheter followed by oxytocin. Median time to delivery was shorter in both nulliparous (20.9 vs. 26.1 h; p &lt; 0.001) and multiparous (14.9 vs. 18.6 h; p = 0.01) women who received Foley’s catheter with concurrent oxytocin. No significant differences in cesarean section rates, postpartum hemorrhage, chorioamnionitis, or NICU admission.</td>
</tr>
<tr>
<td>Connolly et al, a randomized trial of Foley balloon induction of labor trial in nulliparas (FIAT-N)</td>
<td>Foley’s catheter with concurrent oxytocin infusion vs. Foley’s catheter followed by oxytocin infusion</td>
<td>Single-center RCT</td>
<td>n = 166</td>
<td>Nulliparous women with a singleton pregnancy at ≥ 24 wk gestation. Cervical dilation &lt; 3 cm.</td>
<td>Primary: Time to delivery Secondary: Cesarean delivery rate Chorioamnionitis Estimated blood loss Postpartum hemorrhage Composite neonatal outcome (≥ 1 of 5-min Apgar’s score &lt; 5, umbilical artery pH &lt; 7.1, NICU admission, NEC, or neonatal death) Women who received Foley’s catheter with concurrent oxytocin had a shorter mean time to delivery vs. Foley’s catheter followed by oxytocin (15.9 vs. 18.9 h; p = 0.004). No significant differences in cesarean delivery rates, chorioamnionitis, estimated blood loss, postpartum hemorrhage, or composite neonatal outcome.</td>
</tr>
<tr>
<td>Levine et al, mechanical and pharmacologic methods of labor induction: a randomized controlled trial</td>
<td>Foley’s catheter plus vaginal misoprostol tablet vs. Foley’s catheter plus oxytocin infusion vs. Foley’s catheter alone vs. vaginal misoprostol tablet alone</td>
<td>Single-center RCT</td>
<td>n = 491</td>
<td>Nulliparous and multiparous women with a singleton pregnancy at ≥ 37 wk gestation. Bishop’s score &lt; 6 Cervical dilation ≤ 2 cm.</td>
<td>Primary: Time to delivery Secondary: Cesarean delivery rate Time to vaginal delivery Time to delivery censored for cesarean section Time to active labor Delivery ≤ 12 h Delivery ≤ 24 h Maternal length of stay Indication for cesarean section Maternal morbidity composite (≥ 1 of third or fourth degree perineal laceration, blood transfusion, endometritis, wound separation infection, venous thromboembolism, hysterectomy, ICU admission, death) Chorioamnionitis Use of terbutaline Intrauterine pressure catheter Amnioinfusion. Median time to delivery was faster with combination methods vs. single agents (Foley’s plus misoprostol, 13.1 h; Foley’s plus oxytocin, 14.5; misoprostol, 17.6 h; Foley’s, 17.7 h; p &lt; 0.001). After censoring for cesarean delivery and adjusting for parity, women who received Foley’s plus misoprostol were almost twice as likely to deliver before those receiving misoprostol or Foley’s alone (hazard ratio = 1.92 and 1.87, respectively). No significant differences in cesarean delivery rates, indication for cesarean delivery, composite maternal outcome, or neonatal outcomes.</td>
</tr>
</tbody>
</table>
### Table 1 (Continued)

<table>
<thead>
<tr>
<th>Clinical trial</th>
<th>Drug/device</th>
<th>Trial design</th>
<th>Trial population</th>
<th>Outcomes</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edwards et al, Foley catheter compared with the controlled-release dinoprostone insert: a randomized controlled trial[^9]</td>
<td>Foley’s catheter vs. vaginal dinoprostone insert</td>
<td>Multicenter RCT</td>
<td>n = 376</td>
<td>• Nulliparous and multiparous women with a singleton pregnancy at ≥ 36 wk gestation</td>
<td>• Women in the Foley’s catheter group had a shorter median time to delivery (21.6 vs. 26.6 h; p = 0.003) and vaginal delivery (20.1 vs. 24.3 h; p = 0.005) vs. women who were treated with dinoprostone</td>
</tr>
<tr>
<td></td>
<td>Dosing:</td>
<td></td>
<td></td>
<td>• Cervical dilation &lt; 3 cm; if 2 cm dilated, &lt; 80% effaced</td>
<td>• More women who received the Foley’s catheter delivered ≤ 24 h (56 vs. 40%; p = 0.003) and delivered vaginally ≤ 24 h (44 vs. 30%; p = 0.004)</td>
</tr>
<tr>
<td></td>
<td>• 16F Foley’s catheter inflated to 30 mL</td>
<td></td>
<td></td>
<td>Primary:</td>
<td>• No significant differences in other secondary outcomes</td>
</tr>
<tr>
<td></td>
<td>• 10 mg vaginal dinoprostone insert, max 12 h</td>
<td></td>
<td></td>
<td>Time to delivery</td>
<td></td>
</tr>
<tr>
<td>Suffecool et al, labor induction in nulliparous women with an unfavorable cervix: double balloon catheter versus dinoprostone[^6]</td>
<td>Vaginal dinoprostone insert vs. double-balloon catheter</td>
<td>Single-center RCT</td>
<td>n = 62</td>
<td>• Nulliparous women with a singleton pregnancy at ≥ 37 wk gestation</td>
<td>• Women who received the double-balloon catheter had a shorter mean time to delivery (17.9 vs. 26.3 h; p = 0.001) and mean time to vaginal delivery (19.1 vs. 24.4 h; p = 0.05) vs. those who received vaginal dinoprostone</td>
</tr>
<tr>
<td></td>
<td>Dosing:</td>
<td></td>
<td></td>
<td>• Bishop’s score &lt; 6</td>
<td>• More women in the double-balloon group delivered ≤ 24 h (87 vs. 48%; p = 0.002)</td>
</tr>
<tr>
<td></td>
<td>• Both balloons of the double-balloon catheter inflated to 80 mL</td>
<td></td>
<td></td>
<td>Primary:</td>
<td>• Cesarean delivery rates were similar in both treatment groups</td>
</tr>
<tr>
<td></td>
<td>• 10 mg vaginal dinoprostone insert, max 12 h</td>
<td></td>
<td></td>
<td>Time to delivery</td>
<td></td>
</tr>
<tr>
<td>Carbone et al, combination of Foley bulb and vaginal misoprostol compared with vaginal misoprostol alone for cervical ripening and labor induction: a randomized controlled trial[^9]</td>
<td>Foley’s catheter plus vaginal misoprostol tablet vs. vaginal misoprostol tablet alone</td>
<td>Single-center RCT</td>
<td>n = 123</td>
<td>• Nulliparous women with a singleton pregnancy at ≥ 24 wk gestation</td>
<td>• Women who received Foley’s plus misoprostol had a shorter mean time to delivery vs. misoprostol alone (15.3 vs. 18.3 h; p = 0.03)</td>
</tr>
<tr>
<td></td>
<td>Dosing:</td>
<td></td>
<td></td>
<td>• Bishop’s score &lt; 6</td>
<td>• There were no significant differences in mode of delivery, labor complications, or adverse maternal and neonatal outcomes</td>
</tr>
<tr>
<td></td>
<td>• Foley’s catheter inflated to 60 mL</td>
<td></td>
<td></td>
<td>Primary:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 25 μg vaginal misoprostol tablet given every 4 h</td>
<td></td>
<td></td>
<td>Time to delivery</td>
<td></td>
</tr>
<tr>
<td>Wing et al (2013), misoprostol vaginal insert and time to vaginal delivery: a randomized controlled trial[^4]</td>
<td>Vaginal misoprostol insert vs. vaginal dinoprostone insert</td>
<td>Double-blind multicenter RCT</td>
<td>n = 1,358</td>
<td>• Nulliparous and multiparous women with a singleton pregnancy at ≥ 36 wk gestation</td>
<td>• Compared with the dinoprostone insert, women who received the misoprostol insert had a shorter median time to vaginal delivery (32.8 vs. 21.5 h), any delivery (27.3 vs. 18.3 h), and onset to active labor (18.6 vs. 12.1 h; p &lt; 0.001 for all comparisons)</td>
</tr>
<tr>
<td></td>
<td>Dosing:</td>
<td></td>
<td></td>
<td>• Bishop’s score &lt; 4</td>
<td>• Time to vaginal delivery</td>
</tr>
<tr>
<td></td>
<td>• 200 μg vaginal misoprostol insert, max 24 h</td>
<td></td>
<td></td>
<td></td>
<td>• Cesarean delivery rate</td>
</tr>
<tr>
<td></td>
<td>• 10 mg vaginal dinoprostone insert, max 24 h</td>
<td></td>
<td></td>
<td></td>
<td>• Secondary:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time to delivery</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Tachysystole</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Maternal or fetal adverse events (e.g., tachysystole)</td>
</tr>
</tbody>
</table>

[^9]: Clinical trial
[^6]: Drug/device
[^20]: Trial design
[^21]: Trial population
[^3]: Outcomes
[^4]: Main results

- Nulliparous: Women who have never given birth before.
- Multiparous: Women who have given birth before.
- Singleton pregnancy: A pregnancy involving one fetus.
- Cervical dilation: The opening or dilatation of the cervix (the lower part of the uterus). It is measured in centimeters.
- BMI: Body Mass Index. It is calculated using the formula weight in kilograms divided by height in meters squared (BMI = weight(kg)/height(m)^2).
- Parity: Number of previous live births.
- NICU: Neonatal Intensive Care Unit. It is a specialized unit providing intensive medical and nursing care for newborn infants who are critically ill.
- NICU admission: The need for admission to a Neonatal Intensive Care Unit.
- Apgar score: A scoring system for evaluating newborn infants right after birth. It is measured at 1 and 5 minutes of age.
- Chorioamnionitis: An inflammation of the placenta and amniotic membranes that can occur during pregnancy.
- Tachysystole: An abnormal condition where the rate of uterine contractions is unusually high.
- Operative vaginal delivery: Delivery that requires the use of medical instruments or surgery.
- Cesarean delivery rate: The percentage of births that were Cesarean deliveries.
- Maternal or fetal adverse events: Malevolent conditions resulting from labor or delivery.
- NICU admission: Neonatal Intensive Care Unit admission.
- Time to delivery: The time from the start of labor to the delivery of the baby.
- Requirement for postpartum cooling: The need for postpartum cooling measures.
- Maternal or fetal complications: Malevolent conditions resulting from labor or delivery.
- NICU admission: Neonatal Intensive Care Unit admission.
- Time to vaginal delivery: The time from the start of labor to the delivery of the baby via vaginal route.
- Time to active labor: The time from the start of labor to the establishment of active labor.
- NICU admission: Neonatal Intensive Care Unit admission.
- NICU admission: Neonatal Intensive Care Unit admission.
- NICU admission: Neonatal Intensive Care Unit admission.
- NICU admission: Neonatal Intensive Care Unit admission.
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Clinical trial</th>
<th>Drug/device</th>
<th>Trial design</th>
<th>Trial population</th>
<th>Outcomes</th>
<th>Main results</th>
</tr>
</thead>
</table>
| **Fitzpatrick et al, cervical ripening with Foley balloon plus fixed vs. incremental low-dose oxytocin: a randomized controlled trial**58 | Foley’s catheter plus fixed low-dose oxytocin infusion vs. Foley’s catheter plus standard incremental low-dose oxytocin infusion | Single-center RCT | n = 116 | Primary: | • Cesarean delivery rates were similar with both treatments  
• Tachysystole requiring intervention occurred in 13.3% of women receiving misoprostol and 4.0% of women receiving dinoprostone (p < 0.001) |
| | Dosing:  
• Fixed low-dose oxytocin infusion of 2 mL/min  
• Incremental low-dose oxytocin infusion starting at 1 mL/min and increasing by 2 mL/min every 30 min, max dose 20 mL/min | | | Secondary: | • There were no significant differences in median time to delivery in women who received the Foley’s catheter plus fixed low-dose or incremental low-dose oxytocin (23.7 vs. 19.2 h; p = 0.388)  
• No significant differences in other secondary outcomes |
| **Hill et al, a randomized clinical trial comparing vaginal misoprostol versus cervical Foley plus oral misoprostol for cervical ripening and labor induction**72 | Foley’s catheter plus oral misoprostol vs. vaginal misoprostol | Single-center RCT | n = 126 | Primary: | • Women who received Foley’s plus oral misoprostol had a shorter time to delivery (12.9 vs. 17.8 h; p = 0.001) and were more likely to deliver ≤ 24 h (relative risk = 1.36) vs. those who received vaginal misoprostol  
• Cesarean delivery rate and indications for cesarean delivery were similar in both treatment groups  
• A higher rate of tachysystole was seen with Foley’s plus oral misoprostol vs. vaginal misoprostol (39 vs. 21%; p = 0.015)  
• No significant differences in neonatal outcomes |
| | Dosing:  
• 24F Foley catheter inflated to 50 mL  
• 100 μg oral misoprostol at 4 to 6-h intervals, max 4 doses  
• Initial dose of 25 μg vaginal misoprostol followed by 50 μg doses at 3 to 6-h intervals, max 8 doses | | | Secondary: | • Women who received Foley’s plus oral misoprostol and vaginal (48 vs. 46%; p = 0.82) deliveries ≤ 24 h in women who received Foley’s plus oxytocin and those who received the Foley’s alone |
| **Fonseca et al, randomized trial of preinduction cervical ripening: misoprostol vs oxytocin**79 | Vaginal misoprostol vs. oxytocin infusion | Single-center RCT | n = 361 | Primary: | • Vaginal delivery rates were similar in women treated with misoprostol or oxytocin (81 vs. 87%; p = 0.31)  
• Mean time to delivery was shorter with oxytocin vs. misoprostol (13.1 vs. 16.3 h; p = 0.005)  
• No significant differences in neonatal outcomes |
| | Dosing:  
• 25 μg vaginal misoprostol every 4 h  
• Oxytocin infusion started at 4 mL/min for nulliparous women and 2 mL/min for multiparous women; increased by 4 mL/min or 2 mL/min, respectively, every 15 min, max dose 40 mL/min | | | | |
| **Pettler et al, transcervical Foley catheter with and without oxytocin for cervical ripening: a randomized controlled trial**81 | Foley’s catheter plus low-dose oxytocin infusion vs. Foley’s catheter alone | Single-center RCT | n = 183 | Primary: | • There were no significant differences in the proportion of overall (65 vs. 60%; p = 0.50) and vaginal (48 vs. 46%; p = 0.82) deliveries ≤ 24 h in women who received Foley’s plus oxytocin and those who received the Foley’s alone |
| | Dosing:  
• 20F Foley’s catheter inflated to 30 mL  
• Oxytocin infusion started at 1 mL/min and increased by 1 mL/min every 15 min, max 10 mL/min | | | | |

(Continued)
### Table 1 (Continued)

<table>
<thead>
<tr>
<th>Clinical trial</th>
<th>Drug/device</th>
<th>Trial design</th>
<th>Trial population</th>
<th>Outcomes</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing (2008), misoprostol vaginal insert compared with dinoprostone vaginal insert: a randomized controlled trial</td>
<td>Vaginal misoprostol insert vs. vaginal dinoprostone insert</td>
<td>Double-blind multicenter RCT</td>
<td><em>n</em> = 1,308</td>
<td>Primary:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dosing:</td>
<td></td>
<td></td>
<td>Time to vaginal delivery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50 μg or 100 μg vaginal misoprostol insert, max 24 h</td>
<td></td>
<td></td>
<td>Cesarean delivery rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 mg vaginal dinoprostone insert, max 24 h</td>
<td></td>
<td></td>
<td>Secondary:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Composite modified Bishop’s score at 12 h (vaginal delivery ≤ 12 h, increase in Bishop’s score ≥ 3 from baseline, Bishop’s score ≥ 6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Delivery ≤ 12 and ≤ 24 h</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Predelivery oxytocin administration</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Time to active labor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in cesarean delivery rates or rates of maternal complications</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Women in the Foley’s plus oxytocin group required more regional analgesia during induction (23 vs. 9%; <em>p</em> = 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Median time to vaginal delivery was similar in women treated with the misoprostol 100 μg vaginal insert and the dinoprostone insert (1,596 vs. 1,650 min; <em>p</em> = 0.97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Women in the misoprostol 50 μg group had a longer time to vaginal delivery vs. those treated with dinoprostone (2,127 vs. 1,650 min; <em>p</em> = 0.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No significant differences in cesarean delivery rates between either misoprostol dose and dinoprostone</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>All three treatments had similar safety profiles</td>
</tr>
</tbody>
</table>

### Table 2 Summary of limitations in published labor induction clinical trials (all RCTs)

<table>
<thead>
<tr>
<th>Trial design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited number of RCTs</td>
</tr>
<tr>
<td>Most were single center and conducted in an academic setting</td>
</tr>
<tr>
<td>Many trials not registered</td>
</tr>
<tr>
<td>Small sample sizes</td>
</tr>
<tr>
<td>Wide variation in the methods of induction assessed</td>
</tr>
<tr>
<td>Different doses, dosing intervals, and routes of administration across trials</td>
</tr>
<tr>
<td>Inappropriate comparators</td>
</tr>
<tr>
<td>Inconsistent labor management and delivery practices</td>
</tr>
</tbody>
</table>

Limited applicability to practice in the U.S.:
- Majority of RCTs were international
- Clinical practice guidelines may vary between countries
- Demographic and cultural differences may influence outcomes
- Differences in the methods of induction evaluated
- More trials in the U.S. assessed combination methods of labor induction

<table>
<thead>
<tr>
<th>Trial population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous patient populations</td>
</tr>
<tr>
<td>Most trials included all women undergoing elective or medically indicated induction</td>
</tr>
<tr>
<td>Variation in required gestational age and Bishop’s score between trials</td>
</tr>
<tr>
<td>Most trials recruited both nulliparous and multiparous women but were not designed to assess these groups separately</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wide variability in the reporting of outcomes between trials</td>
</tr>
<tr>
<td>Most common primary outcomes focused on reducing the duration of induction (i.e., time to delivery, delivery ≤ 24 h)</td>
</tr>
<tr>
<td>The majority of trials were not powered to evaluate more clinically relevant efficacy and safety outcomes (e.g., cesarean delivery)</td>
</tr>
<tr>
<td>Inconsistent definitions of many efficacy and safety parameters across trials</td>
</tr>
<tr>
<td>Few trials included prespecified subgroup analyses</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; ICU, intensive care unit; NEC, necrotizing enterocolitis; NICU, neonatal intensive care unit; pPROM, preterm premature rupture of membranes; RCT, randomized controlled trial.

---

American Journal of Perinatology Reports Vol. 8 No. 4/2018
generalizability of the results to other community-based populations. Among the 12 multicenter trials, the number of centers ranged from 2 to 49 and tended to include both university-affiliated and community hospitals.

Sample size estimation and power analysis are crucial components of any well-designed trial. A sufficiently large sample is needed for reliable accurate results; however, sample size was limited in nearly all trials included in our search. Although power-based sample size calculations were used in 53 RCTs, most trials were underpowered to detect significant differences in secondary efficacy and safety outcomes. In addition to sample size, proper randomization is another important aspect of clinical trial design used to balance treatment groups and eliminate selection bias. Most of the RCTs (n = 58) included in our search described the method of randomization, with the majority using simple randomization techniques (e.g., computer-generated random numbers). Few RCTs included in our search were blinded. However, in most cases, it was not possible to blind patients and medical staff due to the nature of the intervention (e.g., Foley’s catheter vs. vaginal dinoprostone).

**Trial Design: Treatment Protocols**

There was variation in the methods (pharmacological, mechanical, or both) and protocols for labor induction among the RCTs included in our search. Pharmacological interventions varied by dose, dose intervals, and route of administration. For example, misoprostol was administered through sublingual, oral, or vaginal routes. Most trials used 25 or 50 μg misoprostol tablets administered every 3 to 6 hours; two trials used a 50, 100, or 200 μg controlled-release misoprostol vaginal insert. Oral misoprostol was given as a static dose (tablet or solution) or titrated solution. Dinoprostone was typically administered as a 10 mg controlled-release vaginal insert kept in place for 12 to 24 hours. Some trials used dinoprostone gel or tablets administered vaginally (ranging from 1–3 mg every 4–8 hours) or dinoprostone gel administered intracervically (0.5 mg every 6 or 12 hours). Protocols for oxytocin infusion varied, with starting doses ranging from 1 to 4 mU/min, rate increases every 15 to 30 minutes, and maximum dosing ranging from 10 to 40 mU/min. Among the trials that used mechanical methods for induction, most used the Foley catheter (n = 23); different sizes of catheters were used (e.g., 16F, 18F) and were inflated with varying volumes of water or saline solution (30–60 mL). Double-balloon catheters were used in 11 trials and each balloon was filled with up to 80 mL of fluid. The variation in doses and routes of administration with pharmacological agents, as well as the different types of balloon catheters and inflation volumes used in the labor induction trials can lead to different outcomes, making it difficult to compare methods across trials. It should be noted that findings from a recent meta-analysis of six trials suggest there are no differences in labor induction outcomes (i.e., time to delivery, vaginal and cesarean delivery rates, maternal satisfaction) with single versus double balloon catheters.

It is also difficult to compare treatments within some trials due to the use of inappropriate comparators. Comparing methods with different mechanisms of action and efficacy/safety profiles complicate results and makes it difficult to interpret the data. For example, two trials compared the Foley catheter plus oxytocin versus vaginal misoprostol. Another compared vaginal dinoprostone with or without sequential oxytocin versus vaginal misoprostol plus sequential oxytocin. These trials may have yielded more meaningful data if the combination agents were compared with the same single agents (e.g., Foley’s catheter plus oxytocin vs. Foley’s catheter alone or oxytocin alone).

**Trial Design: Labor Management**

Differences in labor management following labor induction also varied between the trials included in our search. In some trials, oxytocin administration was permitted if a patient did not go into active labor but explicit guidelines were typically not provided. Active labor and delivery, including the use of oxytocin for labor augmentation, were managed according to standard institutional practices. For example, a trial comparing the Foley catheter with or without oxytocin used a low-dose oxytocin regimen for induction and augmentation that was specific to that institution. This trial found that the addition of oxytocin to the Foley catheter did not shorten the time to delivery. However, subsequent studies that used a higher dose of oxytocin per institutional protocol reported a shorter time to delivery with the combination method. Other differences in protocols (e.g., inflation of the Foley balloon to 30 mL vs. 60 mL) and patient populations likely contributed to the differences in study outcomes. However, differences in labor management protocols across trials or within some multicenter trials may also influence the outcomes of labor induction studies, making it more difficult to compare treatment methods.

**Applicability to Clinical Practice in the U.S.: Clinical Trial Location**

Most RCTs identified in our search were international, with the majority conducted in South Asian (n = 15), West Asian (n = 11), European (n = 10), and African (n = 8) countries. Just 12 of the 63 RCTs were conducted within the U.S. Results from international and U.S. trials should be interpreted with caution because clinical practice guidelines in other countries (e.g., WHO, Society of Obstetricians and Gynaecologists of Canada [SOGC], Royal College of Obstetricians and Gynaecologists [RCOG], National Institute of Health and Care Excellence [NICE]) can differ from those in the U.S. where ACOG recommendations determine standard of care. For example, WHO and ACOG recommend the use of misoprostol or dinoprostone for induction of labor; while NICE guidelines prefer dinoprostone and recommend misoprostol only for women with intrauterine fetal death or in the context of a clinical trial. Although guidelines may recommend the use of a particular agent, other agents may be preferred due to availability or physician preference and experience. WHO guidelines, which are intended for a global audience, recommend the use of misoprostol or...
dineoprostone as first-line induction agents, yet oxytocin alone is the preferred method in many Latin American, African, and Asian countries. It is likely that demographic and cultural differences also influence treatment; the potential impact of these differences on trial outcomes is discussed in the “Trial population” section of this review.

**Applicability to Clinical Practice in the U.S.: Use of Combination Methods**

More than half of the clinical trials conducted in the U.S. (58%; \( n = 7/12 \)) evaluated combination methods of labor induction compared with just 22% of the international trials (\( n = 11/51 \)). The trials in the U.S. assessed the Foley catheter in combination with oxytocin or misoprostol versus various comparators. The Foley catheter plus oxytocin was compared with the Foley catheter alone (\( n = 1 \)). Sequential use of the Foley catheter and oxytocin (\( n = 2 \)) or a different oxytocin dosing protocol (i.e., fixed vs. incremental low-dose oxytocin; \( n = 1 \)). The Foley catheter plus vaginal misoprostol was compared with vaginal misoprostol alone (\( n = 1 \)); another clinical trial compared the Foley catheter plus oral misoprostol versus vaginal misoprostol alone. A separate trial compared four different treatment groups: the Foley catheter plus oxytocin versus the Foley catheter plus vaginal misoprostol versus the Foley catheter alone versus vaginal misoprostol alone. Despite the differences in trial design, all but one of the trials found that the combination method was more effective in inducing labor: the primary efficacy outcomes reported in these trials were time to delivery (\( n = 5 \)) or delivery within 24 hours (\( n = 2 \)). No significant differences in maternal complications or adverse neonatal outcomes between the combination methods and comparators were reported in any of the trials.

The international trials assessed a wider variety of combination methods, with the majority evaluating mechanical methods plus oxytocin or misoprostol: the Foley catheter plus oxytocin versus vaginal misoprostol (\( n = 2 \)), the Foley catheter plus vaginal misoprostol versus vaginal misoprostol or the Foley catheter alone (\( n = 3 \)), and double-balloon catheter plus oral misoprostol versus oral misoprostol alone (\( n = 2 \)). A few of the international trials assessed combinations of pharmacological methods, including the dinoprostone vaginal insert plus concurrent oxytocin versus sequential use of the dinoprostone insert and oxytocin (\( n = 1 \)), the dinoprostone vaginal insert with or without sequential oxytocin versus vaginal misoprostol plus sequential oxytocin (\( n = 1 \)), and vaginal misoprostol plus sequential oxytocin versus oxytocin alone (\( n = 2 \)). The primary outcomes reported in the international combination method trials included time to delivery (\( n = 8 \)) and vaginal delivery within 24 hours (\( n = 1 \) or 48 hours (\( n = 2 \)). Just more than half the trials (\( n = 6/11 \)) reported an improvement in time to vaginal delivery or vaginal delivery within 24 or 48 hours with the combination methods; however, two trials were not adequately powered. There were no significant differences in maternal complications or neonatal outcomes, except for a higher number of infants with 5-minute Apgar’s scores < 7 with the double-balloon catheter plus sequential oral misoprostol versus oral misoprostol alone (8 vs. 1; \( p = 0.04 \)), and lower incidences of uterine hyperstimulation (8 vs. 40%; \( p < 0.001 \)) and meconium-stained liquor (6 vs. 27%; \( p = 0.001 \)) with the Foley catheter plus vaginal misoprostol versus vaginal misoprostol alone.

It is unclear why combination methods (i.e., the Foley catheter plus oxytocin or misoprostol) were used more often in trials conducted within the U.S. As described previously, the use of oxytocin alone is the preferred method of labor induction in many regions outside the U.S., thus differences in the frequency and types of combination methods used in the U.S. and international trials may be reflective of different demographics and cultural preferences among clinical trial populations.

**Applicability to Clinical Practice in the U.S.: Use of Misoprostol**

The majority of the international trials (75%; \( n = 38/51 \)) included misoprostol as a comparator compared with 6 of the 12 trials conducted in the U.S. Most of the international trials assessed different routes of administration (e.g., oral vs. vaginal) or compared misoprostol with dinoprostone. Among the international trials that assessed misoprostol versus dinoprostone, nearly all compared vaginal misoprostol with vaginal dinoprostone gel or the vaginal dinoprostone insert; three trials compared vaginal misoprostol with intracervical dinoprostone gel. In most of these trials, misoprostol was associated with improved efficacy outcomes compared with dinoprostone, including a shorter time to delivery, more vaginal deliveries ≤ 12 or ≤ 24 hours, improved Bishop’s scores, and less requirement for oxytocin. Safety profiles were generally similar with misoprostol and dinoprostone, except for increased frequency of fetal heart rate abnormalities with misoprostol (11 vs 0%; \( p = 0.03 \)). It is important to note that many of these trials were not adequately powered for all efficacy and most safety outcomes.

Of the six U.S. trials that included misoprostol as a comparator, three assessed vaginal misoprostol in combination with the Foley catheter versus either vaginal or oral misoprostol alone or the Foley catheter alone, and another three trials assessed vaginal misoprostol versus oxytocin or the dinoprostone vaginal insert. The largest U.S. trials identified in the search (\( n = 1,308 \); \( n = 1,358 \)) were multicenter trials comparing vaginal misoprostol with vaginal dinoprostone. Compared with the vaginal dinoprostone insert, the 100 μg vaginal misoprostol insert was associated with a similar time to vaginal delivery (27.5 vs. 26.6 hours; \( p = 0.97 \)), while the 200 μg misoprostol vaginal insert significantly reduced time to vaginal delivery (32.8 vs. 21.5 hours; \( p < 0.001 \)). Cesarean delivery rates were similar with the dinoprostone insert and the 100 μg (25.4 vs. 27.8%; \( p = 0.64 \)) and 200 μg (27.1 vs. 26.0%; \( p = 0.05 \)) misoprostol insert doses. All treatments had a similar safety profile; however, the incidence of uterine tachysystole was higher in women receiving the 200 μg misoprostol vaginal insert compared with those receiving the dinoprostone insert (13.3 vs. 4.0%; \( p < 0.001 \)).
In summary, more international trials evaluated misoprostol compared with those conducted in the U.S. Misoprostol is widely used for labor induction both in the U.S. and globally, and in some countries misoprostol may be preferred over other pharmacological methods because it is less expensive and does not require refrigeration.91–93 Misoprostol can also be administered through various routes (oral, sublingual, vaginal), which may provide an advantage in hospitals or clinics with less experienced staff.91

**Trial Population**

The majority of the RCTs included in our search used vague enrollment criteria and typically included a broad population of women presenting at term or postterm for induction of labor, regardless of indication. Required gestational ages ranged from 24 to 42 weeks. Required Bishop’s scores ranged from < 4 to < 8; some trials specified a cervical dilation of < 2 or < 3 cm. Of note, lower Bishop’s scores (e.g., < 5) and limited dilation at the start of induction have been associated with increased risk of cesarean section following labor induction.94–96

Although trials conducted exclusively in women with PROM and pPROM or previous cesarean sections were excluded from our search, a few of the RCTs with very broad patient inclusion criteria included women with ruptured membranes or prior cesarean delivery. A few trials limited the trial population to those with specific indications for labor induction (e.g., hypertension, diabetes, oligohydramnios, postterm pregnancy); thus the results from these trials may be different compared with those that included a broader population of women requiring labor induction. Most trials recruited both nulliparous and multiparous (low and high parity) women; however, very few were powered to detect differences in outcomes between the two groups.

The heterogeneous patient populations in the labor induction trials make it difficult to compare outcomes between trials. Treatment efficacy and safety can vary across the patient population and many factors (e.g., age, parity, indication, medical history) can differentially affect outcomes. Furthermore, as described previously, patient demographic and cultural preferences vary between trial populations. Labor induction rates can also vary; rates in Africa and Asia, where more than half of the RCTs identified in this search were conducted, are approximately 4 and 12%, respectively, compared with approximately 24% in the U.S.1,88 These differences can be attributed to patient and physician factors, as well as availability of adequate facilities and monitoring and the ability to perform a safe cesarean delivery.88 Differences among populations within countries may also influence labor induction outcomes. A post hoc analysis of data from a trial comparing the misoprostol vaginal insert with the dinoprostone vaginal insert showed differences in cesarean delivery rates, indications for cesarean delivery, birth weights, and incidences of postpartum hemorrhage among women who identified themselves as black, white, or Hispanic.97 Thus, heterogeneity in induction practices and resources, as well as potential disparities based on race and ethnicity, should be considered when interpreting findings from trials of labor induction.

**Outcomes Assessments: Efficacy and Safety**

There was wide variability in the reporting of efficacy outcomes between the trials included in our search (Table 3). The most common primary outcome reported was time from the start of induction to delivery (vaginal or cesarean; n = 28 trials). Other frequently reported efficacy outcomes were delivery within 24 hours, cesarean deliveries, time to active labor, mode of delivery, and the need for oxytocin; most trials were not adequately powered to assess significant differences in these outcomes. Many trials considered an induction successful if vaginal delivery occurred within 24 hours, yet this timeline may not be realistic considering that the latent phase of labor is longer in women undergoing induction versus spontaneous labor.98,99 ACOG recommends longer duration of the latent phase (up to 24 hours or longer) with oxytocin administration for at least 12 to 18 hours after

<table>
<thead>
<tr>
<th>Table 3 Common presupposed primary and secondary outcome measures in published labor induction trials (all RCTs)23–85</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcomes</strong></td>
</tr>
<tr>
<td>• Time to delivery (n = 21)</td>
</tr>
<tr>
<td>• Time to vaginal delivery (n = 7)</td>
</tr>
<tr>
<td>• Delivery ≤ 24 h (n = 6)</td>
</tr>
<tr>
<td>• Vaginal delivery ≤ 24 h (n = 6)</td>
</tr>
<tr>
<td>• Rate of cesarean delivery (n = 6)</td>
</tr>
<tr>
<td>• Rate of vaginal delivery (n = 4)</td>
</tr>
<tr>
<td>• Time to active labor (n = 4)</td>
</tr>
<tr>
<td>• Umbilical artery cord pH (n = 9)</td>
</tr>
<tr>
<td>• Mode of delivery (n = 15)</td>
</tr>
<tr>
<td>• Fetal heart rate abnormalities (n = 13)</td>
</tr>
<tr>
<td>• Gastrointestinal symptoms (nausea, vomiting, diarrhea, etc.; n = 12)</td>
</tr>
<tr>
<td>• Change in Bishop’s score (n = 8)</td>
</tr>
<tr>
<td>• Delivery ≤ 24 h (n = 8)</td>
</tr>
<tr>
<td>• Fetal or neonatal death (n = 8)</td>
</tr>
<tr>
<td>• Delivery ≤ 12 h (n = 7)</td>
</tr>
<tr>
<td>• Endometritis (n = 7)</td>
</tr>
<tr>
<td>• Vaginal delivery ≤ 24 h (n = 7)</td>
</tr>
<tr>
<td>• Indication for cesarean delivery (n = 7)</td>
</tr>
<tr>
<td>• Maternal fever (n = 6)</td>
</tr>
<tr>
<td>• Vaginal delivery ≤ 12 h (n = 6)</td>
</tr>
<tr>
<td>• Failed induction (n = 5)</td>
</tr>
</tbody>
</table>

Abbreviations: NICU, neonatal intensive care unit; RCT, randomized controlled trial.
Limitations of Labor Induction Trials

Viteri, Sibai

membrane rupture and a reassuring nonstress test before an induction is considered a failure. Long labor can help prevent unnecessary cesarean deliveries which is particularly important in nulliparous women because lowering the rate of primary cesarean delivery rates will in turn reduce repeat cesarean delivery rates.

Nearly all trials included in the search reported at least some safety findings; however, many trials did not include a comprehensive assessment of safety outcomes. In many trials, it was unclear whether any safety outcomes were prespecified. Commonly reported prespecified safety outcomes included maternal and fetal complications (e.g., tachysystole/uterine hyperstimulation, abnormal fetal heart rate tracings, uterine rupture, postpartum hemorrhage) and neonatal outcomes (e.g., Apgar's scores, arterial cord pH, neonatal intensive care unit [NICU] admission). These safety outcomes were identified as secondary outcomes in the majority of trials and not adequately powered.

Outcomes Assessments: Inconsistent Definitions

A major limitation of all trials identified in our search was that the definitions of many efficacy and safety parameters (e.g., time to labor, time to delivery, failed induction, uterine hyperstimulation) were inconsistent between trials. For example, most trials defined tachysystole as more than five uterine contractions in a 10-minute period, either with or without fetal heart rate abnormalities; however, other trials defined tachysystole as more than five contractions in 5 minutes, more than 12 contractions in 20 minutes, or failed to define the term altogether. The criteria for failed labor induction also varied between trials, as failed induction was commonly defined as failure to progress to the active phase of labor 12, 24, or 48 hours after treatment initiation; failure to reach the active phase after rupture of membranes and 10 or 12 hours of oxytocin infusion; failure to progress to the active phase despite adequate contractions; or unfavorable cervix following labor induction protocol. Inconsistent definitions for these and other efficacy and safety parameters limit the ability to compare outcomes across trials.

Outcomes Assessments: Clinical Relevancy

Another major limitation was that in most trials, the primary outcomes were not always the most clinically relevant (e.g., time to delivery, delivery within 24 hours). Duration from induction to delivery might have clinical significance among those with medically indicated delivery before 37 weeks, yet outcomes such as the rate of cesarean delivery or severe maternal or neonatal morbidity are more relevant measures of efficacy and safety. WHO guidelines for labor induction consider cesarean section, severe maternal morbidity or death, perinatal death, and serious neonatal morbidity as top priority outcomes for labor induction trials. The goal of labor induction should be vaginal delivery because cesarean deliveries are associated with increased risk to the mother and infant, as well as higher health care costs. Of the 63 RCTs we reviewed, just six used cesarean section rate as a primary outcome measure; however, two of the trials were not adequately powered.

Outcomes Assessments: Statistical Power

As noted previously, nearly all RCTs identified in our search were underpowered to detect differences in secondary efficacy and safety outcomes due to small sample sizes. Much larger trial populations are needed to identify statistically significant differences in rare maternal and neonatal adverse events. For example, a clinical trial comparing oral misoprostol with the Foley catheter assessed the noninferiority of misoprostol with respect to the composite outcome of neonatal asphyxia or postpartum hemorrhage. Assuming that the composite outcome would occur in 13.7% of patients in the misoprostol group and 12.7% of patients in the Foley catheter group, it was determined that 1,860 patients (930 women per group) were needed to provide 80% power to demonstrate noninferiority of misoprostol (prespecified margin of 5%). A recently completed well-designed clinical trial assessing labor induction versus expectant management in nulliparous women (ARRIVE) enrolled more than 6,000 patients to detect differences in a composite of several severe neonatal morbidity and perinatal mortality outcomes, further highlighting the need for an adequate sample size to ensure sufficient statistical power. Assuming 92.5% adherence to the assigned protocol (i.e., labor induction or expectant management) and that the composite outcome of adverse perinatal events would occur in 3.5% of patients assigned to induction of labor, it was determined that at least 6,000 women were needed to provide 85% power to detect a 40% difference in the primary outcome. While obtaining a sufficiently large sample size is important, it may not always be possible to achieve, given the significant amount of time and resources needed to conduct such a study and this was noted as a limitation by the authors of several published studies.

There are also challenges in defining appropriate assumptions to adequately power trials. In most of the trials included in our search, sample size calculations were based on assumptions derived from previous trials with different trial populations and in some cases, different comparators. Other trials based their assumptions on the investigators’ own experience at their institution. Thus, the criteria for demonstrating superiority or noninferiority were fairly subjective.

Outcomes Assessments: Subgroup Analyses

In addition, of the five trials identified in our search that included prespecified subgroup analyses, only two were powered to detect differences between groups. While subgroup analyses can be useful for evaluating treatment effects in specific patient populations (e.g., nulliparous vs. multiparous women), they should be identified a priori rather than data driven. Furthermore, each subgroup needs to have a large enough sample size to achieve adequate power, which may not always be practical or feasible.

Outcomes Assessments: Patient Satisfaction

Last, in addition to efficacy and safety, the ideal method for labor induction should minimize patient pain and discomfort. Patient pain and/or satisfaction with the labor induction procedure was assessed in just nine trials (none were
conducted in the U.S.

None of the trials used a validated patient-reported outcome instrument for labor induction; however, until very recently, there have been no validated questionnaires for labor induction. While WHO and NICE guidelines recommend that patient needs and preferences be taken into consideration when selecting a method of labor induction, ACOG guidelines do not specify that maternal preference be considered. In cases, in which labor induction methods have similar efficacy and safety profiles, we recommend that physicians and health care providers take into account patient preference and overall labor experience.101,102

Conclusions and Recommendations for Future Clinical Trial Designs

The optimal methods for labor induction remain unclear due to the significant limitations in the published literature on labor induction. For many trials, the reported conclusions were not supported by the results. Wide variation in trial design, poorly defined inclusion/exclusion criteria, inconsistent definitions and reporting of outcomes, lack of clinically relevant outcomes, and inadequate statistical power limit the ability for evidence-driven treatment decisions. The majority of trials focused on reducing the duration of labor induction and included primary outcomes related to time to delivery, delivery within 12 to 24 hours, or time to active labor. Clinically important outcomes, including the rate of cesarean delivery and nearly all potential maternal and fetal complications (e.g., uterine hyperstimulation, uterine rupture, abnormal fetal heart rate, postpartum hemorrhage) and neonatal outcomes (e.g., Apgar’s scores, NICU admission) were consistently underpowered. In addition, because most of the trials were international, their findings may not be generalizable to many U.S. populations that may differ in terms of age, body mass index (BMI), race/ethnicity, preference for certain labor induction methods, and labor and delivery practices. Regulatory requirements for the design, conduct, and oversight of clinical trials also varied between countries; it may be difficult to ascertain the validity of results for studies conducted without a governing body to provide standardization and audit. It is important to note that, although the conclusions of this narrative review were based on evidence, as well as our own perspectives and experiences in clinical practice, a systematic review would provide a more rigorous, quantitative assessment of the literature with minimal bias and a comprehensive evaluation of study quality.

Ideally, we recommend that future trials of labor induction methods should be head-to-head RCTs with populations large enough to adequately power primary and secondary outcomes, as well as prespecified subgroup analyses (e.g., parity, BMI, age). Ideally, trials should compare agents with similar mechanisms of action, and combination methods should be compared with the same single agents administered using the same route and dosage. Trials should be multicenter trials with clear unambiguous inclusion/exclusion criteria and consistent protocols for induction. If feasible, trials should be double-blinded but this may not be possible depending on the methods of labor induction used. All physicians and nurses should follow the guidelines established by ACOG and the Society for Maternal-Fetal Medicine for the safe prevention of primary cesarean delivery.99

Future labor induction trials should clearly define efficacy and safety outcomes and use consistent definitions between trials. Trials should report on clinically relevant outcomes because the downstream effects of induction (i.e., cesarean section, maternal and neonatal complications) are more important in determining efficacy and safety. The primary outcome of all labor induction trials should be the rate of cesarean delivery. Cesarean deliveries are associated with increased risk of adverse maternal and neonatal outcomes and are a significant burden to health care systems.99,100 The most effective labor induction methods are those that allow safe vaginal delivery, minimizing risk to the mother and infant.

Addressing all these limitations, future clinical trial designs for labor induction will be challenging and may not be possible in some cases. Detecting significant differences in rare adverse events can require thousands of patients and may not be achievable. The management of active labor and delivery based on standard institutional practices will invariably affect all labor induction studies as they are most likely to be pragmatic in nature. Another unavoidable limitation is that many non-clinical factors (e.g., patient preference, obstetrician experience) cannot be regulated through protocol and likely impact labor induction outcomes. Although certain limitations of published labor induction trials may be inevitable, identifying these limitations will hopefully lead to improved trial designs with outcomes that can be better applied to real-world obstetric practice. Given that the best method of labor induction is still being debated and may vary based on indication and other patient characteristics; we recommend that health care providers need to be trained in all labor induction methods to allow for a more personalized approach to labor induction that considers the unique needs and preferences of the patient.

Conflict of Interest
The authors have nothing to disclose.

Acknowledgments
Medical writing and editorial assistance for the development of this manuscript were provided by Kimberly Fuller, PhD, of SciFluent Communications, and were financially supported by Ferring Pharmaceuticals, Inc.

References
15 Wood S, Cooper S, Ross S. Does induction of labour increase the risk of caesarean section? A systematic review and meta-analysis of trials in women with intact membranes. BJOG 2014;121(06):674–685, discussion 685
19 West HM, Jozwiak M, Dodd JM. Methods of term labour induction for women with a previous caesarean section. Cochrane Database Syst Rev 2017;6:CD009792
Limitations of Labor Induction Trials


49 Carbone JF, Tuuli MG, Fogerty PJ, Roehl KA, Macones GA. Combination of Foley bulb and vaginal misoprostol compared with vaginal misoprostol alone for cervical ripening and labor induction: a randomized controlled trial. Obstet Gynecol 2013;121(2, Pt 1):247–252


63 Gungörük K, Yıldırım G, Gungörük O, Ark C, Tekirdağ İ. Sustained-release dinoprostone vaginal pessary with concurrent high-dose oxytocin infusion compared to sustained-release dinoprostone vaginal pessary followed 6 h later by high-dose oxytocin infusion for labor induction in women at term with unfavorable cervix: a randomized controlled trial. Gynecol Obstet Invest 2011;71(03):32–40


71 Tan TC, Yan SY, Chua TM, Biswas A, Chong YS. A randomised controlled trial of low-dose misoprostol and dinoprostone vaginal pessaries for cervical priming. BJOG 2010;117(10):1270–1277


77 Calder AA, Loughney AD, Weir CJ, Barber JW. Induction of labour in nulliparous and multiparous women: a UK, multicentre, open-label study of intravaginal misoprostol in comparison with dinoprostone. BJOG 2008;115(10):1279–1288
84 Wing DA; Misoprostol Vaginal Insert Consortium. Misoprostol vaginal insert compared with dinoprostone vaginal insert: a randomized controlled trial. Obstet Gynecol 2008;112(04):801–812
100 Truven Health Analytics. The cost of having a baby in the United States. 2013