Cervical Ripening in the Outpatient Setting

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

and

Patient-Centered Outcomes Research Institute
1828 L Street, NW, Ste. 900
Washington, DC 20036
www.pcori.org

This information is distributed solely for the purposes of predissemination peer review. It has not been formally disseminated by the Agency for Healthcare Research and Quality. The findings are subject to change based on the literature identified in the interim and peer-review/public comments and should not be referenced as definitive. It does not represent and should not be construed to represent an Agency for Healthcare Research and Quality or Department of Health and Human Services (AHRQ) determination or policy.

Contract No. To be added for final version.

Prepared by: To be added for final version.

Investigators: To be added for final version.
This report is based on research conducted by the XXXXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The Patient-Centered Outcomes Research Institute (PCORI) funded the report (PCORI Publication No. 2020-SR-03). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ or PCORI. Therefore, no statement in this report should be construed as an official position of PCORI, AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report is made available to the public under the terms of a licensing agreement between the author and the Agency for Healthcare Research and Quality. This report may be used and reprinted without permission except those copyrighted materials that are clearly noted in the report. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

PCORI, AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies, may not be stated or implied.

This report may periodically be assessed for the currency of conclusions. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program website at www.effectivehealthcare.ahrq.gov. Search on the title of the report.

People using assistive technology may not be able to fully access information in this report. For assistance contact EPC@ahrq.hhs.gov.


ii
Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new healthcare technologies and strategies.

The Patient-Centered Outcomes Research Institute (PCORI) was established to fund research that can help patients and those who care for them make better informed health care choices. PCORI partnered with AHRQ to help fulfill its authorizing mandate to engage in evidence synthesis and make comparative effectiveness research results more available to patients and providers.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the healthcare system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the website (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

If you have comments on this systematic review, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Gopal Khanna, M.B.A.
Director
Agency for Healthcare Research and Quality

Arlene Bierman, M.D., M.S.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Nakela Cook, M.D., M.P.H.
Executive Director
Patient-Centered Outcomes Research Institute

Jean Slutsky, P.A., M.S.P.H.
Chief Engagement and Dissemination Officer
Patient-Centered Outcomes Research Institute

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Jill Huppert, M.D.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Michelle Althuis, Ph.D.
Associate Director
Research Synthesis and New Technology
Patient-Centered Outcomes Research Institute

Jennie Dalton Bowen, M.P.H.
Program Officer
Research Synthesis and New Technology
Patient-Centered Outcomes Research Institute
**Acknowledgments**
To be added for final version.

**Key Informants**
In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants will be provided in the final report.

**Technical Expert Panel**
In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts will be provided in the final report.

**Peer Reviewers**
Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report does not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers will be provided in the final report.
Cervical Ripening in the Outpatient Setting

Structured Abstract

**Objectives.** To assess the comparative effectiveness and potential harms of cervical ripening in the outpatient setting (versus inpatient, versus other outpatient intervention) and of fetal surveillance when a prostaglandin was used for cervical ripening.

**Data sources.** Electronic databases (Ovid® MEDLINE®, Embase®, CINAHL®, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews) to March 2020; reference lists; and a Federal Register notice.

**Review methods.** Using predefined criteria and dual review, we selected randomized controlled trials (RCTs) and cohort studies of cervical ripening comparing prostaglandins and mechanical methods in outpatient versus inpatient settings; one outpatient method versus another (including placebo or expectant management); different methods/protocols for fetal surveillance in cervical ripening using prostaglandins. When data from similar study designs, populations, and outcomes were available, random effects using profile likelihood meta-analyses were conducted. Inconsistency (using $I^2$) and small sample size bias (publication bias; if ≥7 studies) was assessed. Strength of evidence (SOE) was assessed. All review methods followed AHRQ EPC Methods guidance.

**Results.** We included 30 RCTs and 9 cohort studies (72% fair-quality) involving 9,465 women. The evidence is most applicable to women aged 25 to 30 years, with singleton, vertex presentation, and low-risk pregnancies. No studies on fetal surveillance were found. For studies that compared outpatient to inpatient settings for cervical ripening, the frequency of cesarean delivery was not significantly different using dinoprostone (2 RCTs, 1 cohort study, SOE: low). The frequency of infection, hypoxic-ischemic encephalopathy, meconium aspiration syndrome, and postpartum hemorrhage requiring transfusion was low overall, and not significantly different in outpatient versus inpatient cervical ripening using dinoprostone (2 RCTs, 4 cohort studies, SOE: low). In comparisons of outpatient versus inpatient single-balloon catheters for cervical ripening (3 RCTs, 2 cohort studies), and for outpatient catheters versus inpatient dinoprostone (1 double-balloon RCT, 1 single-balloon RCT), differences between groups on all outcomes were small and not statistically significant (SOE: low). Evidence on misoprostol, double-balloon catheters and hygroscopic dilators was insufficient to draw conclusions for outpatient versus inpatient settings.

In the outpatient setting, head-to-head comparisons found the frequency of cesarean delivery or receiving antibiotics for suspected uterine infection was not significantly different for dinoprostone gel 2.5 mg versus 5 mg, or for latex versus silicone single-balloon catheters (1 RCT each, SOE: low). Differences between prostaglandin and placebo for cervical ripening were small and not significantly different for cesarean delivery (7 RCTs), fetal encephalopathy (1 RCT), birth trauma (3 RCTs), or uterine infection (7 RCTs) (SOE: low). These findings did not change according to the specific prostaglandin, route of administration, study quality, or gestational age. Small, nonsignificant differences in the frequency of cesarean delivery and uterine infection were also found between dinoprostone versus either membrane sweeping or expectant management (6 RCTs, SOE: low). These findings did not change according to the specific prostaglandin or study quality.
For all comparisons, there was insufficient evidence on many important outcomes such as perinatal mortality and time from admission to vaginal birth. Limitations of the evidence include the quantity, quality and sample sizes of trials for specific interventions, particularly rare harm outcomes.

**Conclusions.** In women with low-risk pregnancies, the risk of cesarean delivery and fetal, neonatal, or maternal harms using either dinoprostone or single-balloon catheters were not significantly different for cervical ripening in the outpatient versus inpatient setting, and similar when compared with placebo, expectant management, or membrane sweeping in the outpatient setting.
# Contents

Introduction ...................................................................................................................................... i
Background ..................................................................................................................................... 1
Purpose and Scope of the Systematic Review ............................................................................ 2
Methods......................................................................................................................................... 3
Review Approach........................................................................................................................ 3
   Key Questions......................................................................................................................... 3
      Analytic Framework ............................................................................................................... 4
Study Selection ............................................................................................................................ 4
Data Extraction and Risk of Bias Assessment ......................................................................... 5
Data Synthesis and Analysis ..................................................................................................... 5
Grading the Strength of the Body of Evidence ....................................................................... 6
Results......................................................................................................................................... 7
   Description of Included Evidence ............................................................................................. 7
   Key Question 1. Comparative effectiveness and harms of cervical ripening using prostaglandins in the outpatient versus inpatient setting ............................................................ 8
      Key Points............................................................................................................................... 8
      Summary of Findings................................................................................................................ 9
   Key Question 2. Comparative effectiveness and harms of cervical ripening using mechanical methods in the outpatient versus inpatient setting .......................................................... 15
      Key Points............................................................................................................................... 15
      Summary of Findings................................................................................................................ 15
   Key Question 3. Comparative effectiveness and harms of cervical ripening in the outpatient setting ........................................................................................................................................ 24
      Key Points............................................................................................................................... 24
      Summary of Findings................................................................................................................ 24
   Key Question 4. Comparative effectiveness and harms of different methods and protocols for fetal surveillance in pregnant women undergoing cervical ripening with prostaglandins .................................................................................................................. 39
      Summary Tables: Evidence for Efficacy/Effectiveness and Harms ......................................... 40
      Insufficient Evidence ............................................................................................................. 43
Discussion ..................................................................................................................................... 44
   Findings in Relation to the Decisional Dilemma(s)................................................................. 44
   Strengths and Limitations ........................................................................................................ 47
   Applicability ............................................................................................................................. 48
   Implications for Clinical Practice, Education, Research, or Health Policy ....................... 50
      Considerations for Clinical Practice .................................................................................... 50
      Research Recommendations ............................................................................................... 50
      Conclusions ........................................................................................................................... 51
References..................................................................................................................................... 52
Abbreviations and Acronyms ...................................................................................................... 58
Tables
Table A. Primary birth-related efficacy/effectiveness outcomes.................................................... 2
Table B. Primary fetal harms outcomes.......................................................................................... 3
Table C. Primary maternal harms outcomes................................................................................... 3
Table 1. Criteria for population, intervention, comparison, and outcomes of eligible studies a,b ... 5
Table 2. Baseline characteristics of women enrolled in included studies................................. 8
Table 3. Secondary maternal harms outcomes: outpatient versus inpatient cervical ripening with dinoprostone.......................................................................................................................... 13
Table 4. Subgroup analyses of cesarean delivery frequency according to parity (prostaglandin versus placebo)............................................................................................................................ 30
Table 5. Primary birth-related efficacy/effectiveness outcomes.................................................... 40
Table 6. Primary fetal harms outcomes ..................................................................................... 40
Table 7. Primary maternal harms outcomes ............................................................................. 42

Figures
Figure 1. Analytic framework....................................................................................................... 4
Figure 2. Meta-analysis of cesarean delivery with prostaglandins for cervical ripening: outpatient versus inpatient .......................................................................................................................... 11
Figure 3. Meta-analysis of cesarean delivery with catheters for cervical ripening: outpatient versus inpatient ........................................................................................................................................ 20
Figure 4. Meta-analysis of cesarean delivery with outpatient catheters versus inpatient dinoprostone for cervical ripening ........................................................................................................ 23
Figure 5. Meta-analysis of cesarean delivery prostaglandins versus placebo for cervical ripening in the outpatient setting ........................................................................................................ 29
Figure 6. Meta-analysis of cesarean delivery with prostaglandins versus expectant management and membrane sweeping for cervical ripening in the outpatient setting ........................................... 36

Appendixes
Appendix A. Introduction
Appendix B. Methods
Appendix C. Results
Appendix D. Included Studies
Appendix E. Evidence Tables
Appendix F. Forest Plots
Appendix G. Risk of Bias Assessment
Appendix H. Strength of Evidence
Appendix I. Excluded Studies
Appendix J. Appendix References
Evidence Summary

Main Points

- Low strength evidence suggests that outpatient cervical ripening with dinoprostone (intravaginal insert or intracervical gel) or single-balloon catheters (30–50 ml fill) were not significantly different for cesarean delivery and fetal/neonatal or maternal infection in comparison with the same intervention in the inpatient setting.
- The evidence on outpatient cervical ripening with misoprostol, double-balloon catheters, or hygroscopic dilators was insufficient.
- Low strength evidence suggests that the risk of cesarean delivery with dinoprostone intracervical gel 2.5 mg versus 5.0 mg, and with silicone versus latex single-balloon catheters in the outpatient setting was not significantly different. Evidence was insufficient to draw conclusions on other outcomes or other direct comparisons of interventions.
- Low strength evidence suggests that the risk of cesarean delivery with prostaglandins (dinoprostone or misoprostol) compared with placebo, expectant management, and membrane sweeping in the outpatient setting was not significantly different. The incidence of hypoxic-ischemic encephalopathy, meconium aspiration syndrome, postpartum hemorrhage and uterine infection, primarily with dinoprostone versus placebo were not significantly different.
- There was no evidence comparing different mechanical methods with each other, with membrane sweeping or with expectant management in the outpatient setting.
- For all comparisons, there was insufficient evidence on time from admission to vaginal birth, perinatal mortality, fetal/neonatal intracranial or subgaleal hemorrhage, and maternal hemorrhage requiring transfusion.
- Comparative evidence on fetal surveillance for cervical ripening with a prostaglandin was not found.

Background and Purpose

The purpose of this review is to assess the comparative effectiveness and potential harms of cervical ripening in the outpatient versus the inpatient setting. The intended audience includes the American College of Obstetricians and Gynecologists’ (ACOG) guideline developers, clinicians who deliver neonates (e.g., obstetricians, nurse midwives, family physicians), other personnel who administer and monitor cervical ripening, and health system policymakers. In addition to these clinical implications, we hope to inform the future research necessary to provide high-quality, evidence-based care to all pregnant women.

Methods

We employed methods consistent with those outlined in the Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program Methods Guidance (https://effectivehealthcare.ahrq.gov/products/cer-methods-guide/overview), and we describe these in the full report. Our searches covered publication dates up to March 2020.
Results

We included 39 mostly fair-quality studies (30 RCTs, 9 cohort studies), with 9,465 women. The majority of the evidence pertained to KQ 3 (22 RCTs). Participants’ mean age was 28.7 years, most were nulliparous, mean baseline Bishop score was 3.4 and gestational age was 40.6 weeks. Most studies exclude women with prior cesarean delivery, but few studies excluded women with diabetes or hypertension of any type. Postterm pregnancy was the most common reason for cervical ripening. Tables A–C summarize our findings; the full report provides more outcomes and details. If a prespecified, primary outcome is not listed in a table below that means that no study reported on that outcome (e.g., time from admission to vaginal delivery) or the evidence was insufficient to draw conclusions (i.e., due to imprecise estimates [too few patients or events], lack of corroborating evidence [a single study], and study limitations). The highest strength of evidence found for any outcome was low-strength. No studies that met inclusion criteria were identified for KQ4.

Table A. Primary birth outcome: Cesarean delivery

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Findingsa</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
<th>I² for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Question 1: Prostaglandin Outpatient vs. Inpatient</td>
<td>Dinoprostone outpatient vs. inpatient</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=1,120)</td>
<td>23% vs 23%</td>
<td>RR 0.97 (0.75 to 1.25), I²= 0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Cohort studies (n=2,511)</td>
<td>33% vs. 33%</td>
<td>RR 0.79 (0.67 to 0.98), I²=0%</td>
<td></td>
</tr>
<tr>
<td>Key Question 2: Mechanical Method Outpatient vs. Inpatient</td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>3 RCTs (n=370)</td>
<td>12% vs. 20%</td>
<td>RR 0.59 (0.21 to 1.03), I²=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Cohort studies (n=1,057)</td>
<td>33% vs. 30%</td>
<td>RR 0.95 (0.72 to 1.22), I²=0%</td>
<td></td>
</tr>
<tr>
<td>Key Question 3: Outpatient Comparison of Methods</td>
<td>Outpatient catheter vs. inpatient dinoprostone</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=549)</td>
<td>33% vs. 26%</td>
<td>RR 1.24 (0.88 to 1.70), I²=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dinoprostone gel 2.5 mg vs. 5.0 mg</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=116)</td>
<td>20% vs. 19%</td>
<td>RR 0.90 (0.44 to 1.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandin vs. placebo</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>12 RCTs (n=924)</td>
<td>16% vs. 21%</td>
<td>RR 0.80 (0.58 to 1.09), I²=4.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandin vs. expectant management</td>
<td>Low strength of evidence of little or no difference</td>
<td>4 RCTs (n=615)</td>
<td>27% vs. 26%</td>
<td>RR 0.95 (0.68 to 1.33), I²=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dinoprostone vs. membrane sweeping</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>3 RCTs (n=339)</td>
<td>22% vs. 15%</td>
<td>RR 1.44 (0.85 to 2.36), I²=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silicone vs. latex Single-balloon catheters</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=534)</td>
<td>39% vs. 40%</td>
<td>RR 0.98 (0.80 to 1.22)</td>
<td></td>
</tr>
</tbody>
</table>

a Difference of < 5% = little or no difference; 5% to 10% = small difference; 11% to 20% = moderate difference; >20% = large difference.

CI = confidence interval; RCT = randomized controlled trial; RR = relative risk
### Table B. Primary fetal harms outcomes

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Outcome</th>
<th>Findings&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
<th>I&lt;sup&gt;2&lt;/sup&gt; for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 1:</strong> Prostaglandin Outpatient vs. Inpatient</td>
<td>Dinoprost</td>
<td>Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=1,120)</td>
<td>4% vs. 3%</td>
<td>RR 1.39 (0.67 to 3.03), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Birth Trauma</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=129)</td>
<td>2% vs. 3%</td>
<td>RR 0.49 (0.05 to 5.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Shoulder dystocia</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>1 RCT (n=129)</td>
<td>3% vs. 11%</td>
<td>RR 0.28 (0.06 to 1.30)</td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 2:</strong> Mechanical Method Outpatient vs. Inpatient</td>
<td>Dinoprost vs. placebo</td>
<td>Meconium Aspiration Syndrome&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=134)</td>
<td>2% vs. 4%</td>
<td>RR 0.76 (0.03 to 22.33), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dinoprost vs. placebo</td>
<td>Birth Trauma</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>3 RCTs (n=270)</td>
<td>6% vs. 0.69%</td>
<td>RR 7.88 (0.98 to 63.20)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Difference of <1% = little or no difference; >1% to 3% = small difference; >3% to 8% = moderate difference; >8% = large difference

<sup>b</sup> Neonatal intensive care unit (NICU) admission, not specified as the Syndrome

CI = confidence interval; RCT = randomized controlled trial; RR = relative risk

### Table C. Primary maternal harms outcomes

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Outcome</th>
<th>Findings&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
<th>I&lt;sup&gt;2&lt;/sup&gt; for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 2:</strong> Mechanical Method Outpatient vs. Inpatient</td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Uterine Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=259)</td>
<td>5% vs. 5%</td>
<td>RR 0.33 (0.36 to 2.74), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outpatient catheter vs. inpatient dinoprostone</td>
<td>Postpartum Hemorrhage</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>2 RCTs (n=549)</td>
<td>28% vs. 22%</td>
<td>RR 1.10 (0.83 to 1.45), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 3:</strong> Outpatient Comparison of Methods</td>
<td>Prostaglandins vs. placebo</td>
<td>Uterine Infection</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>7 RCTs (n=771)</td>
<td>7% vs. 5%</td>
<td>RR 0.75 (0.40 to 1.39), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandins vs. expected management</td>
<td>Uterine Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=294)</td>
<td>6% vs. 5%</td>
<td>RR 1.21 (0.45 to 3.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandins vs. membrane sweeping</td>
<td>Uterine Infection</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=269)</td>
<td>7% vs. 4%</td>
<td>RR 1.22 (0.56 to 2.75), I&lt;sup&gt;2&lt;/sup&gt;=0%</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Difference of <1% = little or no difference; >1% to 3% = small difference; >3% to 8% = moderate difference; >8% = large difference

CI = confidence interval; RCT = randomized controlled trial; RR = relative risk
Strengths and Limitations

The evidence comparing interventions in the outpatient and inpatient settings suffers from too few RCTs and too small of sample sizes (range 48 to 827; mean 172), particularly when assessing harms that are rare. Evidence quantity and quality is low for specific interventions. These are: 1) misoprostol and double-balloon catheters (comparing each in the outpatient versus inpatient settings), 2) direct comparisons of single- and double-balloon catheters, 3) hygroscopic dilators and 4) the various formulations and routes of administration of dinoprostone or misoprostol. These studies enrolled narrowly defined populations and did not analyze effects in important subgroups such as women over 30 or 35, Group B Streptococcus (GBS) status, diabetes, hypertension, fetal growth restriction, and gestational age categories. The studies generally either excluded women with such characteristics, or failed to report on them in detail. There was variation in outcome definition and reporting across the studies, with many reporting outcomes not defined as specified in the protocol for this review. Differences in rare harms, such as hypoxic-ischemic encephalopathy, would require much larger studies (i.e., statistical power) than are currently available.

Implications and Conclusions

This report can provide more guidance to clinicians and pregnant women on the relative benefits and harms of outpatient cervical ripening. This report found low strength of evidence that outpatient cervical ripening with dinoprostone and single-balloon catheters does not impose increased risk of cesarean delivery. We also found no indications of important signals of increased risk of fetal/neonatal and maternal harms, although not all such harms were adequately studied. The evidence is most applicable to younger women with singleton, vertex presentation pregnancies and low or no obstetric or medical risk factors. It does not identify the characteristics of pregnant women and fetuses that will benefit most or have the lowest risk of harm. There is evidence that women prefer, and were satisfied with, outpatient cervical ripening, although the decision-making process is complex. Filling the gaps in the evidence will require RCTs with sample sizes large enough to evaluate important harms; that evaluate important subgroups of the population; and study outpatient misoprostol, double-balloon catheters. Observational studies should be prospective and use appropriate methods to control for confounding and effect modification.
Introduction

Background

Induction of labor (IOL) is the process of initiating labor by using medications, mechanical methods (devices), or other techniques, with the goal of achieving vaginal birth with minimal risks.\(^1\) IOL has shown maternal/child benefit when the health of a pregnant woman or fetus is at risk (e.g., maternal hypertension or diabetes, fetal growth restriction, and in postterm pregnancies).\(^2,3\) In addition to these medically-indicated deliveries, IOL is also done on an elective basis; reasons include scheduling at the request of pregnant women or to ensure availability of appropriate providers.\(^4\) A recent large randomized study of low-risk nulliparous pregnant women (the ARRIVE trial)\(^5\) demonstrated that induction of labor at 39 weeks, compared with expectant management, resulted in lower cesarean delivery rates and no difference in neonatal outcomes.\(^2,3\) IOL rates are rising dramatically in the United States, reaching 25.7 percent in 2017,\(^6\) with tertiary care centers that provide high-risk obstetric care, such as Oregon Health & Science University (OHSU), reaching 46.5 percent in 2019.\(^7\) Labor induction occurs in approximately one-quarter of term pregnancies, with estimates of 77 to 85 percent due to medical indications.\(^8-10\)

Cervical ripening, often an initial component of labor induction, is the process of softening and effacing the cervix as well as stimulating early cervical dilation. Based on data from trials of labor induction, approximately 83 to 85 percent of women with an indication for induction require cervical ripening.\(^11,12\) Common cervical ripening methods include pharmacologic options, such as prostaglandins (misoprostol and prostaglandin E2), and mechanical options, such as inserting a balloon catheter or hygroscopic dilator into the endocervix. See Appendix A for descriptions of commonly used interventions, including contraindications for their use.

Traditionally cervical ripening has been performed as an inpatient procedure, and while there is variation, it can require substantial time and resources to accomplish successfully due to multiple factors. While prostaglandins (vaginal or oral) and mechanical methods (e.g., balloon catheters) are the most commonly used methods of cervical ripening in the inpatient setting, there is institution and provider-level variation in the dose, administration route and frequency of administration. Some women’s cervixes will rapidly respond to a cervical ripening intervention, while others require extended time with more than one intervention attempted if the first one fails. While interventions used for cervical ripening are generally not costly, the hospital inpatient resources used, including highly skilled labor and delivery staff, contribute to increased costs when cervical ripening care is provided in the inpatient setting.

For a variety of reasons, some women may prefer to be at home during the cervical ripening process, and because of the resources and variation involved, providers are also interested in exploring methods of cervical ripening in the outpatient setting that have minimal or no increased risk. Informed by these considerations, there is growing interest in, and evidence for, outpatient cervical ripening. It has been proposed that outpatient cervical ripening may facilitate more efficient and more satisfying IOL, also reducing inpatient length of stay compared to inpatient cervical ripening.

There are concerns regarding potential maternal/fetal risks of outpatient cervical ripening in comparison with the inpatient setting. These risks may be compounded by known and theoretical iatrogenic effects of medication and mechanical cervical stimulation. However, the risks of may be mitigated through the choice of cervical ripening method and clinical management. For example, prostaglandin use has been associated with tachysystole and fetal distress. Careful
A review of existing literature is needed to elucidate whether these outcomes occur more frequently when cervical ripening is accomplished in the outpatient versus inpatient setting and whether they increase fetal or maternal morbidity. In addition, maternal or fetal characteristics may differentially affect these risks. Finally, understanding the range of feasible outpatient cervical ripening options, and what form of fetal surveillance should be used (if any), is an important aspect of this review.

A woman’s preferences and satisfaction related to the setting of cervical ripening also need to be considered. Some may actively seek outpatient cervical ripening and others may strongly prefer inpatient cervical ripening. This likely variation in preferences and satisfaction has been identified as an important contextual question of this review.

Despite potential cost saving and sometimes strong personal preferences favoring outpatient cervical ripening, this approach to care is still debated. Controversy is driven by interpretation of risk, clinician’s discipline and experience (e.g., obstetrician vs. midwife),13,14 and geographical practice variation. Clinician and institutional risk-aversion driven by potential legal litigation is also a consideration. The 2009 American College of Obstetricians and Gynecologists (ACOG) Practice Bulletin on induction of labor summarized evidence on cervical ripening in the outpatient setting, based on only two studies available at that time (one on a prostaglandin, one on a single-balloon catheter).15 ultimately not reaching a recommendation. A 2017 Cochrane review found that evidence on outpatient versus inpatient cervical ripening was insufficient to address differences in maternal and fetal/neonatal health outcomes, such as cesarean delivery, between settings.16 This review included only randomized controlled trials (RCTs), and included interventions not available in the United States, or that are used primarily to stimulate or maintain contractions rather than primarily for cervical ripening. Many cervical ripening studies have been conducted in non-U.S. settings, where patient acceptance and understanding of risk may be different, in addition to variation in provider philosophy and health system resources. There is a need to assess the benefits of outpatient versus inpatient cervical ripening, without increasing risk (rise in cesarean delivery rate, adverse neonatal outcomes), framed within considerations of cost, patient autotomy, and satisfaction. This is the crux of the decisional dilemma. When cervical ripening is indicated, what methods can be recommended as effective, but with no increased risks, in the outpatient setting and what surveillance best serves women induced with prostaglandin in the outpatient setting?

**Purpose and Scope of the Systematic Review**

This systematic review assessed the comparative effectiveness and potential harms of cervical ripening in the outpatient versus the inpatient setting, comparison of benefits and harms of different methods of cervical ripening in the outpatient setting, and evidence on benefits and harms of fetal surveillance during labor when a prostaglandin was used for cervical ripening. The intended audience includes the ACOG’s guideline developers, practitioners who deliver infants (e.g., obstetricians, family physicians, nurse midwives), other personnel who administer and monitor cervical ripening, and health system policymakers. In addition to these clinical implications, we hope to inform the future research necessary to provide high-quality, evidence-based care to all pregnant women.
Methods

Review Approach

The methods for this systematic review followed the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (available at https://effectivehealthcare.ahrq.gov/topics/cer-methods-guide/overview). This systematic review reports in accordance with the Preferred Items for Reporting in Systematic Reviews and Meta-Analyses (PRISMA).17

Key Questions

An initial set of Key Questions were posted on the AHRQ Effective Health Care (EHC) Program website for public input from May 10 to May 30, 2019, prior to the initiation of this review, and a public stakeholder webinar was held by the Patient-Centered Outcomes Research Institute (PCORI). Changes to the Key Questions based on public comment include expanding defined subgroups and removing specific brand names for mechanical devices. Subsequently, a group of Key Informants and a separate group of Technical Experts Panel (TEP), including representatives of American College of Obstetrics and Gynecologists’ (ACOG) guideline group, provided comments on the scope of the review. The following Key Questions and inclusion criteria reflect suggestions received, and are in the final protocol. The final protocol was posted on the EHC website on January 16, 2020 (https://effectivehealthcare.ahrq.gov/products/cervical-ripening/protocol) and submitted to PROSPERO for registration (ID CRD42020167406).

KQ1: How do the effectiveness and harms of cervical ripening using prostaglandins compare in the outpatient vs. inpatient setting?

1a: How do effectiveness and harms vary by choice of prostaglandin?

1b: Do effectiveness and harms vary by important patient characteristics (such as gestational age, parity, uncomplicated pregnancy, prior cesarean delivery, etc.)?

KQ2: How do the effectiveness and harms of cervical ripening using mechanical methods (e.g., balloon catheters) compare in the outpatient vs. inpatient setting?

2a: How do effectiveness and harms vary by choice of mechanical method in the inpatient versus the outpatient setting?

2b: Do effectiveness and harms vary by important patient characteristics (such as gestational age, parity, uncomplicated pregnancy, prior cesarean delivery, etc.)?

KQ3: How do the effectiveness and harms of cervical ripening in the outpatient setting vary by method of cervical ripening compared with each other, placebo or expectant management?

3a: Do effectiveness and harms vary by important patient characteristics (such as gestational age, parity, uncomplicated pregnancy, prior cesarean delivery, etc.)?

KQ4: How do the effectiveness and harms of different methods and protocols for fetal surveillance compare with each other or with no monitoring in pregnant women undergoing cervical ripening with prostaglandins?

4a. Do effectiveness and harms vary by important patient characteristics (such as gestational age, parity, uncomplicated pregnancy, prior cesarean delivery, etc.)?
**Contextual Question:** What evidence informs preference for or satisfaction with different methods of cervical ripening in the outpatient setting or outpatient compared to the inpatient setting? (This question is addressed in the Discussion of this report)

**Analytic Framework**

**Figure 1. Analytic framework**

**Population**
Term pregnant women undergoing cervical ripening as the first step in induction of labor, either medically indicated or elective

**Interventions**
- **KQ1.** Pharmacologic agents; outpatient vs. inpatient
- **KQ2.** Mechanical methods; outpatient vs. inpatient
- **KQ3.** Pharmacologic agent or mechanical methods; outpatient
- **KQ4.** Method of fetal surveillance

**Fetal Harms**
- Perinatal mortality, hypoxic-ischemic encephalopathy, seizure, infection, meconium aspiration syndrome, birth trauma, intracranial or subgaleal hemorrhage, need for respiratory support within 72 hours after birth, Apgar score ≤3 at 5 minutes, hypotension requiring vasopressor support, umbilical cord gas < pH 7.0 or 7.10

**Maternal Harms**
- Hemorrhage requiring transfusion, postpartum hemorrhage by mode (i.e., vaginal or cesarean delivery), uterine infection, placental abruption, uterine rupture, umbilical cord prolapse, insufficient duration of time from admission to birth for GBS prophylaxis antibiotics administration

**Effectiveness**
- Total time admission to vaginal birth; total labor and delivery length of stay
- Cesarean delivery rate overall
- Vaginal birth within 24 hours
- Failed induction rate, defined as:
  - Cesarean delivery in patient at <6 cm dilation excluding fetal distress (labour dystocia, failure to progress, etc.)
  - Cesarean delivery in patient at <6 cm dilation for fetal distress
  - Cervical assessment at time of admission (e.g., latent vs. active phase, Bishop score, cervical dilation)
  - Time from ROM to delivery

**KQ3 includes:**
- Breastfeeding
- Maternal mood
- Mother-baby attachment

**Study Selection**

We searched Ovid® MEDLINE®, Embase®, CINAHL®, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews from database inception to October 2019 for randomized controlled trials (RCTs), to January 2020 for initial observational studies, and an expanded additional search in March 2020 for Key Question 4, to identify studies eligible for this review, according to the criteria listed in Table 1 for eligible populations, interventions, comparisons, and outcomes. We included RCTs and comparative cohort studies or case-control studies that attempted to control for potential confounding. In our protocol, we indicated that we would include only observational studies with more than 200 women enrolled, if inadequate evidence was found in RCT evidence for primary outcomes on any Key Question. We ultimately decided to remove the sample size threshold due to limited evidence and all abstracts of previously excluded observational studies were subsequently re-reviewed. We excluded interventions not available in the United States (unless listed in Table 1) and outcomes not listed. While we initially planned to exclude women with contraindications to outpatient cervical ripening, specifically multiple pregnancy, prior uterine rupture, and breech presentation.
of the fetus, we ultimately included a few studies that did enroll women with twin pregnancies and with breech presentation fetuses. A Federal Register notice requesting Supplemental Evidence and Data for Systematic review (SEADS) did not result in any new evidence being identified. We used dual review to select studies. Full details on review methods, including complete search strategies, can be found in the Methods Appendix B.

### Table 1. Criteria for population, intervention, comparison, and outcomes of eligible studies

<table>
<thead>
<tr>
<th>PICOTs</th>
<th>Inclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Question 1: Prostaglandin Outpatient vs. Inpatient</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>Pregnant women ≥37 weeks undergoing cervical ripening</td>
</tr>
</tbody>
</table>

| Interventions and Comparisons | KQ 1 and 3: Pharmacologic agents (prostaglandins)  
|                              | KQ 2 and 3: Mechanical methods (single- and double-balloon catheters, dilators)  
|                              | KQ 1 and 2: The comparison is setting (inpatient versus outpatient)  
|                              | KQ 3: The comparison is different methods in the outpatient setting  
|                              | KQ 4: Any method of fetal surveillance during labor compared with another method/protocol, or no monitoring. |

| Outcomes Birth-related | Primary: Total time admission to vaginal birth; Total L&D length of stay; and Cesarean delivery rate overall  
|                       | Secondary: Vaginal birth within 24 hours; Failed induction, defined as cesarean delivery in patient at <6cm dilation: 1) for fetal distress, and 2) excluding fetal distress (e.g., labor dystocia); Cervical assessment at time of admission; Time from ROM to delivery. In addition, for KQ3, breastfeeding, maternal mood, and mother-baby attachment |

| Outcomes Fetal/Neonate Harms | Primary: Perinatal Mortality; Hypoxic-ischemic encephalopathy; Seizure; Infection (confirmed sepsis or pneumonia); Meconium aspiration syndrome; Birth trauma (e.g., bone fracture); and Intracranial or subgaleal hemorrhage  
|                              | Secondary: Respiratory support within 72 hours of birth; Apgar score ≤3 at 5 minutes; Hypotension requiring vasopressor support; and Umbilical cord gas < pH 7.0 or 7.10 |

| Outcomes Maternal Harms | Primary: Hemorrhage requiring transfusion; Postpartum hemorrhage by mode (vaginal, cesarean delivery); Uterine infection (i.e., chorioamnionitis)  
|                         | Secondary: Placental abruption; Uterine rupture; Umbilical cord prolapse; and Inadequate time for GBS prophylaxis antibiotics |

**a** (Bolded) items indicate Primary Outcomes  
**b** See Appendix A for descriptions of interventions  
**c** Reduced rate is the desired direction of this outcome  
GBS = Group B Streptococcus; KQ = Key Question; L&D = labor and delivery; ROM = rupture of membrane

### Data Extraction and Risk of Bias Assessment

Data were abstracted from included studies into evidence tables, including study and patient characteristics and study results, with data verified for accuracy and completeness by a second team member. The risk of bias of included studies was assessed according to established methods, with RCTs assessed based on criteria established in the *Cochrane Handbook for Systematic Reviews of Interventions*, and cohort studies evaluated using criteria developed by the U.S. Preventive Services Task Force. Based on the risk of bias assessment. Individual included studies were rated as being “good,” “fair,” or “poor” quality.

### Data Synthesis and Analysis

We analyzed the evidence according to Key Question, using both narrative (qualitative) and quantitative (meta-analysis) methods (where possible). In both approaches, we grouped the drugs (Key Questions 1 and 3) by type of prostaglandin (PGE1, misoprostol versus PGE2, dinoprostone). We evaluated any variation in results according to the dose, formulation, or route within those categories using qualitative synthesis, as there were too few studies to conduct subgroup or sensitivity analyses on these factors. For the mechanical methods (Key Questions 2 and 3), we evaluated single- and double-balloon catheters separate from hygroscopic dilators,
and evaluated variation by type of catheter (single- versus double-balloon). We conducted meta-
analysis on the prioritized (primary) outcomes noted in Table 1 above, when there were at least
two studies reporting the same outcome, within the intervention groups described here.
Identification of the primary outcomes was done with input from our Technical Expert Panel
(TEP), partner (ACOG), and sponsor (PCORI) of this report. For the outcome of cesarean
delivery, a positive effect would be to reduce, or at least not increase, the incidence. Secondary
outcomes were only pooled if there were at least three studies available. Profile-likelihood
random effects models were used for meta-analysis, with heterogeneity assessed using both the
$\chi^2$ test and the I-squared ($I^2$) statistic. Small study effects (potential publication bias) was
analyzed using Funnel plots and the Egger test where there were at least 10 studies combined in
meta-analyses. For dichotomous outcomes, we calculated relative risks and 95% confidence
intervals. We calculated relative risks rather than absolute risk differences (and 95% CI) to
account for variation in the underlying risk for the outcome in different study populations. To
give the clearest picture, we present both the absolute incidence and change in each group,
alongside the RR (95% CI). However, to narratively describe the comparisons, we described the
magnitude of absolute differences in the following terms. For cesarean delivery, a difference of
less than 5% was “little or no difference”, differences of 5% to 10% “small difference”, 11% to
20%, “moderate difference”, and greater than 20%, “large difference”. For harm outcomes, a
difference of 1% or less was “little or no difference”, differences of >1% to 3% “small
difference”, >3% to 8%, “moderate difference”, and greater than 8%, “large difference”. For this
assessment, incidences were rounded. We are not certain of that these thresholds translate
directly into clinically meaningful differences. They are an attempt to provide a framework for
interpretation of the findings.

Our a priori plan for subgroup analysis included the population characteristics laid out in the
subquestions of the Key Questions above. Important maternal subgroups: parity, maternal age,
Group B Streptococcus (GBS) status, diabetes (pre-gestational, gestational), hypertension
(chronic, preeclampsia without severe features, gestational). Important fetal subgroups: fetal
growth restriction, gestational age (<39 weeks, 39 to 41 weeks, >41 weeks), though only a few
subgroup analyses were conducted given the limited data. We assessed applicability based on the
source of potential study participants, number of women randomized relative to the number of
women enrolled, and characteristics of the population, intervention (including process details),
and care setting.

**Grading the Strength of the Body of Evidence**

The strength of evidence (SOE) of primary outcome-intervention pairs were evaluated using
the AHRQ methods.\(^2^3\) For example, the SOE on the risk for cesarean delivery with dinoprostone
(PGE2) compared with placebo in the outpatient setting was evaluated separately to the evidence
on this population, intervention and setting with misoprostol (PGE1). Details on the methods
used are presented in the Methods Appendix B and primary outcomes are those bolded in Table
2, above. We note that where there was both RCTs and observational study evidence, we
evaluated these separately and used the observational study evidence to supplement the RCT
evidence in order to come to a final rating. Additionally, for bodies of evidence with only a
single study, we rated consistency as unknown (rather than not applicable). In these cases we did
not automatically downgrade the evidence to “insufficient”, but considered the sample size or
number of events available for analysis.
Results

Description of Included Evidence

Searches identified 10,437 references, from which 639 articles were selected for full-text review after dual review of abstracts. Following dual review of full-text of these articles, 39 unique studies (in 42 publications) were included. Thirty randomized controlled trials (RCTs) and nine cohort studies were included, evaluating 9,465 women. The majority of the evidence pertained to Key Question 3 comparing interventions in the outpatient setting (22 RCTs). The Key Questions comparing cervical ripening in the inpatient and outpatient setting included eight RCTs (2 for Key Question 1 on prostaglandins and 6 for Key Question 2 on mechanical methods) and all nine cohort studies (6 for Key Question 1 and 3 for Key Question 2). We did not identify any studies eligible for Key Question 4, addressing fetal surveillance during labor when a prostaglandin was used for cervical ripening. Four studies were rated good quality, 28 fair quality, and 7 poor quality. In addition, eight studies were identified to help address the Contextual Question (CQ). A flow diagram of the search results and selection of studies, a list of included studies, and a list of excluded studies with reason for exclusion can be found in Appendix C, D and I, respectively.

The characteristics of women enrolled in the included studies are summarized in Table 2, below, and more detailed information by study in Appendix C, Tabled C-2. Participants’ weighted mean age was 28.7 years and weighted body mass index (BMI) was 26.7. BMI was reported in only 20% of trials, and the timing of measurement was not reported. Race was reported in 33% of studies. While more than half included mostly white women (64% to 84%), three included mostly African American women (61% to 88%), and one included mostly Latino women (96%). The majority of participants were nulliparous (65%); only five studies reported mean parity (weighted mean 0.25). Most studies (64%) excluded women with prior cesarean delivery, one RCT limited recruitment to women with prior vaginal birth while another RCT only recruited women with prior cesarean delivery. Relatively few excluded women with preexisting diabetes (13%), gestational diabetes (10%), chronic hypertension (18%), or gestational hypertension (21%), hence a small proportion of women enrolled had gestational diabetes (GDM, 6%), though one RCT reported 69 percent of participants had GDM. Postterm pregnancy was the most frequently reported reason for cervical ripening (61% of all participants). Weighted mean Bishop score at baseline was 3.4 and mean gestational age (GA) was 40.6 weeks. Details of the interventions used in each study can be found in Appendix C, Tables C-3 to C-5. Most studies were conducted in the United States (62%), and less than half (46%) reported funding source; a nonprofit organization was the most prevalent source of funding for those that did report the source (50%). Evidence Tables of study and patient population characteristics and study results can be found in Appendix E, and of risk of bias assessment for individual studies are available in Appendix G.
Table 2. Baseline characteristics of women enrolled in included studies

<table>
<thead>
<tr>
<th>Weighted Means</th>
<th>KQ 1 RCTs</th>
<th>KQ 1 Cohort Studies</th>
<th>KQ 2 RCTs</th>
<th>KQ 2 Cohort Studies</th>
<th>KQ 3 RCTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of studies</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>N population</td>
<td>1127</td>
<td>3963</td>
<td>1214</td>
<td>1142</td>
<td>2741</td>
</tr>
<tr>
<td>N range</td>
<td>300–827</td>
<td>76–1343</td>
<td>48–695</td>
<td>42–615</td>
<td>49–534</td>
</tr>
<tr>
<td>N mean</td>
<td>564</td>
<td>661</td>
<td>202</td>
<td>381</td>
<td>125</td>
</tr>
<tr>
<td>Age (years)</td>
<td>28.2</td>
<td>30.5</td>
<td>29.8</td>
<td>24.2</td>
<td>26.1</td>
</tr>
<tr>
<td>Race, non-white (n studies)</td>
<td>NR</td>
<td>43.1% (2)</td>
<td>41.4% (3)</td>
<td>NR</td>
<td>63.7 (8)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>NR</td>
<td>25.8</td>
<td>27.3</td>
<td>NR</td>
<td>28.5</td>
</tr>
<tr>
<td>Parity</td>
<td>NR</td>
<td>0.23</td>
<td>NR</td>
<td>0.5</td>
<td>0.81c</td>
</tr>
<tr>
<td>Bishop score (0 to 13)</td>
<td>4a</td>
<td>3.3</td>
<td>2.9</td>
<td>NR</td>
<td>3.6</td>
</tr>
<tr>
<td>Gestational age (weeks)</td>
<td>NRb</td>
<td>41.2</td>
<td>40.5</td>
<td>40.3</td>
<td>40.1</td>
</tr>
<tr>
<td>Nulliparous</td>
<td>68.6%</td>
<td>79.1%</td>
<td>62.6%</td>
<td>54.4%</td>
<td>51.8%</td>
</tr>
<tr>
<td>Prior cesarean delivery</td>
<td>0%</td>
<td>0%</td>
<td>6.3%c</td>
<td>15.7%c</td>
<td>35.4%d</td>
</tr>
<tr>
<td>Elective IOL</td>
<td>10.1%</td>
<td>0.6%</td>
<td>24.0%</td>
<td>3.3%</td>
<td>43.6%</td>
</tr>
<tr>
<td>Postterm IOL</td>
<td>83.6%</td>
<td>72.3%</td>
<td>57.5%</td>
<td>51.8%</td>
<td>32.8%</td>
</tr>
<tr>
<td>Medically-indicated IOL</td>
<td>4.6%</td>
<td>26.6%</td>
<td>18.1%</td>
<td>39.5%</td>
<td>21.1%</td>
</tr>
</tbody>
</table>

BMI = body mass index; IOL = induction of labor; KQ = Key Question; RCT = randomized controlled trial; NR = not reported

a Only 1 study reported the median Bishop score at baseline
b One RCT reported mean 40.71 weeks, the other RCT reported median 40.14 weeks

c Based on only one study. All other studies did not report percentage of participants with cesarean delivery or excluded them.
d Based on three trials that included participants with prior cesarean delivery. Twelve other trials excluded participants with prior cesarean delivery.

Results are organized by Key Question and then by comparison. For Key Question 1 and Key Question 3, we refer to the prostaglandins as dinoprostone (PGE2) and misoprostol (PGE1). Outcomes are reported in the following order: birth outcomes, fetal/neonate harms outcomes, and maternal harms outcomes, with primary outcomes of interest within each category listed first followed by any secondary outcomes reported. In many cases the primary (and/or secondary) outcomes of interest were not reported as specified for this report. For example, for the prioritized primary outcome of time from admission to vaginal birth, many studies reported only time from admission to delivery by any mode. We included such related outcomes for completeness when the authors reported them, but they did not contribute to the conclusions (i.e., strength of evidence [SOE]). Our rule for reporting percentages in the text of the report was to report the percent per group as a whole number, except if the incidence was ≤1 percent in either group, when we report two decimal places and the n/N for each group. Information on subgroups, if available, is reported where relevant by outcome. No RCT data were available to evaluate differential effectiveness or harms (i.e., effect modification) for subgroups specified in our protocol for any Key Question. Outcomes that were not reported are not listed or noted below; if a prespecified outcome does not appear within a section that means that no study reported on that outcome.

Key Question 1. Comparative effectiveness and harms of cervical ripening using prostaglandins in the outpatient versus inpatient setting

Key Points

- Based on 2 RCTs and 1 cohort study, there was little to no difference in the frequency of cesarean delivery for outpatient versus inpatient cervical ripening using dinoprostone (SOE: low).
Based on 2 RCTs, the frequency of infection was low overall, with little to no difference between outpatient versus inpatient cervical ripening using dinoprostone (SOE: low).

Evidence on misoprostol was insufficient to draw conclusions, based on 1 small (n=273) cohort study.

Summary of Findings

Two fair-quality RCTs\(^\text{26,65}\) involving 1,122 women, one conducted in Canada and the other in Australia, compared outpatient versus inpatient dinoprostone use for cervical ripening (Appendix E). Funding was government in one\(^\text{65}\) and multiple sources in the other.\(^\text{26}\) Six cohort studies compared prostaglandin use for cervical ripening in the outpatient versus inpatient setting; five\(^\text{24,29,31,54,60}\) assessed dinoprostone (N=3,690, with two studies having an overlap of 793 women in their study populations, out of 1,343 in one and 1,179 in the other\(^\text{29,54}\)), and one (N=273) assessed misoprostol.\(^\text{28}\) For meta-analyses and SOE assessments, only the larger of the two studies with overlapping data were considered.\(^\text{54}\) One cohort study of dinoprostone was rated poor quality due to concerns over patient selection, dissimilarity in baseline patient characteristics without control for confounding, questions on attrition (see Appendix G). The other cohort studies were fair-quality. Three cohort studies did not report their source of funding\(^\text{24,28,60}\), two were industry funded,\(^\text{29,54}\) and one received support from a variety of sources.\(^\text{31}\) Two of the cohort studies were performed in the United States,\(^\text{28,31}\) one in the United Kingdom,\(^\text{60}\) and three in Canada.\(^\text{24,29,54}\) Details of these studies are presented below.

Dinoprostone (PGE2)

Two fair-quality RCTs (N=1,127) compared outpatient versus inpatient dinoprostone for cervical ripening.\(^\text{26,65}\) In one trial,\(^\text{26}\) a single dose of intravaginal controlled release dinoprostone insert 10 mg was administered. In the other trial,\(^\text{65}\) nulliparous women received 2 mg and multiparous women received 1 mg intravaginal dinoprostone gel. Only 51 percent of those randomized received an initial dose of dinoprostone, and of these, 30 percent required a second dose and 6 percent required a third dose. In the outpatient groups, women were monitored for 40 to 60 minutes following dinoprostone insertion, and returned for reassessment after 12 hours (or the next morning). In one trial, women had telephone assessment every four hours, and nonstress testing at the followup visit.\(^\text{26}\) Women assigned to the inpatient group had monitoring that mirrored the outpatient protocols.

Both trials enrolled women with uncomplicated singleton term pregnancies, with cephalic presenting fetuses and excluded women with a uterine scar as well as those with fetal growth restriction, preeclampsia, or rupture of membrane (ROM). Other exclusions varied between trials; exclusion criteria in one trial were modified halfway through to exclude women with a BMI >35 and those with diet-controlled diabetes.\(^\text{65}\) Neither trial reported mean BMI. The weighted mean patient age was 28 years, and the weighted mean gestational age was 40 weeks. Median baseline Bishop score was 4 for both outpatients and inpatients in one trial\(^\text{26}\) (the other trial did not report baseline Bishop score).\(^\text{65}\) Sixty-nine percent of women were nulliparous.

Reasons for induction included postterm pregnancy (not defined) in 84 percent, elective/social in 10 percent (1 trial),\(^\text{65}\) medical necessity in 3.6 percent (1 trial),\(^\text{26}\) and other in 3.3 percent of women (2 trials).\(^\text{26,65}\) Characteristics of enrolled women are in Appendix E-1.

Five retrospective cohort studies compared outpatient and inpatient dinoprostone use for cervical ripening.\(^\text{24,29,31,54,60}\) The two overlapping fair-quality studies investigated intravaginal, controlled release 10 mg dinoprostone inserts,\(^\text{29,54}\) and the other studies evaluated dinoprostone
gel: one fair-quality study used 1 mg intravaginal and the other two used the gel intracervically (a fair-quality study used 0.5 mg and a poor-quality study used 2 mg). In one large cohort study, only 59 percent (597/907) completed in-home cervical ripening. Except for one cohort study where a single dose was used, additional doses were given as determined at the followup visits (criteria for administering addition doses not provided). All included women carrying singleton, vertex presenting fetuses and while all excluded women with prior uterine surgery, exclusions across studies varied. Three excluded women with ROM, and two excluded women with medical or fetal risk factors (e.g., hypertension). The two cohort studies with overlapping populations differed in their inclusion criteria, with the more recent study including only women who presented for induction for postterm pregnancy or premature rupture of membrane (PROM). Continuous fetal heart rate (FHR) monitoring was done for 30 minutes to 1 hour post dinoprostone insertion; one study reported a maximum of 2 hours of monitoring. Across four studies, the weighted mean age was 31 years and mean gestational age was 41 weeks, while a fifth study reported that 54 percent of women had a gestational age between 37 and 41 weeks and 46 percent had a gestational age >41 weeks. The weighted mean Bishop score at baseline was 3.9 (3 studies) and 55 percent of women were nulliparous (3 studies). Reason for induction was postterm pregnancy in 72 percent, elective in 0.6 percent (2 studies), and medical necessity in 27 percent of women (4 studies). Postterm pregnancy was defined as >41 weeks (2 studies), between 41 weeks and 3 days and 42 weeks (1 study), and not defined in two studies. Only one study reported BMI (mean 25.8 kg/m²).

Characteristics of enrolled women are in Appendix E-1.

The only primary birth outcome reported was cesarean delivery, which was found to be not significantly different between outpatient and inpatient use (SOE: low). Fetal, neonatal, and maternal harms appeared to be rare across included studies but were poorly reported; studies generally did not report on most of our prespecified outcomes. The frequencies of infection (not confirmed), hypoxic-ischemic encephalopathy, meconium aspiration (not specified as the syndrome) and postpartum hemorrhage requiring transfusion were not significantly different for outpatients and inpatients (SOE: low). However, individual study sample sizes were likely too small detect differences in rare events.

Detailed study results can be found in Appendix E-2.

Birth Outcomes

Total Time from Admission to Vaginal Birth

None of the included studies reported the prespecified primary birth outcomes related to duration of labor according to mode (i.e., vaginal birth). One cohort study (N=992) reported the time from admission to delivery (not specified by mode), finding 26.25 hours in the outpatient group and 24.28 hours in the inpatient group, with the difference not being statistically significant (mean difference [MD] 1.97 hours, 95% confidence interval [CI] −1.18 to 5.13). Other studies reported only the total length of hospital stay, again not stratified by mode of delivery. Two RCTs reported similar durations of stay for outpatients and inpatients (MD −0.58 hours, 95% CI −6.40 to 4.73), I²=0%). Across two small cohort studies, the duration of hospital stay was significantly shorter among outpatients compared with inpatients (pooled MD −17.34 hours, 95% CI −32.90 to −6.08, I²=23.3%; Appendix F).
Cesarean Delivery

Based on pooled analysis, the frequency of cesarean delivery in outpatients and inpatients was the same (2 RCTs, 23% vs. 23%; relative risk [RR] 0.97, 95% CI 0.75 to 1.25, I²=0%).²⁶,⁶⁵ Although overall cesarean delivery frequency was higher in cohort studies (33% vs. 23% for RCTs), pooled incidence was again the same for outpatients and inpatients (4 cohort studies, 33% vs. 33%; RR 0.79, 95% CI 0.67 to 0.98, I²=0%; Figure 2). Exclusion of the poor-quality cohort study did not change the estimate (3 cohort studies, 34% vs. 34%; pooled RR 0.80, 95% CI 0.68 to 1.01, I²=0%). In a subgroup analysis in one cohort study²⁹ the frequency of cesarean delivery in women with postterm pregnancies (adjusted odds ratio [OR] 0.74, 95% CI 0.54 to 1.01) was not significantly different to that of the full population (postterm and PROM, adjusted OR 0.71, 95% CI 0.54 to 0.95). This study was not included in the meta-analysis due to substantial overlap in patient populations with another study from the same institution.⁵⁴

Figure 2. Meta-analysis of cesarean delivery with prostaglandins for cervical ripening: outpatient versus inpatient

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Quality Rating</th>
<th>Treatment, Control,</th>
<th>Risk Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author, Year</td>
<td>Intervention</td>
<td>n/N</td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biem 2003</td>
<td>Dinoprostone</td>
<td>35/149</td>
<td>0.95 (0.64, 1.42)</td>
</tr>
<tr>
<td>Wilkinson 2015a</td>
<td>Dinoprostone</td>
<td>91/407</td>
<td>0.97 (0.76, 1.25)</td>
</tr>
<tr>
<td>Subgroup (I-squared = 0.0%, p = 0.925)</td>
<td></td>
<td></td>
<td>0.97 (0.75, 1.25)</td>
</tr>
<tr>
<td>Cohort - Dinoprostone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awartani 1999</td>
<td>Dinoprostone</td>
<td>2/50</td>
<td>0.40 (0.08, 1.97)</td>
</tr>
<tr>
<td>Farmer 1996</td>
<td>Dinoprostone</td>
<td>14/40</td>
<td>1.15 (0.60, 2.19)</td>
</tr>
<tr>
<td>Salvador 2009*</td>
<td>Dinoprostone</td>
<td>183/567</td>
<td>0.75 (0.61, 0.92)</td>
</tr>
<tr>
<td>Stock 2014</td>
<td>Dinoprostone</td>
<td>319/907</td>
<td>0.83 (0.64, 1.08)</td>
</tr>
<tr>
<td>Subgroup (I-squared = 0.0%, p = 0.504)</td>
<td></td>
<td></td>
<td>0.79 (0.67, 0.98)</td>
</tr>
<tr>
<td>Cohort - Misoprostol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang 2005</td>
<td>Misoprostol</td>
<td>25/177</td>
<td>0.75 (0.43, 1.31)</td>
</tr>
<tr>
<td>Subgroup (I-squared = NA, p = NA)</td>
<td></td>
<td></td>
<td>0.75 (0.43, 1.31)</td>
</tr>
<tr>
<td>Cohort subgroup (I-squared = 0.0%, p = 0.668)</td>
<td></td>
<td></td>
<td>0.79 (0.68, 0.94)</td>
</tr>
</tbody>
</table>

*RR estimate calculated from author’s adjusted odds ratio comparing inpatient with outpatient.
CI = confidence interval; IP = inpatient; NA = not applicable; OP = outpatient
Other Birth Outcomes

Few studies reported secondary outcomes prioritized for this review. Three studies reported vaginal birth within 24 hours of cervical ripening. In a fair-quality RCT (N=425), although the incidence was numerically lower in the outpatient group (41% vs. 50%), the difference was not statistically significant (RR 0.84, 95% CI 0.68 to 1.03). The two cohort studies with overlapping populations reported a significantly lower proportion of women delivered vaginally within 24 hours of a single dose of dinoprostone in the outpatient (12%) versus the inpatient (41%) group (reported in the study as inpatient versus outpatient), the adjusted OR is 2.16 (95% CI 1.57 to 2.97). The other cohort study (N=992) reported on the frequency of delivery via any mode by 24 hours, with similar rates across the groups (23% vs. 27%; RR 0.84, 95% CI 0.58 to 1.22).

Failed induction was variably defined and no study’s definition met our protocol specification that the cervical dilation must be <6 cm. Three studies reported on the frequency of cesarean delivery for fetal distress, with none finding a difference between outpatient and inpatient groups that reached statistical significance. A RCT (N=821) found somewhat fewer cesarean deliveries due to fetal distress in outpatients compared with inpatients (9% vs. 12%; RR 0.78, 95% CI 0.52 to 1.18). Two cohort studies also reported on cesarean delivery for fetal distress, where the incidence was much lower in both groups (0.5%) and not significantly different between outpatient and inpatient groups (RR 0.88, 95% CI 0.11 to 7.35, I²=0%).

Fetal/Neonate Harm Outcomes

Harms were poorly reported across studies. Studies were likely underpowered to detect most fetal and neonatal outcomes or differences between outpatients and inpatients, given their rare frequency.

Perinatal Mortality

Death at three months occurred in one infant in the outpatient group due to meconium aspiration following a long labor involving additional dinoprostone insertion, use of oxytocin, and cesarean delivery in one cohort study.

Hypoxic-Ischemic Encephalopathy

The risk of encephalopathy was very low in two studies reporting this outcome, with too few patients to determine differences between groups. A fair-quality RCT reported low risk of hypoxic-ischemic encephalopathy (0.74% [3/407] vs. 0.72% [3/414]) and one cohort study reported “neonatal encephalopathy” (cause not specified, 0.11% [1/907] vs. 0% [0/85]).

Infection

None of the included studies reported confirmed neonatal infection, but neonatal infection as a cause for admission to an intensive care unit was uncommon and similar between outpatients (4%) and inpatients (3%) across two RCTs (N=1,120; RR 1.39, 95% CI 0.67 to 3.03, I²=0%; Appendix F).

Meconium Aspiration Syndrome

None of the studies reported explicitly on meconium aspiration syndrome. In a single RCT, one neonate from the outpatient group required neonatal intensive care unit (NICU) admission for meconium aspiration: 0.67% (1/149) vs. 0% (0/150). One cohort study reported a neonatal
death in the outpatient group that was attributed to meconium aspiration (0.11% [1/907] vs. 0% [0/85]).

**Other Fetal/Neonatal Outcomes**

No study reported the need for respiratory support, but one RCT reported respiratory problems (not specified) requiring admission to a “special care nursery” (presumed similar to NICU) in the same proportions between women having outpatient versus inpatient cervical ripening (4% vs. 4%; RR 0.90, 95% CI 0.47 to 1.75). The frequency of Apgar scores ≤3 at 5 minutes was also similar between outpatients and inpatients (0% [0/149] vs. 0.67% [1/150]) in one RCT, as were Apgar scores <7 in the other RCT and across four cohort studies (RR 1.00, 95% CI 0.29 to 4.61). One cohort study reported the mean Apgar score at 5 minutes was 8.9 in both groups. One RCT (N=299) reported the mean umbilical cord pH to be very comparable between groups (7.25 versus 7.24). A retrospective cohort study (N=992) reported that 3 percent of the outpatient group had cord pH <7.0 compared with none in the inpatient group. However, 70 percent of neonate records were missing this outcome. As a proxy for other important neonatal outcomes, two studies reported admission to NICU (at varying timepoint after birth), with little difference in rates between groups.

**Maternal Harm Outcomes**

Harms were poorly reported across the studies. Maternal hemorrhage requiring transfusion was the only outcome reported as defined by our protocol. A single RCT reported that maternal hemorrhage requiring transfusion was rare, with little difference between outpatients and inpatients (0.67% [1/149] vs. 0% [0/150]; RR 3.02, 95% CI 0.12 to 73.55). The findings were similar when limited to hemorrhage requiring hysterectomy following a cesarean delivery (0% vs. 0.76%; RR 0.34, 95% CI 0.01 to 8.17). One small cohort study (N=76) reported the incidence of infection, not confirmed or specifically uterine, with no cases in either group. Secondary maternal harm outcomes were also rare, with little differences between groups (Table 3).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design Sample Size</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postpartum hemorrhage &gt;1000 mL</td>
<td>Cohort study N=992</td>
<td>13% vs. 9%</td>
<td>RR 1.34 (.68 to 2.64)</td>
</tr>
<tr>
<td>Placental abruption</td>
<td>RCT N=299</td>
<td>0.7% vs. 1.3%</td>
<td>RR 0.50 (0.05 to 5.49)</td>
</tr>
<tr>
<td></td>
<td>Cohort study N=100</td>
<td>0% vs. 0%</td>
<td>Not calculated</td>
</tr>
<tr>
<td>Uterine rupture</td>
<td>RCT N=299</td>
<td>0% vs. 0.7%</td>
<td>RR 0.34 (0.01 to 8.17)</td>
</tr>
<tr>
<td></td>
<td>Cohort study N=992</td>
<td>0% vs. 0%</td>
<td>Not calculated</td>
</tr>
</tbody>
</table>

CI = confidence interval; mL = milliliter; RCT = randomized controlled trial; RR = relative risk

In addition to the study that conducted subgroup analysis for time to delivery based on parity, another cohort study compared outpatient versus inpatient dinoprostone use in women who either had postterm pregnancies or PROM. Results from analyses confined to postterm pregnancies were similar to those of full population (postterm and PROM) for cesarean delivery, vaginal birth within 24 hours of first dinoprostone insertion, 5-minute Apgar score ≤7, and for NICU admission >12 hours.
**Misoprostol (PGE1)**

Evidence on misoprostol in the outpatient versus inpatient setting was very limited, with one small, fair-quality prospective cohort study (N=273)\(^28\) of a single dose of intravaginal misoprostol 50 µg for cervical ripening given the evening prior to scheduled induction of labor (IOL). All women were monitored for one hour post-dose, and women in the outpatient group were discharged if the FHR was reactive and no regular contractions were noted, with explicit instructions to return if contractions started (every 5 minutes), suspected ROM, lack of perceived fetal movement, or vaginal bleeding. All women were reassessed the following morning to determine the need for a repeat dose of misoprostol. Those with history of rapid delivery, grand multiparity, active medical problems, FHR decelerations, prior uterine surgery, ROM, placenta previa, or vaginal bleeding were excluded from outpatient cervical ripening. Inpatients had coexisting complications (e.g., diabetes, gestational hypertension, precipitous delivery, poor obstetric history) but frequencies were not reported. More women in the outpatient group had baseline cervical dilation of ≥2 cm (40% versus 29%). The mean patient age was 26 years, 46 percent werenulliparous, 64 percent were classified as non-Hispanic white, 22 percent as Hispanic, and 14 percent American Indian. The mean gestational age was between 40 and 41 weeks and 6 days in 59.6 percent of women, and baseline cervical dilation was ≥2 cm in 36 percent of women. This study did not report reasons for induction, BMI, and Group B Streptococcus (GBS) colonization. No women had a history of cesarean delivery. (See Appendix E-1 for patient and study characteristics.)

The primary birth outcomes reported were time from hospital admission to vaginal birth and cesarean delivery. Outpatient cervical ripening was associated with a shorter time from admission to vaginal birth compared with inpatient cervical ripening in both nulliparous and multiparous women. Slightly lower cesarean delivery rates (not statistically significant) were also observed in outpatients. The evidence was insufficient to draw conclusions.

Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

**Time from Admission to Vaginal Birth and Cesarean Delivery Frequency**

Mean time from hospital admission to vaginal birth was somewhat shorter for both nulliparous (MD –3.1 hours, 95% CI –4.74 to –1.46) and multiparous outpatients (MD –5.30 hours, 95% CI –6.84 to –3.76) versus inpatients. Spontaneous vaginal birth was more common in outpatients (80% vs. 64%; RR 1.26, 95% CI 1.07 to 1.50). The study did not adjust for the differences in baseline cervical dilation between groups. The frequency of cesarean delivery was slightly lower for outpatients (14%) versus inpatients (19%), but differences were not statistically significant (RR 0.75, 95% CI 0.43 to 1.31; Appendix E-2).

**Other Birth Outcomes**

There were no cesarean deliveries for failed induction (defined as cervical dilation of ≤3 cm after intervention) in either group. Risk of cesarean delivery for failure to progress was similar between groups (5% vs. 6%). On admission, more outpatients had advanced cervical dilation (4 to 9 cm) compared with inpatients (10% vs. 2%; RR 5.0, 95% CI 1.2 to 21.5). (See Appendix E-2 and Appendix F)
Fetal/Neonate Harm Outcomes

Fetal and neonatal harms were poorly reported. Suspected neonatal sepsis (confirmation not reported) was rare (0.40% [1/177] versus 0% [0/96]). Neonatal breathing difficulties (not defined) were less common in outpatients (3% vs. 7%; RR 0.39, 95% CI 0.13 to 1.19) as was nonreassuring FHR on admission (9% vs. 14%; RR 0.65, 95% CI 0.34 to 1.28) but neither were statistically significant. No neonates in either group had 5-minute Apgar scores ≤3; scores between 4 and 6 occurred in 0.8 percent versus 1.8 percent of neonates. There was little difference in the frequency of neonates with Apgar scores <7 at 5 minutes between outpatients and inpatients. Authors reported that meconium was “uncommon and not more frequent in outpatients” and that the frequency of newborn complications requiring >3-day hospitalization was “unaffected by” outpatient status. (Appendix E-2)

Maternal Harm Outcomes

Authors only reported that no cases of placental abruption or rupture were recorded.

Key Question 2. Comparative effectiveness and harms of cervical ripening using mechanical methods in the outpatient versus inpatient setting

Key Points

- Based on three RCTs and two cohort studies comparing outpatient versus inpatient single-balloon catheter for cervical ripening, differences on all outcomes between groups were small and did not reach statistical significance (SOE: low).
- The evidence comparing outpatient versus inpatient double-balloon catheter, based on a single small (n=48) trial, was insufficient to draw conclusions.
- Evidence comparing any outpatient catheter (1 double-balloon RCT, 1 single-balloon RCT) versus inpatient dinoprostone found small differences that did not reach statistical significance for birth outcome, uterine infection and postpartum hemorrhage (SOE: low). Evidence on fetal/neonatal harms was insufficient to draw conclusions.
- Evidence on hygroscopic dilators was based on 1 small cohort study (n=42) and was insufficient to draw conclusions.

Summary of Findings

Single-balloon Catheter

Three fair-quality RCTs (N=370) compared outpatient with inpatient single-balloon (Foley) catheters (Appendix E).39,51,57 Two studies specified that catheters were placed by physicians or residents, one study did not report the type of provider inserting the catheter; balloons were filled to 30 ml in two studies and 40 ml in one study. One trial included concomitant oxytocin in the inpatient group.39 Following catheter placement, all three trials conducted FHR monitoring prior to sending the outpatient group home (for 20 to 30 minutes in 2 trials39,57 and duration not reported in the third51). In the absence of labor or clinical events (e.g., vaginal bleeding, decreased fetal movements, ROM) women in the outpatient arms were instructed to return to the hospital the next day39,57 or in 24 hours.51 Women in the inpatient arms in all three trials were admitted to the hospital following catheter placement. Study inclusion criteria required gestational age ≥37 weeks in one study,57 39 to 42 weeks in one study,39 and ≥41 weeks in the
third. Two studies\textsuperscript{51,57} required Bishop score \( \leq 5 \) for enrollment, and the third required cervical dilation \(<3\) cm, or if dilation was at least \(2\) cm, cervical effacement had to be \(<80\) percent.\textsuperscript{39} One study required enrollment of parous women.\textsuperscript{39} Two RCTs were conducted in the United States and one in Portugal; none of the trials reported funding source.

We also included a fair-quality retrospective cohort study\textsuperscript{46} (\(N=615\)) that compared outpatient versus inpatient single-balloon catheter at a single center in Australia (Appendix E). The study did not report the type of provider placing the catheter, and balloons were reported to be filled to 30 ml. In both groups, fetal monitoring was conducted before and after catheter placement (duration of monitoring was not reported). Following monitoring, women in the outpatient group were sent home with instructions to return the next morning, and women in the inpatient group were admitted to the hospital with continuous fetal monitoring. One other poor-quality cohort study was identified (Appendix E).\textsuperscript{38}

The characteristics of women enrolled in these studies are detailed in Appendix E-1. Across the three trials, weighted mean maternal age was 28.8 years and mean gestational age was 39.7 weeks. One trial only enrolled parous women,\textsuperscript{39} one trial included 74.6 percent nulliparous women,\textsuperscript{51} and the last trial did not report parity but reported a mean gravidity of 1.8.\textsuperscript{37} Mean Bishop score was 2.2 in one trial,\textsuperscript{51} median of 2 in another,\textsuperscript{57} and median of 2.0 using modified Bishop score in the third study.\textsuperscript{39} The proportion of women with comorbidities such as diabetes and hypertension were mostly not reported except for gestational diabetes, which affected 3 percent of women across two trials.\textsuperscript{39,57} The reason for cervical ripening varied by study; with all undergoing elective induction in one,\textsuperscript{39} reported as elective in 43 percent in a second, with other reasons for induction being post due date (33\%) or medically-indicated (19\%).

In the cohort study, mean maternal age was 24 years and mean gestational age was 40.4 weeks. The mean modified Bishop score at time of catheter placement was 1.85. The authors did not report BMI. Forty-seven percent of women were nulliparous. A slightly smaller proportion of women in the outpatient group were parous (48.5\%) compared with the inpatient group (58.4\%; \(p=0.02\)). There were also numerous between-group differences in indication for induction between the outpatient and inpatient groups, including the proportion of postterm pregnancies (41\% vs. 4\%, \(p<0.001\)), oligohydramnios (11\% vs. 28\%, \(p<0.001\)), preeclampsia (0.3\% vs. 32\%, \(p<0.001\)), abnormal FHR (2\% vs. 10\%, \(p=0.001\)), and maternal diabetes (17\% vs. 5\%, \(p<0.001\)). The study was conducted in the United States and funding was not reported.

All studies reported cesarean delivery, and while there was a trend towards lower risk in the outpatient catheter groups, this difference was not statistically significant. These studies sporadically reported on maternal and fetal/neonatal harm outcomes, and were likely too small to identify real differences in risk. For all outcomes, the evidence did not identify clear differences between outpatient and inpatient single-balloon catheters, and is low strength or insufficient.

Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

**Total Time from Admission to Vaginal Birth and Total L&D Length of Stay**

One RCT reported time from admission to the labor and delivery unit to delivery (vaginal or cesarean delivery).\textsuperscript{39} The study found that women who had undergone outpatient and inpatient cervical ripening had similar time from admission to delivery (12.4 vs. 13.5 hours; MD –1.10, 95\% CI –3.59 to 1.39).

Although not a prespecified primary outcome for this review, total time of hospital stay was shorter in the outpatient groups in two RCTs\textsuperscript{39,51} and one cohort study.\textsuperscript{46} In all three studies, the
inpatient group was admitted immediately following catheter placement, while the outpatient group was sent home with instructions to return the next day. When the RCTs were pooled, the mean difference between outpatient and inpatient groups was −7.15 hours (95% CI −18.94 to 4.47, I²=87%). There is a high degree of heterogeneity in this analysis, as the two studies reported magnitudes of difference (favoring outpatient setting) that were quite different (−12 hours in one and −2.4 hours in the other). The difference between groups in the cohort study was longer (−24 hours) and statistically significant (95% CI −32.81 to −15.19). One study conducted subgroup analysis, comparing the total time from labor unit admission to delivery based on cervical dilation <2 cm (n=28 outpatient and 36 inpatient), finding a non-statistically significant mean difference between groups of −1.2 hours (95% CI −4.85 to 2.45).39

**Cesarean Delivery**

Evidence from three RCTs consistently showed that outpatient single-balloon catheter use was associated with a lower incidence of cesarean delivery compared with inpatient use, although the difference did not reach statistical significance (Appendix E-2).39,51,57 When pooled, the risk of cesarean delivery remained lower in the outpatient group, but the estimate was also not statistically significant (3 RCTs, pooled RR 0.59, 95% CI 0.21 to 1.03, I²=0%; Figure 3). This meta-analysis includes one study that did not report the overall incidence of cesarean delivery, but the incidence of cesarean delivery due to failed induction, which was significantly lower in the outpatient group (3% vs. 17%; RR 0.18, 95% CI 0.41 to 1.10).51 Pooled results from two cohort studies were similar to the RCT finding, with a nonsignificant lower risk of cesarean delivery in the outpatient catheter group (33% vs. 30%; RR 0.95, 95% CI 0.72 to 1.22, I²=0%).38,46

**Other Birth Outcomes**

One RCT reported a lower risk of cesarean delivery due to failed induction in the outpatient single-balloon catheter group versus inpatient (3% vs. 17%; RR 0.18, 95% CI 0.41 to 1.10).51 A fair-quality cohort study (N=615) found a significantly more women in the outpatient group requiring cesarean delivery due to arrest of labor, when limiting the analysis to women with a cesarean delivery.46 When analyzing across all women in each group, the difference was not significant, likely due to the imbalance in the numbers of women in each group (25% vs. 20%; RR 1.21, 95% CI 0.90 to 1.64). Cesarean delivery due to fetal distress was also similar when considering all patients in each group (15% vs. 11%; RR 1.27, 95% CI 0.74 to 2.19).

One RCT reported higher Bishop scores at the time of admission in the outpatient group (5.70) than the inpatient group (2.10) (MD 3.60, 95% CI 2.95 to 4.25).51 Two RCTs and the cohort study reported similar or identical Bishop scores on admission in both outpatient and inpatient groups.39,46,57

**Fetal/Neonatal Harm Outcomes**

None of the RCTs reported the incidence of neonatal infection, but the cohort study found no difference between outpatient and inpatient groups in risk of culture-confirmed neonatal sepsis (2% vs. 2%; RR 0.75, 95% CI 0.24 to 2.34).46 In one trial, there was little difference in the incidence of birth trauma (cephalohematoma, subgaleal hematoma, brachial plexus injury, clavicle or humerus fracture, and scalp laceration) between outpatient and inpatient groups (2% [1/65] vs. 3% [2/64]; RR 0.49, 95% CI 0.05 to 5.30). In the same study, shoulder dystocia was less likely to occur in the outpatient group but the difference was not statistically significant (3%
vs. 11% [7/64]; RR 0.28, 95% CI 0.06 to 1.30). These findings are insufficient to draw conclusions.

Regarding other fetal/neonatal harm outcomes, one trial reported that one infant in each group required respiratory support within 72 hours after birth (2% vs. 2%; RR 1.00, 95% CI 0.06 to 15.65). The numbers of neonates with Apgar score of <7 at 5 minutes were very small (0 to 1 per intervention group) in two trials, resulting in imprecise risk estimates that indicated no difference between groups (2 RCTs, 0.77% [1/130] vs. 1.6% [2/129]; pooled RR 0.62, 95% CI 0.08 to 4.98, I²=0%). In the cohort study, results for Apgar score <7 at 5 minutes were consistent with the trials (1.3% [4/300] vs. 1.3% [4/315], RR 1.05, 95% CI 0.26 to 4.16).

**Maternal Harm Outcomes**

One trial found that postpartum hemorrhage, defined as blood loss >1,000 ml, rarely occurred in either the outpatient or inpatient groups (0% [0/65] vs. 2% [1/64]; RR 0.33, 95% CI 0.01 to 7.91). This evidence is insufficient.

Two trials reported the same incidence of chorioamnionitis in the outpatient and inpatient groups (5% for both, pooled RR 0.99, 95% CI 0.36 to 2.74, I²=0%; Appendix F). Two cohort studies reported any uterine infection. While the incidence was slightly greater in the outpatient group (7% vs 4%), pooled analysis indicates that the difference was not statistically significant (RR 1.70, 95% CI 0.90 to 3.69, I²=0%; Appendix F). This evidence is low strength.

**Double-balloon Catheter**

One small (N=48) fair-quality trial compared cervical ripening with a double-balloon catheter (Cook®) in the outpatient versus inpatient setting (Appendix E). In both groups, women underwent 20 minutes of cardiotocograph monitoring, followed by catheter insertion by a doctor or midwife trained in catheter assertion if cardiotocographic monitoring was reassuring. Monitoring continued for 20 minutes after catheter insertion in both groups. Women in the outpatient group were then discharged home with instructions to return to the hospital the following morning or at the onset of labor or clinical complications (e.g., vaginal bleeding). Outpatient group participants did not need to return to the hospital early in the event of catheter expulsion. Both groups received a double-balloon catheter; each balloon was inflated with 70 to 80 milliliters of water. Study eligibility criteria included ≥37 to ≤42 weeks gestation, Bishop score <7, intact membranes, and singleton pregnancy. Women with a history of cesarean delivery were excluded. The study was conducted in Australia and funded by a non-profit organization.

The details on characteristics of women enrolled in this study can be found in Appendix E-1. The mean maternal age was 29 years, and mean gestational age was 40.75 weeks. Authors did not report BMI or other anthropomorphic measures. Seventy-five percent of women were nulliparous. Cervical dilation at the time of catheter insertion was not reported, but mean Bishop score was 0 to 2 in 27 percent of the population, 3 to 4 in 52 percent, ≥5 in 17 percent, and not reported for 4 percent. Nearly one-third of catheters (29%) were spontaneously expelled.

Outpatient double-balloon catheter use was associated with shorter times from hospital admission to vaginal birth compared with inpatient catheterization. There were no statistically significant differences between outpatient and inpatient groups for other outcomes. Low event rates resulted in risk estimates that were generally imprecise, and study authors noted that the trial was not designed nor adequately powered to detect differences between groups for clinical outcomes. Due to these limitations and combined with the small sample size, the strength of
evidence is insufficient to recommend outpatient versus inpatient double-balloon catheter for all outcomes. Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

**Total Time from Admission to Vaginal Birth**

Time from admission to vaginal birth was significantly shorter in the outpatient group (mean 14.25 hours) than the inpatient group (mean 21.45 hours; MD −7.2 hours, 95% CI −11.45 to −2.95). Related outcomes such as time from catheter insertion to both active labor (17.5 vs. 19.5 hours; MD −2.0 hours, 95% CI −4.02 to 0.02) and vaginal birth (24.5 vs. 29.0 hours; MD −4.5 hours, 95% CI −8.96 to −0.04) were also shorter in the outpatient group, though only time to vaginal birth reached statistical significance. Due to study limitations, lack of confirmatory evidence, and imprecise estimates, this evidence is insufficient to draw conclusions.

**Cesarean Delivery**

Rate of cesarean delivery was lower in the outpatient catheter group versus the inpatient group (18% vs. 33%) but the difference did not reach statistical significance (RR 0.55, 95% CI 0.20 to 1.51; Figure 3). Due to study limitations, lack of confirmatory evidence, and imprecise estimates, this evidence is insufficient to draw conclusions.

**Other Birth Outcomes**

The proportion of women who delivered within 24 hours of catheter insertion was 33 percent in the outpatient group and 27 percent in the inpatient group (RR 1.25, 95% CI 0.47 to 3.29). The trial narratively reported no cases of failed induction in either group, though there was one cesarean delivery in the inpatient group due to fetal distress (0% [0/33] vs. 7% [1/15]; RR 0.16, 95% CI 0.01 to 3.64).

**Fetal/Neonatal Harm Outcomes**

No perinatal deaths were reported in either the outpatient or inpatient catheter groups, nor were there any cases of neonatal infection reported in either group. There was one case of meconium aspiration in the outpatient catheter group (3%; 1/33) compared with no cases in the inpatient group (0%; 0/15), resulting in an imprecise risk estimate (RR 1.41, 95% CI 0.06 to 32.78). This case was not specified as meeting criteria for meconium aspiration syndrome, however. Due to study limitations, lack of confirmatory evidence, and imprecise estimates, this evidence is insufficient to draw conclusions. Two neonates in the outpatient catheter group (6%) had Apgar score <7 at 5 minutes versus none in the inpatient group; the relative risk was 2.35 (95% CI 0.12 to 46.22).

**Maternal Harm Outcomes**

Postpartum hemorrhage, defined in this study as blood loss of >500 milliliters for vaginal birth and >1 liter for cesarean delivery, was slightly higher, but not significantly so, in the outpatient group (18% vs. 13%; RR 1.36, 95% CI 0.31 to 5.99). Due to study limitations, lack of confirmatory evidence, and imprecise estimates, this evidence is insufficient to draw conclusions.
Birth Outcomes

Total Time of Admission to Vaginal Birth and Frequency of Cesarean Delivery

The findings for time to delivery appeared similar between single-balloon catheters and double-balloon catheters, with a pooled analysis of two RCTs (N=259) of single-balloon catheters finding a difference of −7.15 hours (95% CI −18.94 to 4.47), and a single small RCT (N=49) of the double-balloon catheter finding a mean difference of −7.2 hours (95% CI −11.45 to −2.95). However, the outcome reported in the studies of the single-balloon catheter were for any mode of delivery, while the study of the double-balloon catheter reported the prioritized outcome of vaginal birth only.

Meta-analysis comparing any outpatient versus inpatient catheter use (N=418) found that risk of cesarean delivery was lower in the outpatient group (4 RCTs, 13% vs. 21%; pooled RR 0.58, 95% CI 0.29 to 0.89, I²=0%; Figure 3). When analyzed according to catheter type, use of both outpatient single-balloon (3 RCTs, 12% vs. 20%; RR 0.59, 95% CI 0.21 to 1.03, I²=0%) and double-balloon (1 RCT, 18% vs. 33%; RR 0.55, 95% CI 0.20 to 1.51) catheters resulted in a lower risk of cesarean delivery compared to inpatient use.

Maternal Harm Outcomes

Two trials found mixed, but nonsignificant, results for risk of postpartum hemorrhage (Appendix F). Absolute event rates were small in both studies and risk estimates were imprecise for both types of catheters. In one study, outpatient double-balloon catheter use was associated with increased risk of hemorrhage with vaginal births compared with inpatient use (18% [6/33])
vs. 13% [2/15]; RR 1.36, 95% CI 0.31 to 5.99). Outpatient single-balloon catheter use was associated with a reduced risk of postpartum hemorrhage regardless of mode of delivery relative to inpatient use in the other study (0% [0/65] vs. 1.6% [1/64]; RR 0.33, 95% CI 0.01 to 7.91).

Hygroscopic Dilators

A small, poor-quality cohort study (N=42) compared the use of a hygroscopic dilator (Dilapan®) in outpatient versus inpatient settings. The study reported that the number of sticks administered was “as many as the cervix could accommodate,” and the mean was 6 in the outpatient group and 5 in the inpatient group. The type of provider placing the sticks was not reported. The results are reported here as this was the only study identified comparing a hygroscopic dilator in the outpatient versus inpatient setting. Study inclusion criteria required ≥37 weeks gestation with no active labor or contraindication to labor, Bishop score ≤4 and documented fetal well-being. Details of patient characteristics can be found in Appendix E-1. The mean age of enrolled women was 23 years and 15 percent had BMI ≥30; mean gestational age was 40 weeks. Mean gravidity was 1.65 and mean parity was 0.5. Authors did not report the reason for induction. The study was conducted in the United States and the funding source was not reported.

The only birth outcome of interest reported was total length of hospital stay, which was shorter in the outpatient group. Evidence on other outcomes, reported narratively, was very limited but described as similar in both groups. The strength of evidence is insufficient for outpatient versus inpatient hygroscopic dilator for all outcomes. Detailed study results can be found in Appendix E-2.

Birth Outcomes

The total hospital stay was 51 hours in the outpatient group versus 70 hours in the inpatient group (p=0.01). The study did not report cesarean delivery rate or other birth outcomes.

Fetal/Neonatal Harm Outcomes

Fetal and neonatal harms were not reported, although the study narratively reported as similar rates of nonreassuring FHR in both groups.

Maternal Harm outcomes

Rate of endometritis was narratively reported as similar in outpatient and inpatient groups, but no data were provided.

Catheter versus Dinoprostone (PGE2)

Two fair-quality RCTs compared outpatient double-balloon catheters (80 ml inflation, N=695) or single-balloon catheters (30 ml inflation, N=101) with inpatient dinoprostone (Appendix E). Catheter insertion was performed by a resident or midwife or doctor following a reassuring cardiotocograph. Post-insertion FHR monitoring was performed 30 minutes after insertion in one study, but was not routinely performed in the other. Thirty minutes after insertion, women in the outpatient group was discharged home with instructions to return the following morning. Women were instructed to return sooner if labor began or there were other clinical signs (e.g., vaginal bleeding, unmanageable pain, catheter expulsion). In the inpatient groups, intravaginal dinoprostone gel 2 mg or controlled-release vaginal tape 10 mg was administered by a doctor or midwife in one study, while the other study used intravaginal
dinoprostone gel 2 mg for nulliparous women and 1 mg for parous women, with the type of provider administering not reported. Both studies conducted 30 minutes of cardiotocographic monitoring after dinoprostone insertion, with one reassessing after 6 hours and allowing a second dose (criteria and proportion not reported). Both included women with ≥37 weeks gestation, indications for induction of labor (either “low-risk” indications or “unfavorable cervix”), and Bishop score <7. Both studies also excluded women with a history of cesarean delivery. Both were conducted in Australia. One reported no external funding, and the other did not report its funding source.

Patient characteristics are described in Appendix E-1. The weighted mean age was 30.5 years, and mean BMI was 26.4. In both studies, the majority of women were nulliparous (73% overall) and the mean gestational age was 41 weeks. Five percent of women had gestational diabetes and one trial reported that 14 percent of women were GBS positive. Post due date was the most prevalent reason for induction reported for 72 percent of women, 19 percent for medically-indicated reason, and 9 percent for social/elective reason. Bishop scores at baseline were mean 2.8 in one study and median 3.0 (modified Bishop score) in the other. Scores were similar between randomized groups within these studies.

Differences between outpatient catheter and inpatient dinoprostone were small and not statistically significant for any maternal or fetal/neonatal outcomes, including length of hospital stay, risk of cesarean delivery, risk of failed induction, cervical state at time of admission, time from ROM to delivery, maternal infection, umbilical cord prolapse, neonatal infection, or umbilical cord arterial pH <7.10.

Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

**Total Time from Admission to Vaginal Birth and Total L&D Length of Stay**

Neither study reported on the primary birth outcomes relating to time to delivery prioritized for this report. Both studies reported total duration of hospital stay, finding outpatient catheter use associated with a shorter duration relative to dinoprostone (pooled MD –8.28 hours, 95% CI –21.48 to 4.92).

Other related outcomes were time from induction or time from randomization to delivery (cesarean or vaginal). In one study, there was no difference between outpatient catheters and inpatient dinoprostone in the time from induction of labor to delivery (24.2 vs. 23.7 hours; MD 0.50, 95% CI –8.38 to 9.38). The other study reported time from randomization to delivery, finding a significantly longer duration in the inpatient group (21.3 vs. 32.4 hours; MD –11.1, 95% CI -16.5 to –5.7). When pooled, there was no difference between outpatient and inpatient groups (MD -6.46, 95% CI –19.3 to 8.37) though heterogeneity was high (I²=79%).

**Cesarean Delivery**

Risk of cesarean delivery was slightly higher in the outpatient catheter groups versus inpatient dinoprostone in both studies, but the pooled estimate is not statistically significant (33% vs. 26%; RR 1.24, 95% CI 0.88 to 1.70, I²=0%; Figure 4). Based on type of catheter, the estimates were not significantly different (RR 1.26, 95% CI 0.95 to 1.69 for double-balloon catheter and RR 1.16, 95% CI 0.65 to 2.05 for single-balloon catheter). One study conducted subgroup analyses of women with modified Bishop score >3 at the start of cervical ripening (N=217), finding that outpatient (double-balloon) catheter use increased the risk of cesarean delivery compared with inpatient dinoprostone, though the estimate was not statistically
significant (31% vs. 20%; RR 1.53, 95% CI 0.96 to 2.46). In the same subgroup, women in the outpatient catheter group were less likely to have an unassisted vaginal birth than those in the inpatient dinoprostone group (54% vs. 72%; RR 0.75, 95% CI 0.61 to 0.92). This is low-strength evidence.

**Figure 4. Meta-analysis of cesarean delivery with outpatient catheters versus inpatient dinoprostone for cervical ripening**

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Quality</th>
<th>Treatment, Control,</th>
<th>Risk Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beckmann 2019</td>
<td>Cook catheter Double balloon Fair</td>
<td>70/215 60/233</td>
<td>1.26 (0.95, 1.69)</td>
</tr>
<tr>
<td>Henry 2013</td>
<td>Foley catheter Single balloon Fair</td>
<td>17/50 15/51</td>
<td>1.16 (0.65, 2.05)</td>
</tr>
<tr>
<td>Subgroup (I-squared = 0.0%, p = 0.785)</td>
<td></td>
<td></td>
<td>1.24 (0.88, 1.70)</td>
</tr>
</tbody>
</table>

CI = confidence interval; IP = inpatient; OP = outpatient

**Other Birth Outcomes**

Both studies reported the rate of failed induction, though they were inconsistent in how this was defined. The studies found that use of a double-balloon catheter was associated with small, but not statistically significant increase in risk of cesarean delivery for failed induction (3% vs. 2%; RR 1.63, 95% CI 0.47 to 5.68), cesarean delivery for fetal distress (10% vs. 9%; RR 1.19, 95% CI 0.67 to 2.12) or cesarean delivery due to “fetal concerns” (46% vs. 31%; RR 1.47, 95% CI 0.88 to 2.43) relative to inpatient dinoprostone. One study found little difference in cervical dilation (1 cm in both groups) and modified Bishop score (3 in both groups) on admission and in time from ROM to delivery (10 vs. 8.3 hours; MD 1.70, 95% CI –3.42 to 6.82) (Appendix E-2).

**Fetal/Neonatal Harm Outcomes**

One trial reported that there were no cases of perinatal mortality, hypoxic-ischemic encephalopathy, seizure, or need for respiratory support 72 hours after birth. Confirmed infection was rare, with little difference in event rates between outpatient double-balloon catheters and inpatient dinoprostone in one study (0.93% [2/215] vs. 0% [0/233]; RR 5.42, 95% CI 0.26 to 112). This evidence is insufficient (see Appendix H).

Both studies reported a lower incidence of umbilical cord arterial pH <7.10 in the catheter arms versus dinoprostone, though absolute event rates were low (range 0 to 10%) and risk estimates were not statistically significant in either study (RR 0.38, 95% CI 0.12 to 1.16 and RR 0.51, 95% CI 0.10 to 2.66). Subgroup analysis found that risk of umbilical cord arterial pH <7.10 was consistent with overall risk estimates when analyzed according to parity (nulliparous: RR 0.46, 95% CI 0.15 to 1.44; parous: RR 0.22, 95% CI 0.01 to 4.33).
Maternal Harm Outcomes

Both studies reported the rate of postpartum hemorrhage, although neither study stratified results according to mode of delivery. Outpatient double-balloon and single-balloon catheters were associated with rates of postpartum hemorrhage compared with inpatient dinoprostone that were not significantly different. In the study comparing double-balloon catheters versus dinoprostone, postpartum hemorrhage (blood loss >500 ml) occurred in 30 percent versus 26 percent of study participants (RR 1.15, 95% CI 0.86 to 1.55), while in the study comparing outpatient single-balloon catheters with inpatient dinoprostone, (undefined) postpartum hemorrhage rates were 16 percent versus 22 percent, respectively (RR 0.74, 95% CI 0.33 to 1.69). When pooled, there remained no statistically significant difference between outpatient catheters and inpatient dinoprostone in risk of postpartum hemorrhage (28% vs. 25%; RR 1.10, 95% CI 0.62 to 1.56, I²=0%; Appendix F). This is low strength of evidence.

There was no statistically significant difference between catheters and dinoprostone in maternal infection (0% [0/215] vs. 0.43% [1/233]; RR 0.36, 95% CI 0.01 to 8.82) or umbilical cord prolapse (0.47% [1/215] vs. 0% [0/233]; RR 3.25, 95% CI 0.13 to 79.36) based on imprecise estimates from one study. This evidence is insufficient to draw conclusions (see Appendix H).

Key Question 3. Comparative effectiveness and harms of cervical ripening in the outpatient setting

Key Points

- In head-to-head comparisons for outpatient cervical ripening, there was little difference in the frequency of cesarean delivery for comparisons of dinoprostone gel 2.5 mg versus 5 mg, or latex versus silicone single-balloon catheters (1 study each). Similarly, the risk of receiving antibiotics for suspected uterine infection did not differ between catheter types (SOE: low).
- Based on 12 RCTs comparing a prostaglandin with placebo for outpatient cervical ripening, differences were small and did not reach statistical significance for cesarean delivery (7 RCTs), birth trauma (3 RCTs), or uterine infection (7 RCTs). These findings did not change according to the specific prostaglandin or route of administration, study quality, or gestational age (SOE: low).
- Comparisons of prostaglandins (primarily dinoprostone) with membrane sweeping or expectant management (6 RCTs) found small differences in the frequency of cesarean delivery that were not statistically significant. The incidence of uterine infection was low and not different between groups (SOE: low). These findings did not change according to the specific prostaglandin or study quality.

Summary of Findings

While a few of the studies noted the types of providers that inserted or applied the cervical ripening intervention (i.e., obstetrician or midwife), it was not possible to evaluate any potential differences in outcomes based on this characteristic as results were not stratified based on this characteristic. Details on providers (where reported) specifics on how the drugs were placed or administered are given in Appendix E-1. While there were studies using various routes or methods of administration (e.g., intracervical or intravaginal for dinoprostone, oral or
intravaginal for misoprostol), the results (below) did not indicate variation according to this factor, and subgroup analysis of this characteristic was not undertaken due to small numbers of studies within each outcome.

The most commonly reported primary outcome was cesarean delivery, with few studies reporting on measures of time to vaginal birth or length of stay in labor and delivery. Studies rarely reported primary fetal/neonatal or maternal harms, and when reported were less clearly defined than outlined for this review. For example, neonatal sepsis was reported without indicating that it was confirmed (versus suspected). Some secondary outcomes were reported, but again, with less specificity than desired. For example, failed induction was not defined in some, and defined without a threshold for the degree if dilation in others. In this Key Question, we included additional longer-term outcomes of breastfeeding, maternal mood, and mother-baby attachment, however, none of these studies reported on these outcomes. A few studies reported secondary neonatal harm outcomes of Apgar scores and umbilical cord pH, but none of these studies reported on secondary maternal harms.

**Prostaglandin versus Prostaglandin: Misoprostol (PGE1) Versus Dinoprostone (PGE2)**

Four RCTs compared prostaglandins for cervical ripening in the outpatient setting, involving 297 women. A fair-quality study compared intravaginal misoprostol 25 mcg with intracervical dinoprostone gel 0.5 mg, and a poor-quality study compared an oral tablet of dinoprostone (0.5 mg every hour x 6 hours; not available in the United States) with a pharmacy-compounded intracervical gel 3 mg. Two trials compared different doses of prostaglandins: a good-quality trial evaluated intravaginal dinoprostone gel 2.5 mg versus 5 mg, and a fair-quality study oral misoprostol 25 mg versus 50 mcg. Enrollment criteria varied, and were only partially specific. Reason for induction as an eligibility criterion was postterm pregnancy in one, and not clearly reported in the others. For enrollment, women had to have Bishop scores of <4 in two trials, <5 in one, and <6 in the fourth. All of the trials required a singleton pregnancy. Given the variation in study drugs, the protocols for drug administration and followup also varied (see Appendix E-1). All but one of the studies required a period of 2 to 3 hours of continuous FHR monitoring prior to discharge, with most studies requiring women to return 12 to 24 hours after discharge home (after 3 days for the oral misoprostol dose study). None of the RCTs reported their source of funding.

Characteristics of women enrolled are in Appendix E-1. Across the trials, the weighted mean age of women enrolled was 25 years, one enrolled only nulliparous women, two enrolled 60 percent nulliparous (weighted mean), and one reported that 63 percent were primagravida. Weighted mean gestational age was 40 weeks. None of the studies reported BMI. Prior cesarean delivery was excluded by one, present in 12 percent in another, and not reported in two studies. One study reported that 12 percent of women had diabetes and 20 percent had hypertension/preeclampsia (further details not provided). Postterm pregnancy was the reason for induction in all women in one study, and 35 percent in two other studies (one study did not report on reasons for induction). Bishop score at enrollment was 3 to 4 for most women in these trials.

The primary outcomes for these studies were cesarean delivery rates, with no (or only globally reported) primary fetal or maternal harm outcomes reported. Overall, differences between groups were small and not found to be statistically significant. The bodies of evidence for primary outcomes with each comparison are mainly insufficient to draw conclusions due to
small sample sizes, no corroborating evidence, and imprecision of estimates. The exception was for the dose comparison of dinoprostone gel 2.5 mg vs. 5.0 mg, which was low strength for cesarean delivery.

Detailed study results can be found in Appendix E-2.

Birth Outcomes

Cesarean Delivery

All four trials comparing prostaglandin interventions to each other for outpatient cervical ripening reported on the incidence of cesarean delivery, with none finding statistically significant differences between groups. Across the studies, 19 to 32 percent of women enrolled delivered via cesarean delivery. The highest quality of evidence was for the dose comparison of dinoprostone gel (2.5 mg vs. 5.0 mg), with 20 percent and 19 percent having a cesarean delivery (respectively) (RR 0.90, 95% CI 0.24 to 1.65; SOE: low). For the comparison of misoprostol 25 mcg intravaginal and dinoprostone gel 0.5 mg intracervical, 21 percent and 19 percent (respectively) had a cesarean delivery (RR 1.13, 95% CI 0.48 to 2.63). However, this evidence, and evidence for the other two comparisons, is insufficient to draw conclusions due to small sample size (imprecision), study limitations, and unknown consistency.

Other Birth Outcomes

Three of the trials reported on secondary birth outcomes, however none reported the outcomes as they were specified for this review (see Appendix E-2). Differences in “failed induction” (described as <6 cm dilation after 1 to 3 doses) between groups were small and not statistically significant for any comparison. The comparisons included intravaginal misoprostol 25 mcg and intracervical dinoprostone gel 0.5 mg, dinoprostone gel 2.5 mg and 5.0 mg, and oral versus intracervical dinoprostone gel. The incidence of “failed induction” ranged from 0 percent with intravaginal misoprostol to 32 percent with oral dinoprostone. In the study comparing intravaginal misoprostol and intracervical dinoprostone, failed induction was also reported as cesarean delivery due to dystocia (7% vs. 10%) and again, the difference was not statistically significant. In the study of oral versus intracervical dinoprostone, failed induction defined as cesarean delivery due to fetal distress occurred in 4 percent versus 8 percent, and cesarean delivery excluding fetal distress occurred in 16 percent versus 24 percent, with the differences not being statistically significant.

Similar to failed induction, the trials did not report cervical assessment in the way specified for this review (at the time of admission), though the two trials reporting on Bishop score did not find differences between groups. Among women not in labor 18 to 24 hours after treatment, the mean score was 5 in the intravaginal misoprostol and 4 in the intracervical dinoprostone groups (p=0.28). In the dose-comparison study of dinoprostone gel, the proportion of women whose Bishop score changed more than 3 points from baseline to after the second dose was 44.1 percent versus 45.7 percent (p=0.29).

Fetal/Neonatal Harm Outcomes

No study comparing different prostaglandins for cervical ripening in the outpatient setting reported on the primary fetal/neonatal harm outcomes specified for this review. The dose-comparison study of dinoprostone gel reported that 1.8 percent of infants (1/55) in the 2.5 mg group and 4.7 percent (3/64) in the 5.0 mg group had cord gas pH of <7.2 (RR 0.39, 95% CI 0.04 to 3.62). None of these RCTs reported on Apgar scores <3 at 5 minutes, but two reported on
scores less than 7 at 5 minutes. In the dose-comparison study of dinoprostone gel, 0 percent (0/55) in the 2.5 mg group and 4.7 percent (3/64) in the 5 mg group (RR 0.17, 95% CI 0.01 to 3.14) had Apgar scores <7 at 5 minutes.

In the study of oral versus intracervical dinoprostone gel, 4 percent (1/25) in each group had Apgar scores <7 at 5 minutes (RR 1.00, 95% CI 0.07 to 15.12).

**Maternal Harm Outcomes**

The two head-to-head RCTs did not report maternal harms. The other two trials only reported in text as no adverse events or no complications resulting from the use of prostaglandins.

**Prostaglandin versus Placebo**

Twelve RCTs compared prostaglandins for cervical ripening with placebo in the outpatient setting, involving 1,112 women. Four studies evaluated intravaginal misoprostol (25 mcg), four evaluated intracervical dinoprostone gel 0.5 mg, three evaluated intravaginal dinoprostone gel/suppository 2 mg, and one evaluated oral misoprostol 100 mcg. Enrollment criteria included preventing a postterm pregnancy in seven and unclear or not reported in the others. For enrollment, women had to have Bishop score of ≤8 in four trials, ≤6 in three trials, and ≤4 in three trials. The remaining two trials did not have a required Bishop score for enrollment. At baseline, the mean Bishop score was 4.2 in one and 66 percent of women had a Bishop <6 in the other. Singleton pregnancy was required in nine trials and not reported in three. Protocols for drug administration varied widely (see Appendix E-1), both by drug (misoprostol versus dinoprostone), route (oral, intravaginal, intracervical), and indication (postterm pregnancy) ranging from a single dose, daily doses, every 3 to 4 days, or weekly (for multiple doses, up to 2 to 3 doses or until 42 to 44 weeks gestation allowed). After placement of the drug, FHR monitoring was required from 40 minutes to 2 hours prior to discharge home. Eight studies did not report their source of funding, while one received government, two received nonprofit, and one received industry funding. Three of the trials were rated good quality, one was rated poor, and the rest were fair quality.

Characteristics of the women enrolled are in Appendix E-1. Across the trials, the weighted mean age of women enrolled was 25 years, gestational age was 39.8 weeks, and 49 percent were nulliparous. BMI was reported in only two studies (27.6 and 30.4 mg/kg/m²). Most of the studies did not report on, or excluded prior cesarean delivery and diabetes, with one enrolling only women with diabetes (55% gestational). Other comorbidities of interest were not reported. As noted above, the reason or induction was postterm pregnancy in seven, and either a mix of reasons or unclear in the rest. Bishop score at enrollment ranged from 2.9 to 5.5.

All of the RCTs reported the incidence of cesarean delivery, and most reported on the time from prostaglandin first dose to delivery (mode not specified in most). Eight studies reported primary fetal harms outcomes, including meconium aspiration syndrome, birth injury, infection, and neonatal encephalopathy. Eight studies reported on maternal primary harm outcomes, mainly postpartum hemorrhage and uterine infections.

Overall, differences between groups were small and the differences were not found to be statistically significant. The strength of evidence for primary outcomes with each comparison are low (meaning that it is likely future studies would change the estimate meaningfully) or insufficient to draw conclusions due to small sample sizes, no corroborating evidence, and
imprecision of estimates. Similarly, analysis of prespecified subgroups for this review, using meta-regression of pooled studies or from within-study subgroup analyses, were inadequate to draw conclusions, primarily due to small sample sizes.

Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

**Total Time from Admission to Vaginal Birth and Total L&D Length of Stay**

A single, fair-quality RCT reported on one of the prespecified primary birth outcomes; a study of dinoprostone gel (N=80) reported that the time in labor and delivery was similar between groups, and the difference was not statistically significant (11.0 hours vs. 11.8 hours; MD –0.8, 95% CI –6.21 to 4.61).56 This is insufficient evidence to draw conclusions due to lack of confirmatory evidence, imprecision due to small sample size, and study limitations. Subgroup analysis according to parity found that the difference in time in labor and delivery (L&D) was slightly greater (i.e., favoring the prostaglandin group) in nulliparous women (10.7 hours vs. 15.3 hours; difference 4.6 hours, p=0.035) than in multiparous women (11.2 hours vs. 7.1 hours, difference 4.1 hours, p=not significant [NS]).

Three studies (N=353), two using dinoprostone gel30,41 in women with postterm pregnancies and one using intravaginal misoprostol in women with diabetes at ≥38 weeks gestation,56 reported on the time from admission to delivery. None stratified results by mode of delivery. The difference between groups was not statistically significant, although in two studies the time was somewhat longer in the prostaglandin groups (combined estimate across the dinoprostone studies was a difference of 1.79 hours, 95% CI –2.68 to 6.34, I²=0% and was 30.30 hours, 95% CI –34.15 to 94.75 for misoprostol).

Although not a prespecified outcome of interest, two RCTs (N=117) evaluated time from drug/placebo placement to delivery (any mode). A good-quality RCT of intravaginal misoprostol found that the difference in time from drug/placebo placement to delivery (any mode) was greater in nulliparous women, and shorter in multiparous women. While this was not our prespecified outcome (time from admission to vaginal birth) the findings are not statistically significant.47 A poor-quality RCT found that the difference between dinoprostone and placebo in time from dosage to delivery was smaller, though still not statistically significant, in women with less favorable Bishop scores at the time of drug/placebo placement (difference of –86 hours in the overall group vs. –47 hours in those with Bishop score <5).27

**Cesarean Delivery**

Seven RCTs of dinoprostone27,30,40,41,45,55,56 and five of misoprostol12,36,47,52,59 reported on the frequency of cesarean delivery (N=381 and 543, respectively) compared with placebo. Although the incidence was slightly lower with prostaglandins (13% vs. 16%), the difference in the combined estimates were not statistically significant (overall pooled RR 0.80, 95% CI 0.58 to 1.09, I²=4.3%; Figure 5). For this outcome, we were able to assess for small sample size bias, and did not find evidence of it using a funnel plot of an Egger test (p=0.969, Appendix F).
Figure 5. Meta-analysis of cesarean delivery prostaglandins versus placebo for cervical ripening in the outpatient setting

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Post-term pregnancy</th>
<th>Quality</th>
<th>Treatment, n/N</th>
<th>Control, n/N</th>
<th>Risk Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinoprostone</td>
<td>No</td>
<td>Fair</td>
<td>5/41</td>
<td>10/43</td>
<td>0.52 (0.20, 1.40)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Poor</td>
<td>4/30</td>
<td>3/31</td>
<td>1.38 (0.34, 5.64)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Fair</td>
<td>5/23</td>
<td>7/20</td>
<td>0.62 (0.23, 1.65)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Fair</td>
<td>3/37</td>
<td>1/28</td>
<td>2.27 (0.25, 20.68)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Fair</td>
<td>6/43</td>
<td>8/47</td>
<td>0.82 (0.31, 2.17)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Fair</td>
<td>6/24</td>
<td>4/26</td>
<td>1.63 (0.52, 5.07)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Fair</td>
<td>1/38</td>
<td>6/42</td>
<td>0.18 (0.02, 1.46)</td>
</tr>
<tr>
<td>Subgroup</td>
<td></td>
<td>(I-squared = 0.0%, p = 0.442)</td>
<td></td>
<td></td>
<td>0.80 (0.50, 1.31)</td>
</tr>
</tbody>
</table>

| Misoprostol  | No                   | Fair    | 14/57          | 11/63       | 1.41 (0.70, 2.84)  |
| PonMalar, 2017 | No                 | Good    | 8/63           | 18/63       | 0.44 (0.21, 0.95)  |
| Gaffaney, 2009 | Yes                | Fair    | 8/43           | 13/44       | 0.63 (0.29, 1.37)  |
| McKenna, 2004b | Yes                | Good    | 9/33           | 9/35        | 1.06 (0.48, 2.34)  |
| Stilley, 2000 | Yes                 | Good    | 4/27           | 8/33        | 0.61 (0.21, 1.81)  |
| Subgroup     |                     | (I-squared = 21.5%, p = 0.207) |             |             | 0.79 (0.48, 1.26)  |

| Overall      |                     | (I-squared = 4.3%, p = 0.384) |             |             | 0.80 (0.58, 1.09)  |

CI = confidence interval.

Meta-regression by type of prostaglandin and by gestational age (determined by enrollment of only postterm pregnancies versus mixed populations) found no significant interaction, with p-value >0.90. Subgroup analyses by study quality (excluding poor-quality RCTs), found very similar pooled estimates (Appendix F). Similarly, subgroup analysis of gestational age (7 studies including only postterm pregnancies versus 5 studies enrolling a mixed population, not specifically including or excluding postterm pregnancies) also found little difference in estimates of cesarean delivery (Appendix F).

Two RCTs (one good-quality of misoprostol and one fair-quality of dinoprostone) conducted within-study subgroup analysis of cesarean delivery frequency according the parity. Although the studies were small (total N=118), and the subgroup analyses did not reach statistical significance, the direction of the effect in both studies varied according to parity. Nulliparous women had higher frequency of cesarean delivery with a prostaglandin than with placebo in both studies, while multiparous women had somewhat smaller difference in frequency (Table 4).
Table 4. Subgroup analyses of cesarean delivery frequency according to parity (prostaglandin versus placebo)

<table>
<thead>
<tr>
<th>Study Quality Rating</th>
<th>Intervention</th>
<th>Population Characteristics</th>
<th>Cesarean Delivery Frequency by Parity Relative Risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>McKenna 2004b Good</td>
<td>Misoprostol 25 µg Intravaginal single dose N=68</td>
<td>Postterm pregnancies Enrollment Bishop score &lt; 9 (mean 4.5 overall) Mean age: 29 years Nulliparous: 57%</td>
<td>All: 27.27% (9/33) vs. 25.71% (9/35) RR 1.06 (0.48 to 2.34) Nulliparous: 40.00% (8/20) vs. 36.84% (7/19) RR 1.09 (0.49 to 2.41) Multiparous: 7.69% (1/13) vs. 12.5% (2/16) RR 0.62 (0.06 to 6.05)</td>
</tr>
<tr>
<td>Sawai, 1991 Fair</td>
<td>Dinoprostone gel 2 mg Intravaginal twice weekly N=50</td>
<td>Postterm pregnancies Enrollment Bishop score &lt; 9 (mean 5 nulliparous, 4 multiparous) Mean age: NR Nulliparous: 60%</td>
<td>All: 25.00% (6/24) vs. 15.38% (4/26) RR 1.63 (0.52 to 5.07) Nulliparous: 42.86% (6/14) vs. 18.75% (3/16) RR 2.29 (0.70 to 7.48) Multiparous: 0% (0/10) vs. 10.00% (1/10) RR 0.33 (0.02 to 7.32)</td>
</tr>
</tbody>
</table>

CI = confidence interval; NR = not reported; RR = relative risk

Other Birth Outcomes

All twelve RCTs reported at least one of the secondary birth outcomes, most commonly cesarean delivery for various reasons (failed induction) and cervical assessment at the time of admission (primarily Bishop score). The incidence of failed induction defined as cesarean delivery without fetal distress was similar between groups (10% vs. 11%) and the difference was not statistically significant (8 RCTs; RR 0.71, 95% CI 0.48 to 1.05, I²=0%). There was a small, but not statistically significant, difference in the incidence of failed induction with fetal distress (4% vs. 6%; 7 RCTs; RR 0.81, 95% CI 0.37 to 1.89, I²=0%). Bishop score at the time of admission was reported in ten RCTs. The combined analysis finds a higher score with prostaglandins versus placebo (10 RCTs; MD 0.48, 95% CI 0.07 to 0.95, I²=24.7%). However, in a sensitivity analysis removing the poor-quality RCT, the difference in mean Bishop score at admission is smaller and no longer significant (9 RCTs, MD 0.30, 95% CI –0.06 to 0.71). Subgroup analysis by gestational age (as indicated by inclusion of postterm pregnancies only versus mixed populations) found a smaller difference in mixed populations than postterm only populations, but again, the poor-quality study affected the results. One poor-quality study reported the incidence of delivery within 24 hours, but did not stratify by mode of delivery. The incidence for any mode of delivery within 24 hours was 17 percent with dinoprostone and 0 percent with placebo.

Fetal/Neonatal Harm Outcomes

Hypoxic-ischemic Encephalopathy

A good-quality RCT (N=126) reported on “neonatal encephalopathy,” not specifically defined as hypoxic-ischemic, with no cases in either the misoprostol or placebo group. This is insufficient evidence to draw conclusions.
**Infection (Confirmed Sepsis or Pneumonia)**

Only one fair-quality RCT reported on confirmed sepsis, with one case of GBS sepsis in the dinoprostone group (3%; 1/32), and none in the placebo group (0/42); RR 3.91 (95% CI 0.16 to 92.91). One other fair-quality RCT (N=143) reported on incidence of probable sepsis, with 11 percent in the dinoprostone group and 6 percent in the placebo group (RR 1.51, 95% CI 0.30 to 7.69). This evidence is insufficient to draw conclusions.

**Meconium Aspiration Syndrome**

Two fair-quality RCTs (N=134) reported on meconium aspiration syndrome, with both assessing dinoprostone. The overall incidence was low (2% versus 4%), with no differences between groups in either study or when combined (RR 0.76, 95% CI 0.03 to 22.33, I^2=0%; Appendix F). In one of these studies, with two cases (of 26 infants) in the control group, it was noted that both were infants of multiparous women. This evidence is insufficient to draw conclusions.

**Birth Trauma**

Three RCTs (N= 270, 1 good- and 2 fair-quality) reported birth trauma (shoulder dystocia), with a combined incidence of 6 percent (7/127) versus 0.70 percent (1/143). None of the studies found the differences to be statistically significant; the pooled analysis results in RR 7.88 (95% CI 0.98 to 63.20). This is low strength of evidence.

**Other Fetal/Neonatal Outcomes**

All twelve RCTs reported at least one secondary fetal harm outcome (Appendix E-2). Ten RCTs reported an outcome related to Apgar scores; three reported that there were no infants with a score of ≤3 or <7 at five minutes. Four RCTs (N=293) reported at least one infant with a score <7 at five minutes, with a lower incidence in the prostaglandin groups (1/138, 0.73% vs. 6/155 [4%]) that did not reach statistical significance in individual studies or in the pooled estimate (RR 0.40, 95% CI 0.09 to 1.81, I^2=0%). Three other RCTs reported the mean Apgar score at 5 minutes, with no differences between groups (mean 9.1 in both groups). A good-quality RCT (N=60) reported one case of persistent pulmonary hypertension, requiring extracorporeal membrane oxygenation therapy in the placebo group (1/33 [3%]) and none in the misoprostol group (0/27). One fair-quality RCT (N=82) reported that the mean umbilical cord gas at delivery was identical between dinoprostone and placebo (mean 7.27 in both groups), and no study reported the incidence of umbilical cord gas pH <7.0 or 7.1. Although admission to NICU was not a prespecified secondary outcome, since so few studies reported eligible outcomes, we noted that it was reported in 10 RCTs, with none finding a difference between groups.

**Maternal Harm Outcomes**

**Uterine Infection**

Seven RCTs (N=771) reported on the incidence of uterine infection (chorioamnionitis or endometritis). While there was a small difference between groups favoring placebo, the difference was not statistically significant (7% vs. 5%; RR 0.75, 95% CI 0.40 to 1.39, I^2=0%; Appendix F). This is low strength of evidence. Meta-regression analysis did not find significant interaction based on type of prostaglandin, or on gestational age (as determined by studies that enrolled only women with postterm pregnancies versus those with mixed
populations). Sensitivity analysis, removing two RCTs that reported endometritis, resulted in similar findings. For this outcome we were also able to conduct an analyses of small sample size bias, finding no evidence of it (Egger test p=0.981; Appendix F).

**Postpartum Hemorrhage**

Four RCTs (N=339) reported on overall incidence of postpartum hemorrhage, but not by mode of delivery. The pooled incidence with any delivery mode was 2 percent (3/173) in the prostaglandin groups and 0.60 percent (1/166) in the placebo groups (p=0.36) (Appendix F). No study reported on hemorrhage requiring transfusion. This evidence is insufficient to draw conclusions.

**Other Maternal Outcomes**

One good-quality RCT (N=126) reported no cases of uterine rupture.

**Prostaglandin versus Other Approaches**

Seven RCTs compared prostaglandins for cervical ripening with various other approaches, including expectant management, membrane sweeping, and estradiol cream in the outpatient setting. Six RCTs, involving 919 women, compared dinoprostone (3 trials, N=538) or misoprostol (1 trial, N=77) versus expectant management, and dinoprostone versus membrane sweeping (3 trials, N=339). Four were conducted in the United States, one in Israel, and one in Nigeria. One was funded by industry, two by a nonprofit organization, and the remaining three trials did not report their funding sources. An additional RCT conducted in the United States compared dinoprostone versus estradiol cream; it was funded by a nonprofit organization. Five were rated fair quality and two were rated poor quality due to lack of assessor blinding, poor reporting of patient characteristics, failure to report intent-to-treat (ITT) analysis, and high attrition.

**Dinoprostone (PGE2) and Misoprostol (PGE1) Versus Expectant Management**

Four RCTs, involving 615 women, compared prostaglandins with expectant management for cervical ripening in the outpatient setting (Appendix E). Two trials (1 fair- and 1 poor-quality) administered intracervical dinoprostone gel 0.5 mg, with repeat doses given either daily or weekly as necessary. One poor-quality trial administered an intravaginal dinoprostone tablet 3 mg, with women returning 3 to 4 days later for a second dose. A fourth fair-quality trial administered a single dose of an intravaginal misoprostol tablet 25 mcg. Only one trial reported that the mean number of doses used was 1.5. Treatment protocols for expectant management varied across the studies, but generally included a vaginal examination, either daily, weekly, or bi-weekly until spontaneous labor or the need for induction occurred. All of the trials required a period of continuous FHR monitoring prior to discharge for those women who received prostaglandins (2 hours in one, 1 hour in one, and 2 did not specify the duration). In one trial, it was noted that if a second dose of study drug was required, FHR monitoring was only completed if clinically indicated. None of the trials reported the type of provider inserting the prostaglandin or conducting examinations.

Characteristics of the women enrolled are in Appendix E-1. Across the trials, the weighted mean age was 27 years and the weighted mean gestational age was 39.8 weeks. One trial enrolled only multiparous women with a history of prior cesarean delivery and another trial did
not specify parity by groups but did report that mean gravidity was 2.3.50 Two trials enrolled 58 percent nulliparous women (weighted mean).43,49 One trial excluded women with prior cesarean delivery49 and two trials43,50 did not report on previous cesarean delivery. One trial53 excluded women with insulin-dependent diabetes and pregnancy-induced hypertension (no additional information provided). No other trials reported on diabetes or hypertension and none of the trials reported BMI. Reason for induction was elective (to prevent postterm pregnancy) in two trials,49,50 postterm pregnancy (>41 weeks) in one trial,43 and unclear in the final trial (main objective of this trial was to increase the rate of vaginal birth after prior cesarean delivery).53 Weighted mean baseline Bishop score was 4.8 (across three trials)43,49,50; median Bishop score was 2.0 in the fourth trial.53

Across all outcomes reported, differences between groups were small and did not reach statistical significance. The only primary birth outcome reported was cesarean delivery, with little to no difference groups. Primary harms as prespecified for this review were infrequently reported; there was little to no difference in uterine infection between groups. This is low-strength evidence. Evidence on other harms outcomes was insufficient to draw conclusions, primary due to small sample sizes and lack of detailed reporting.

Detailed study results can be found in Appendix E-2.

Birth Outcomes

Cesarean Delivery

The incidence of cesarean delivery was similar for prostaglandins (27%) versus expectant management (26%) across four trials (pooled RR 0.95, 95% CI 0.68 to 1.33, I²=0%; Figure 6).43,49,50 Exclusion of the two poor-quality trials43,50 resulted in a somewhat lower incidence of cesarean delivery with any prostaglandin, though an increased incidence overall, (35.4% vs. 39.5%; pooled RR 0.92, 95% CI 0.45 to 1.30, I²=0%) but the difference was not statistically significant (Appendix F). When stratified by type of prostaglandin, the incidence of cesarean delivery was similar between dinoprostone (29%) and expectant management (27%) across three trials (1 fair- and 2 poor-quality; pooled RR 0.98, 95% CI 0.73 to 1.30, I²=0%; Appendix F)43,50,53 but was lower with misoprostol in one fair-quality trial (8% vs. 18%; RR 0.44, 95% CI 0.12 to 1.58).49 Again, none of the differences were statistically significant. This evidence is low strength.

Other Birth Outcomes

With the exception of Bishop score, authors did not report secondary birth outcomes as specified for this review.

One poor-quality trial reported that women who received dinoprostone compared with expectant management were five time more likely to deliver (vaginal or cesarean) within 24 hours from entry into the trial (66% vs. 13%; RR 4.88, 95% CI 2.91 to 8.18).50

Two fair-quality trials reported the incidence of failed induction, defined as cesarean delivery due to and excluding fetal distress. Cesarean delivery due to fetal distress was less common with prostaglandins across both trials (7% vs. 8%; pooled RR 0.79, 95% CI 0.25 to 2.05, I²= 0%)49,53 but the difference did not reach statistical significance. For cesarean delivery excluding fetal distress, results were inconsistent across the trials with the larger RCT reporting a higher incidence with dinoprostone (23% vs. 16%; RR 1.45, 95% CI 0.90 to 2.33)53 and the other reporting a lower incidence with misoprostol (5% vs. 13%; RR 0.41, 95% CI 0.08 to 1.99)49 versus expectant management, although neither reached statistical significance. In addition to the
difference in the specific prostaglandin used, the study populations varied. Only multiparous women with a history of prior cesarean delivery enrolled in the dinoprostone trial; the misoprostol trial enrolled 53 percent nulliparous women and excluded those with prior cesarean delivery. An additional trial evaluating dinoprostone stated that there were no cases of cesarean delivery for failed induction (not further defined) in either group.

Two fair-quality trials reported mean Bishop scores at time of admission which were similar between prostaglandins and expectant management (pooled difference –0.08, 95% CI -0.70 to 0.87, I²=0%). A third, poor-quality trial noted that the Bishop score on admission was significantly more favorable among those who received dinoprostone versus expectant management (p<0.001; data not reported).

**Fetal/Neonatal Harm Outcomes**

**Perinatal Mortality**

One fair-quality trial evaluating misoprostol reported a single case of stillborn birth which occurred in the expectant management group (0% [0/38] vs. 2.6% [1/39]; RR 0.34, 95% CI 0.01 to 8.14); no information regarding the timing or circumstances surrounding the death was reported.

**Infection**

One fair-quality trial (N=294) evaluating dinoprostone indicated that neonates in both groups had prolonged nursery stays for the same reasons, which included suspected sepsis; no other information was provided.

**Other Fetal/Neonate Harm Outcomes**

Secondary outcomes were reported as specified for this review only for Apgar scores at 5 minutes and umbilical artery pH <7.0. A fair-quality RCT (N=294) reported a single neonate in the dinoprostone group had an Apgar score ≤3 at 5 minutes (0.70% [1/143] vs. 0% [0/151] with expectant management). The frequency of neonates with Apgar scores ≤7 at 5 minutes was similar between dinoprostone (7%) and expectant management (8%) across two trials (pooled RR 0.90, 95% CI 0.31 to 2.74). The other two trials (1 dinoprostone and 1 misoprostol) reported mean Apgar score at 5 minutes which was also similar between groups (pooled difference 0.04, 95% CI –0.03 to 0.11), with scores ranging from 9.2 to 9.5 across all groups. One poor-quality trial reported that 20 percent (7/35) of neonates in both the dinoprostone and expectant management group had an umbilical artery pH <7.2.

One fair-quality trial reported that breathing complications that resulted in prolonged nursery stays occurred in both groups; no other information was provided. This same trial reported that 12 percent (17/143) versus 8 percent (11/151) of neonates in the dinoprostone and expectant management groups, respectively, required resuscitation (no further information provided) with no significant difference between groups (RR 1.63, 95% CI 0.79 to 3.36).

**Maternal Harm Outcomes**

**Postpartum Hemorrhage**

One fair-quality trial reported two cases of postpartum hemorrhage (undefined, mode of delivery not reported), one (3%) in each group (misoprostol [1/38] vs. expectant management [1/39]; RR 1.03, 95% CI 0.07 to 15.82).
Uterine Infection

Endometritis occurred with similar frequency in the dinoprostone (6% [8/143]) versus the expectant management group (5% [7/151]) in one fair-quality trial (RR 1.21, 95% CI 0.45 to 3.24);53 this trial also reported no cases of uterine rupture in either group. A second fair-quality trial evaluating misoprostol reported the frequency of "infectious morbidity" which was also similar between groups, respectively (5% [2/38] vs. 8% [3/39]; RR 0.68, 95% CI 0.12 to 3.87).49 This is low-strength evidence.

Dinoprostone (PGE2) Versus Membrane Sweeping

Three RCTs,30,42,43 involving 339 women, compared dinoprostone versus membrane sweeping for cervical ripening in the outpatient setting (Appendix E-1). The routes of administration and dosages for dinoprostone varied across the trials and included intravaginal gel 2 mg (1 fair-quality trial),30 intracervical gel 0.5 mg (one poor-quality trial),43 and an intravaginal insert 10 mg (1 fair-quality trial).42 In two trials,42,43 both treatments were administered in clinic on a daily basis until spontaneous labor or rupture of membranes; if women achieved a Bishop score of ≥8 or reached 42 weeks gestation, they were admitted for labor induction. In the third trial,30 if necessary, repeat doses/sweepings were administered 1 week after the first administration and then 3 to 4 days thereafter to a maximum gestational age of 43 weeks, at which time they were admitted. None of the RCTs reported the type of provider administering either treatment. In one trial,30 approximately 34 percent of women required more than one dose/sweep (number of administrations not reported in the other trials). One trial30 administered a placebo gel in addition to membrane sweeping. Two trials42,43 performed fetal monitoring (nonstress test and amniotic fluid index) daily in the dinoprostone groups; in one of these trials, identical testing was performed every 3 days in the membrane sweeping group43 while the second trial did not specify their fetal monitoring protocol in this group. A third trial performed FHR monitoring continuously for a minimum of 1 hour after treatment; women were instructed to perform daily kick counts and repeat fetal testing was performed at 42 weeks and then every 3 to 4 days until 43.9 weeks gestation.30

Characteristics of the women enrolled are in Appendix E-1. Across the trials, the weighted mean age of women enrolled was 26 years and 57 percent were nulliparous. Gestational age was similar across all trials (weighted mean 41 weeks in 2 trials42,43; median 41 weeks in 1 trial).30 Women with prior cesarean delivery were excluded by one trial30 and not reported in the others.42,43 None of the studies reported BMI or proportion with comorbidities such as diabetes or hypertension. Reason for induction as an eligibility criterion was postterm pregnancy (≥41 weeks) in all three trials.30,42,43 For enrollment, women had to have Bishop scores of <4 in two trials (weighted mean 2.8)42,43; the third trial30 did not have a required Bishop score for enrollment, though 66 percent of women had a Bishop <6 at baseline. There were no between-group differences in Bishop scores at enrollment.

Few of the prespecified outcomes were reported. There was low-strength evidence that the risk of cesarean delivery did not differ between groups. Other outcomes were generally similar between groups, and mostly deemed insufficient to draw conclusions due to small sample sizes and imprecision.

Detailed study results can be found in Appendix E-2.
Birth Outcomes

Total Time from Admission to Vaginal Birth

Two fair-quality trials reported the time interval from admission to delivery for any mode of delivery, not the prespecified outcome of vaginal birth. Including cesarean deliveries, it was significantly longer in the dinoprostone versus the membrane sweeping group across the trials (pooled difference 2.64 hours, 95% CI 0.47 to 4.88, I²=0%).\(^{30,42}\)

Cesarean Delivery

The incidence of cesarean delivery was greater with dinoprostone (22%) compared with membrane sweeping (15%) across three RCTs, however the difference did not reach statistical significance (pooled RR 1.44, 95% CI 0.85 to 2.36, I²=0%; Figure 6).\(^{30,42,43}\) Exclusion of the poor-quality trial\(^{43}\) resulted in a similar estimate (2 RCTs, pooled RR 1.40, 95% CI 0.64 to 2.65, I²=0%; Appendix F).\(^{30,42}\) This evidence is low strength.

Figure 6. Meta-analysis of cesarean delivery with prostaglandins versus expectant management and membrane sweeping for cervical ripening in the outpatient setting

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Quality Rating</th>
<th>Treatment, n/N</th>
<th>Control, n/N</th>
<th>Risk Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectant Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magann, 1998</td>
<td>Dinoprost</td>
<td>Poor</td>
<td>8/35</td>
<td>5/35</td>
</tr>
<tr>
<td>Ohel, 1996</td>
<td>Dinoprost</td>
<td>Poor</td>
<td>4/70</td>
<td>6/104</td>
</tr>
<tr>
<td>Rayburn, 1999</td>
<td>Dinoprost</td>
<td>Fair</td>
<td>61/143</td>
<td>68/151</td>
</tr>
<tr>
<td>Oboro, 2005</td>
<td>Misoprostol</td>
<td>Fair</td>
<td>3/38</td>
<td>7/39</td>
</tr>
<tr>
<td>Subgroup (I-squared = 0.0%, p = 0.490)</td>
<td></td>
<td></td>
<td></td>
<td>0.95 (0.68, 1.33)</td>
</tr>
<tr>
<td>Membrane Sweeping</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doany, 1997</td>
<td>Dinoprost</td>
<td>Fair</td>
<td>3/37</td>
<td>4/50</td>
</tr>
<tr>
<td>Magann, 1998</td>
<td>Dinoprost</td>
<td>Poor</td>
<td>8/35</td>
<td>5/35</td>
</tr>
<tr>
<td>Magann, 1999</td>
<td>Dinoprost</td>
<td>Fair</td>
<td>25/91</td>
<td>17/91</td>
</tr>
<tr>
<td>Subgroup (I-squared = 0.0%, p = 0.870)</td>
<td></td>
<td></td>
<td></td>
<td>1.44 (0.85, 2.36)</td>
</tr>
<tr>
<td>Overall (I-squared = 0.0%, p = 0.436)</td>
<td></td>
<td></td>
<td></td>
<td>1.03 (0.80, 1.46)</td>
</tr>
</tbody>
</table>

CI = confidence interval.

Other Birth Outcomes

Two of the trials reported on secondary birth outcomes, however neither reported the outcomes as they were specified for this review (see Appendix E-2). A fair-quality trial\(^{42}\) reported that failed induction due to fetal distress requiring operative delivery (i.e., cesarean or forceps) was twice as frequent with dinoprostone (12%) than membrane sweeping (6%), but the difference was not statistically significant (RR 2.20, 95% CI 0.80 to 6.08). The frequency of operative delivery excluding fetal distress (i.e., cephalopelvis disproportion or transverse arrest) was identical in both groups (18%). The second, poor-quality, trial\(^{43}\) stated that there were no
cases of cesarean delivery for failed induction (not further defined) in either group. Both trials reported Bishop score at time of admission; one reported a statistically significant difference between groups which favored membrane sweeping (mean 6.6 vs. 8.6; MD −1.93, 95% CI −2.66 to −1.20) while the other simply stated that there was no difference between groups (p>0.05).

**Fetal/Neonate Harm Outcomes**

No trial reported primary fetal harms as specified for this report. The frequency of probable or suspected infection (i.e., sepsis) was similar in the dinoprostone and membrane sweeping groups across two fair-quality trials (2 RCTs; 4% vs. 2%; pooled RR 1.96, 95% CI 0.39 to 11.80). A fair-quality trial also reported NICU admission due to meconium (“syndrome” not specified), which occurred in one neonate (1%) each group.

Two trials reported on secondary fetal harms of interest to this report, none of which differed statistically between groups. In both trials, the incidence of neonates with Apgar scores <7 at 5 minutes was low, and not significantly different between groups (none in one study, and one in each group in the other). The incidence of umbilical artery pH <7.2 was identical between groups (24%) across both trials (RR 1.0, 95% CI 0.56 to 1.69, I²=0%).

**Maternal Harm Outcomes**

In one fair-quality trial there were no cases of postpartum hemorrhage (undefined) in either group. The frequency of uterine infection was similar, and not statistically significantly different between groups in two fair-quality trials (4.7% vs 4.3%; RR 1.22, 95% CI 0.56 to 2.75, I²=0%; Appendix F). This is low strength of evidence.

**Dinoprostone (PGE2) Versus Estradiol Cream**

One fair-quality RCT (N=85) compared intracervical dinoprostone gel 0.5 mg versus intravaginal estradiol cream 4 mg. The study also had a placebo arm, and is included in the section above. Women received doses on a weekly basis until the onset of spontaneous labor, rupture of membranes, or an indication for delivery arose. Authors did not report the type of provider administering the treatments or the mean number of doses received. Reason for induction of labor was unclear; women with >37 weeks gestation and Bishop score ≤6, and an uncomplicated pregnancy (no comorbidities) were enrolled. The first 20 women underwent 2 hours of FHR monitoring following drug administration (discontinued because no abnormalities or significant uterine activity were noted in any of the 20 women). The study was conducted in the United States and had funding from a nonprofit organization. Characteristics of women enrolled are in Appendix E-1. Mean maternal age was 22 years, 19 percent were nulliparous, mean gestational age was 37.2 weeks, and mean Bishop score at baseline was 2.9.

Primary outcomes, as specified for this review, were cesarean delivery and uterine infection. However, although the incidence of both were lower with dinoprostone, as the sample size was very small (41 to 43 per group) the differences were not statistically significant and the evidence is insufficient to draw conclusions.

Detailed study results can be found in Appendix E-2.

**Birth Outcomes**

The cesarean delivery rate was lower with dinoprostone (12%) versus estradiol cream (31%), however the difference did not reach statistical significance (RR 0.41, 95% CI 0.16 to 1.06). Regarding secondary birth outcomes, failed induction was not reported as specified for this review and was defined as cesarean delivery for fetal distress (2% in both groups) and cesarean
delivery excluding fetal distress (e.g., dystocia, abnormal presentation) (7% vs. 23%, respectively; RR 0.32, 95% CI 0.10 to 1.09). The mean Bishop score at the time of admission was similar (7.9 vs. 8.0).

Fetal/Neonate Harm Outcomes
The trial did not report primary and secondary fetal harms as specified for this review. One case of meconium aspiration (“syndrome” not specified) occurred in the dinoprostone group (1/41 vs. 0/44; RR 3.21, 95% CI 0.13 to 76.74). The frequency of secondary fetal harm outcomes were similar between groups: mean Apgar score at 5 minutes (9.4 vs. 9.2) and mean arterial cord blood pH (7.32 vs. 7.35).

Maternal Harm Outcomes
The incidence of uterine infection was somewhat lower following dinoprostone versus estradiol cream, but the differences were not statistically significant: chorioamnionitis (2.4% vs. 9.1%; RR 0.27, 95% CI 0.03 to 2.30) and endomyometritis (4.9% vs. 6.8%; RR 0.72, 95% CI 0.13 to 4.07). No secondary maternal harms were reported.

Single-balloon Silicone versus Latex Catheter
A fair-quality RCT (n=534) compared single-balloon catheters made of silicone with those made of latex for cervical ripening in the outpatient setting. Catheters were inserted by “trained obstetric and midwifery staff,” and women were asked to return the following day for induction. If insertion with the assigned catheter failed, the other could be tried, and if that failed, medication methods could be tried. As a result, 97 percent of those assigned to the silicone single-balloon and 91 percent of those assigned to the latex single-balloon received the assigned intervention. Analyses were conducted based on ITT assumptions. Enrollment criteria were not specific to reason for induction, but Bishop score <7 and gestational age ≥36 weeks were required, and prior cesarean delivery or multiple pregnancies were not excluded. A 2-hour period of FHR monitoring was required after catheter insertion. The study was locally funded.

Characteristics of women enrolled are in Appendix E-1. Seventy-one percent enrolled were 25 to 35 years old, 59 percent were nulliparous, 30 percent were categorized as overweight and 22 percent as obese. The median gestational age was 39 weeks and 6 days. At baseline, the proportion of women with diabetes or hypertension were not reported, 5 percent had a prior cesarean delivery, and 6 percent received antibiotics for GBS prophylaxis. Postterm pregnancy was the reason for induction in 31 percent, and medically indicated in 65 percent of women. Cervical dilation at baseline was less than 1 centimeter in 65 percent, and 1 to 2 centimeters in the rest; Bishop scores were ≤4 in 35 percent and 5 to 6 in the rest.

The only primary birth outcome reported was cesarean delivery, with the primary fetal harm outcome of infection, and maternal harms of infection and hemorrhage reported. Overall, differences between groups were small and not found to be statistically significant (Appendix E-2). The strength of evidence was low for more common outcomes (>5%; cesarean delivery and chorioamnionitis), and insufficient to draw conclusions for rare outcomes (neonatal infection, postpartum hemorrhage).

Birth Outcomes
The cesarean delivery rate was 39 percent versus 40 percent with silicone versus latex single-balloon catheters (RR 0.98, 95% CI 0.80 to 1.22). The secondary birth outcome of the mean Bishop score at the time of admission (removal of catheter) was 5.9 in both groups.
Fetal/Neonatal Harm Outcomes

Although the study did not report any of the primary fetal harm outcomes for this review, admission to NICU due to “infection risk” was reported, with little to no difference in incidence between groups (1.87% vs. 1.47%; RR 1.27, 95% CI 0.35 to 4.69). Secondary fetal harm outcomes included respiratory distress with high lactate level, requiring NICU admission (10% vs. 7%; RR 1.32, 95% CI 0.76 to 2.31) and incidence of umbilical cord artery lactate level ≥ 6.0 mmol/L (10% vs. 10%; RR 1.06, 95% CI 0.63 to 1.76) for silicone versus latex catheters.

Maternal Harm Outcomes

This trial reported two of the primary maternal harm outcomes: intrapartum antibiotics for suspected chorioamnionitis (14% vs. 10%; RR 1.35, 95% CI 0.85 to 2.16) and postpartum hemorrhage. Postpartum hemorrhage was reported by volume of blood loss, but not according to mode of delivery (vaginal versus cesarean). The differences between the catheters was small and not statistically significant for any of the three categories. For the most severe category (>1500 ml) there were 1.89 percent in the silicone and 1.12 percent in the latex group (RR 1.69, 95% CI 0.41 to 7.01).

The primary outcome specified for this RCT was accidental ROM at the time of catheter insertion. This outcome was not one prespecified for this review, but the study conducted subgroup analyses on parity. Overall, accidental ROM occurred more frequently with the silicone catheter (7% vs. 2%; RR 4.8; 95% CI 1.7 to 14.0). Subgroup analyses found increased risk in nulliparous women (RR 5.4, 95% CI 1.6 to 18.1), but not in parous women (RR 3.1, 95% CI 0.3 to 29.3), although this subgroup was small (N=163, vs. N=371 for nulliparous). No other subgroup analyses were conducted.

Key Question 4. Comparative effectiveness and harms of different methods and protocols for fetal surveillance in pregnant women undergoing cervical ripening with prostaglandins.

No studies comparing different methods and protocols for fetal surveillance during cervical ripening with prostaglandins that met inclusion criteria were identified.
## Summary Tables: Evidence for Efficacy/Effectiveness and Harms

### Table 5. Primary birth-related efficacy/effectiveness outcomes: Cesarean delivery

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Findings</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
<th>$I^2$ for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 1: Prostaglandin</strong></td>
<td>Dinoprostone outpatient vs. inpatient</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=1,120) 23% vs 23% RR 0.97 (0.75 to 1.25), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Cohort studies (n=2,511) 33% vs 33% RR 0.79 (0.67 to 0.98), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 2: Mechanical Method</strong></td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>3 RCTs (n=370) 12% vs. 20% RR 0.59 (0.21 to 1.03), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Cohort studies (n=1,057) 33% vs. 30% RR 0.95 (0.72 to 1.22), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Outpatient catheter vs. inpatient dinoprostone</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=549) 33% vs. 26% RR 1.24 (0.88 to 1.70), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 3: Outpatient</strong></td>
<td>Dinoprostone gel 2.5 mg vs. 5.0 mg</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=116) 20% vs. 19% RR 0.90 (0.44 to 1.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandin vs. placebo</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>12 RCTs (n=924) 16% vs. 21% RR 0.80 (0.58 to 1.09), $I^2=4.3%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prostaglandin vs. expectant management</td>
<td>Low strength of evidence of little or no difference</td>
<td>4 RCTs (n=615) 27% vs. 26% RR 0.95 (0.68 to 1.33), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dinoprostone vs. membrane sweeping</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>3 RCTs (n=339) 22% vs. 15% RR 1.44 (0.85 to 2.36), $I^2=0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silicone vs. latex Single-balloon catheters</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=534) 39% vs. 40% RR 0.98 (0.80 to 1.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Difference of < 5% = little or no difference; 5% to 10% = small difference; 11% to 20% = moderate difference; >20% = large difference.

b Neonatal intensive care unit (NICU) admission, not specified as the Syndrome

CI = confidence interval; RCT = randomized controlled trial; RR = relative risk
Table 6. Primary fetal harms outcomes

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Outcome</th>
<th>Findings*</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI)</th>
<th>I² for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 1:</strong> Prostaglandin</td>
<td>Dinoprostone outpatient vs. inpatient</td>
<td>Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=1,120)</td>
<td>4% vs. 3%</td>
<td>RR 1.39 (0.67 to 3.03), I²=0%</td>
<td></td>
</tr>
<tr>
<td><strong>Outpatient vs. Inpatient</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 2:</strong> Mechanical Method</td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Birth Trauma</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=129)</td>
<td>2% vs. 3%</td>
<td>RR 0.49 (0.05 to 5.30)</td>
<td></td>
</tr>
<tr>
<td>Outpatient vs. Inpatient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single-balloon catheter outpatient vs. inpatient</td>
<td>Shoulder dystocia</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>1 RCT (n=129)</td>
<td>3% vs. 11%</td>
<td>RR 0.28 (0.06 to 1.30)</td>
<td></td>
</tr>
<tr>
<td><strong>Key Question 3:</strong></td>
<td>Dinoprostone vs. placebo</td>
<td>Meconium Aspiration Syndromeb</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=134)</td>
<td>2% vs. 4%</td>
<td>RR 0.76 (0.03 to 22.33), I²=0%</td>
<td></td>
</tr>
<tr>
<td>Outpatient Comparison of Methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dinoprostone vs. placebo</td>
<td>Birth Trauma</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>3 RCTs (n=270)</td>
<td>6% vs. 0.69%</td>
<td>RR 7.88 (0.98 to 63.20)</td>
<td></td>
</tr>
</tbody>
</table>

*Difference of <1% = little or no difference; >1% to 3% = small difference; >3% to 8% = moderate difference; >8% = large difference

b Neonatal intensive care unit (NICU) admission, not specified as the Syndrome

CI = confidence interval; RCT = randomized controlled trial; RR = relative risk
### Table 7. Primary maternal harms outcomes

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Intervention</th>
<th>Outcome</th>
<th>Findings(^a)</th>
<th>Studies</th>
<th>Incidence</th>
<th>Relative Risk (95% CI) for pooled analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 2:</strong> Mechanical Method</td>
<td><strong>Single-balloon catheter outpatient vs. inpatient</strong></td>
<td>Uterine Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>2 RCTs (n=259)</td>
<td>5% vs. 5%</td>
<td>RR 0.33 (0.36 to 2.74), (I^2=0%)</td>
</tr>
<tr>
<td></td>
<td><strong>Outpatient catheter vs. inpatient dinoprostone</strong></td>
<td>Postpartum Hemorrhage</td>
<td>Low strength of evidence of a moderate, but non-significant, difference</td>
<td>2 RCTs (n=549)</td>
<td>28% vs. 22%</td>
<td>RR 1.10 (0.83 to 1.45), (I^2=0%)</td>
</tr>
<tr>
<td><strong>Key Question 3:</strong> Outpatient Comparison of Methods</td>
<td><strong>Prostaglandins vs. placebo</strong></td>
<td>Uterine Infection</td>
<td>Low strength of evidence of a small, non-significant, difference</td>
<td>7 RCTs (n=771)</td>
<td>7% vs. 5%</td>
<td>RR 0.75 (0.40 to 1.39), (I^2=0%)</td>
</tr>
<tr>
<td></td>
<td><strong>Prostaglandins vs. expected management</strong></td>
<td>Uterine Infection</td>
<td>Low strength of evidence of little or no difference</td>
<td>1 RCT (n=294)</td>
<td>6% vs. 5%</td>
<td>RR 1.21 (0.45 to 3.24)</td>
</tr>
<tr>
<td></td>
<td><strong>Prostaglandins vs. membrane sweeping</strong></td>
<td>Uterine Infection</td>
<td>Low strength of evidence of a small, but non-significant, difference</td>
<td>2 RCTs (n=269)</td>
<td>7% vs. 4%</td>
<td>RR 1.22 (0.56 to 2.75), (I^2=0%)</td>
</tr>
</tbody>
</table>

\(^a\)Difference of \(\leq1\% = \text{little or no difference}; >1\% \text{ to } 3\% = \text{small difference}; >3\% \text{ to } 8\% = \text{moderate difference}; >8\% = \text{large difference}\\CI = \text{confidence interval}; RCT = \text{randomized controlled trial}; RR = \text{relative risk}
Insufficient Evidence

For this report, there were several instances where the evidence was insufficient to draw conclusions (see Appendix H). It is important to note these instances for clarity. The Summary Tables (above) include only evidence for which there was at least low-strength evidence. Other important outcomes where evidence was insufficient to draw conclusions included outcomes related to time from admission to vaginal delivery, time in labor and delivery, fetal/neonatal harm outcomes (e.g., hypoxic-ischemic encephalopathy, meconium aspiration syndrome, intracranial/subgaleal hemorrhage) and maternal harm outcomes (e.g., hemorrhage requiring transfusion, postpartum hemorrhage by mode of delivery). For the harm outcomes, the main reason for the evidence being insufficient was inadequate sample sizes for determining rare events. This reason is combined with other issues that reduced the strength of evidence (e.g., study limitations, lack of consistency or directness). For the benefit outcomes, the main reason is that very few studies reported the outcomes prespecified for this report, such that when they were reported the evidence was indirect (i.e., using a different definition). See Appendix H for details of our assessments of the strength of evidence.
Discussion

Findings in Relation to the Decisional Dilemma(s)

The key decisional dilemma identified for this review was – when cervical ripening is indicated, what methods can be recommended as effective, but without increased risks, in the outpatient setting, and what surveillance best serves women having cervical ripening using a prostaglandin setting. More specifically, there was a need to assess the benefits of outpatient versus inpatient cervical ripening, without increasing risk (rise in cesarean delivery rate, adverse neonatal outcomes), framed within considerations of cost, patient autotomy, and satisfaction. The findings of this review can inform an update of guidance from ACOG, as the prior guidance (a 2009 Practice Bulletin on IOL) was unable to make recommendations on outpatient cervical ripening due to too few studies (one each on prostaglandins and catheters). An even more recent Cochrane review from 2017 found the evidence on outpatient versus inpatient cervical ripening to be insufficient to draw conclusions on key outcomes. While current use of outpatient cervical ripening in the United States is not well documented, controversy over its use centers around interpretation of risk. Hence, the findings of this review are useful for informing choices by clinicians and pregnant women by providing better information on the benefits (birth outcomes), and risk of harms (fetal/neonatal and maternal), and some insight into women’s preferences. We did not find evidence on FHR monitoring during cervical ripening with prostaglandins.

Across the primary outcomes prioritized for this review, there was only low-strength of evidence, with many gaps where the evidence is insufficient to draw conclusions. However, for some interventions and outcomes, the evidence is more robust than was available at the time of the prior ACOG guidance, or in the 2017 Cochrane review. The first category of outcomes to consider are the effectiveness outcomes related to birth. For these, we found low strength evidence that outpatient cervical ripening with dinoprostone or single-balloon catheters did not increase the risk of cesarean delivery relative to inpatient cervical ripening. Similarly, outpatient single-balloon catheter use did not increase the risk of cesarean delivery compared with inpatient dinoprostone. While there was not a clear difference in findings based on variables such as type of dinoprostone (gel or insert) or study quality, there are too few studies and participants to draw conclusions on the impact of these factors. Comparisons in the outpatient setting also did not indicate increased risk of cesarean delivery with dinoprostone gel at 2.5 versus 5 mg, latex versus silicone single-balloon catheter, or prostaglandins (either type) versus placebo, expectant management, or membrane sweeping. Analysis of type of prostaglandin, study quality, or of women with postterm pregnancy versus study populations that also included women with other indications for cervical ripening were possible for placebo comparisons and did not alter these findings. Most studies reported on cesarean delivery, but few reported on the time from admission to vaginal delivery. Many studies instead reported time from administration of cervical ripening method to delivery (of any mode). Additionally, much of the evidence for direct comparisons of different interventions in the outpatient setting was insufficient.

Evidence on fetal/neonatal harms was incomplete because some key outcomes were not reported or evidence was insufficient. Because fetal/neonatal harms are rare events, studies with inadequate sample sizes are unlikely to identify statistically significant differences, particularly where the differences are small. With low-strength evidence, our findings suggest no signal of differences in risk of hypoxic-ischemic encephalopathy, infection, or meconium aspiration syndrome with dinoprostone used for cervical ripening in the outpatient setting compared with
Our findings on single-balloon catheters suggest no signal for differences in risk for infection and birth trauma when used for outpatient cervical ripening versus inpatient cervical ripening. Comparing interventions in the outpatient setting, we did not find important differences in the risk of hypoxic-ischemic encephalopathy with misoprostol given intravaginally versus placebo, or meconium aspiration syndrome with dinoprostone versus placebo or membrane sweeping. The low strength of evidence for this conclusion leads to a need for future studies to confirm this finding. Outcomes where evidence was found, but was deemed insufficient, include perinatal mortality and fetal/neonatal infection with the double-balloon catheter (outpatient versus inpatient) and dinoprostone versus expectant management in the outpatient setting. The limited evidence on these outcomes did not suggest obvious increased risk for cervical ripening in the outpatient setting. More evidence is needed to draw firm conclusions. Analysis of effect modifiers such as type of prostaglandin, or study quality were not possible for these outcomes due to too few studies.

For maternal harms, again the findings are incomplete because most key outcomes were not reported or evidence was insufficient. Low-strength evidence suggests that outpatient cervical ripening with dinoprostone did not increase the risk of hemorrhage requiring transfusion compared with use in the inpatient setting. This outcome was not reported for other comparisons. There was not a clearly increased risk of postpartum hemorrhage with single-balloon catheter sin the outpatient setting compared to the inpatient setting or compared with dinoprostone in the inpatient setting. Across three small RCTs, although we found a slightly greater incidence of postpartum hemorrhage with prostaglandins than placebo in the outpatient setting, it was not statistically significant leading to a conclusion that prostaglandins are not associated with increased risk. The low strength of evidence for this conclusion leads to a need for future studies to confirm this finding. A caution in interpreting the evidence on postpartum hemorrhage is that the preferred outcome would be stratified by mode of delivery, so that hemorrhage associated with cesarean delivery could be separated from vaginal births, but the studies did not report the outcome this way. Evidence on double-balloon catheters (outpatient versus inpatient), prostaglandins versus expectant management or membrane sweeping (outpatient), or latex versus silicone single-balloon catheters was insufficient, but did not suggest obvious increased risk for cervical ripening in the outpatient setting. More evidence is also needed for this topic. The evidence on uterine infection (either chorioamnionitis or endometritis) was a bit more robust. Low strength evidence indicated little to no difference uterine infection risk with single-balloon catheters in the outpatient versus inpatient setting. In the outpatient setting, low strength evidence found small, non-significant, differences between prostaglandins and placebo, expectant management or membrane sweeping. Again, evidence for the double-balloon catheter was insufficient, but did not indicate an obvious increase in risk of uterine infection. Analysis by type of prostaglandin, study quality, or of women with postterm pregnancy versus populations of women with this and other indications for cervical ripening were possible for placebo comparisons and did not alter these findings.

Reasons for insufficient evidence were multifactorial, but were largely driven by the small numbers of studies and participants for less common outcomes, where more events are required for statistical power and precise estimates of effect. In addition, for some interventions there was only one RCT and thus, no corroborating evidence to assess consistency of the findings. However, for adverse outcomes, where there was insufficient evidence to draw firm conclusions for guidance, it is important to recognize that we also did not find early signals (i.e., much
greater incidence, although not statistically significant) of increased risk with the outpatient cervical ripening interventions.

We anticipated heterogeneity in the way that outcomes would be defined and reported, and as noted above. For the secondary outcomes, listed in the Methods section, we found that they were often not reported as specified for this review, or with a great deal of heterogeneity, making it more difficult to draw conclusions about them across studies. Failed induction (as evidence by cesarean delivery) was reported in many trials, but most did not describe or define the criteria used to diagnose “failed induction.” Across the Key Questions, differences between groups in incidence of failed induction were small in RCT evidence. One exception was an RCT of single-balloon catheters, where the incidence was lower in the outpatient group, while a cohort study found the opposite. Vaginal birth within 24 hours was not reported, while some studies reported any delivery within 24 hours. For dinoprostone, RCT evidence indicated similar incidence in outpatient and inpatient groups, and slightly higher incidence than placebo in the outpatient setting. Bishop score or cervical dilation on admission was reported in a few studies, favoring outpatient dinoprostone in one RCT, with inconsistent findings with single-balloon catheters (outpatient versus inpatient), and was similar for other comparisons. Secondary fetal/neonatal harm outcomes reported included Apgar scores (mainly <7) at 5 minutes, various outcomes related to respiratory problems that typically required support, and umbilical cord artery gas pH. None of the comparisons were significantly different for these outcomes. Unfortunately, secondary maternal harm outcomes were very infrequently reported, and not as defined in the protocol for this report.

In order to best use evidence on outpatient cervical ripening, adequate information is needed to determine if there is variation in outcomes for specific subgroups. The subgroups identified a priori for this review were parity, maternal age, Group B Streptococcus (GBS) status, diabetes (pre-gestational, gestational), and hypertension (chronic, preeclampsia without severe features, gestational). Important fetal subgroups were fetal growth restriction, and gestational age (by category). No RCT data were available to evaluate differential effectiveness or harms (i.e., effect modification, heterogeneity of treatment effect) for subgroups specified in our protocol based on outpatient versus inpatient status. To effectively evaluate this, data from well-powered high-quality (preferably RCTs) studies that report the outcomes for all treatment groups (e.g., outpatients and inpatients) for all strata of a given subgroup (e.g., for nulliparous and multiparous women) are needed. The limited information available is less than ideal, based on meta-regressions and subgroup analyses within our own meta-analyses (where we had at least 7 studies), and subgroup analyses conducted by the studies we included. These findings are hypothesis-generating and are only appropriate to guide future studies. In comparing outpatient and inpatient dinoprostone use, subgroup analyses of a cohort study suggested that postterm pregnancies had similar outcomes compared with the overall study population. Subgroup analysis of women with Bishop score >3 at enrollment in an RCT comparing outpatient double-balloon catheters with inpatient dinoprostone found a slightly higher incidence of cesarean delivery in the catheter group, though this difference was not statistically significant. In outpatient comparisons of prostaglandins and placebo for cervical ripening, meta-regression by type of prostaglandin and by gestational age (determined by enrollment of only postterm pregnancies versus mixed populations), and analysis of study quality did not find significant subgroup effects on the risk for cesarean delivery. Subgroup analyses conducted by two RCTs (Table 4) of outpatient cervical ripening found that nulliparous women had higher frequency of cesarean delivery with a prostaglandin than with placebo (2.4% and 3.2% greater), while
multiparous women had a lower frequency of cesarean delivery with a prostaglandin versus placebo (4% to 10% lower). Within the subgroups, differences in risk of cesarean delivery between prostaglandin and placebo did not reach statistical significance.

**Strengths and Limitations**

The evidence base on outpatient cervical ripening has multiple important limitations, but there are strengths to be recognized as well. Clearly, the evidence comparing interventions in the outpatient and inpatient settings suffers from too few RCTs and too small of sample sizes (range 48 to 827; mean 172), particularly when assessing harms that are rare. More and better-quality evidence is needed on specific interventions, including misoprostol and double-balloon catheters, direct comparisons of double- and single-balloon catheters, and the various formulations and routes of administration of dinoprostone or misoprostol. The studies are limited by the narrowness of the populations enrolled and inadequate reporting on or analysis of important subgroups such as women over 30 or 35, the effect of GBS status, diabetes, hypertension, fetal growth restriction, and gestational age categories. The studies generally either excluded women with such characteristics, or failed to report on them in detail. RCTs are needed to help address issues of imbalance in characteristics such as these in observational studies. There was variation in how outcomes were defined and reported across the studies. Few studies reported the outcomes as they were specified for this review, based on input from experts. For example, the primary birth outcome of time from admission to vaginal birth was rarely reported as such. More often, studies reported on time from admission to any delivery mode, time from placement of the intervention to delivery (any mode), or placement to the onset of active labor. Input from experts indicates that, in addition to establishing similar efficacy and risk of harm, when comparing outpatient and inpatient cervical ripening, there is interest in reducing the amount of time in the labor and delivery ward. Hence, these other outcomes are not as relevant, but are frequently reported. There were also numerous primary harm outcomes that failed to meet the criteria set by the experts. These outcomes include meconium aspiration (rather than the syndrome), neonatal encephalopathy (not specifically hypoxic-ischemic), infection (typically suspected, rather than confirmed), postpartum hemorrhage (not according to mode of delivery), and uterine infection (again, typically not reported as confirmed infections). Secondary outcomes also suffered from similar lack of specificity, with most not consistent with the definitions for this review and for the comparison of interventions in the outpatient setting versus each other; we were looking for longer-term outcomes related to breastfeeding, maternal mood, and mother-baby attachment, but no study reported these. In addition to these specifics, the studies were limited by inadequate reporting on many characteristics of the populations, interventions, comparators, and outcomes that limited our ability to analyze their impact. For example, women’s cervical dilation when synthetic oxytocin was initiated and the dose/timing of augmentation was not consistently noted in the included studies.

Subgroups analyses were infrequent, and few studies were conducted to directly examine the key subgroups of interest. Examples include race, provider type, parity, and maternal BMI. A major limitation was the lack of studies evaluating fetal monitoring in the context of cervical ripening using a prostaglandin (in any setting).

Strengths of the evidence base included that the available evidence can shed light on the use of outpatient cervical ripening for younger women with singleton pregnancies, primarily nulliparous women either to prevent or for postterm pregnancy. The evidence was adequate to
address the impact of dinoprostone and single-balloon catheters on cesarean delivery, and some fetal/neonatal harms in this population.

Limitations of the review process included our intention to undertake a “best evidence” approach. Initially, we proposed including cohort and case-control studies if RCT evidence was inadequate. Based on this, we included nine cohort studies, most of which were small (range 42 to 1343, mean 567). However, even after including these, we recognized that the evidence base was likely to provide inadequate information to guide recommendations or clinical practice, in particular related to the risk of important harms, and for fetal surveillance studies. After discussion with the AHRQ Task Order Officer (TOO) and the review sponsor, we undertook an additional search for studies with much broader design criteria – single-arm studies of outpatient cervical ripening that reported on any of the harms prioritized for this review. However, this additional search did not identify any relevant studies. While the lack of identifying additional studies may be a limitation, we feel that the approach was a strength. Other limitations of our review process were that we excluded studies published in languages other than English, and were unable to conduct small sample size bias assessments for most outcomes due to small numbers of studies. We consider our approach to meta-analysis, using the profile-likelihood random effects models, to be a strength of our approach because we assumed that there was heterogeneity across the studies and knew we would have small numbers of studies for most outcomes.

Applicability

A number of factors could impact the applicability of our findings. Studies generally included women with singleton pregnancies with vertex presentations, without a history of a prior cesarean delivery or other uterine surgeries, PROM/ROM, comorbidities (e.g., diabetes, hypertension), uterine growth restriction, or other fetal problems. The exception is the evidence comparing prostaglandins in the outpatient setting, where 35 percent of women had a prior cesarean delivery. While studies did not often limit enrollment by parity, most women were nulliparous. The mean age of enrolled women was 26 to 28 years in the prostaglandin studies, and slightly older in studies of catheters (30 years). As such, the studies were not able to evaluate effects in women with older maternal age (e.g., >30 or 35 years). Gestational age was typically 40.6 weeks. The definition of “postterm” varied or was often not defined. However, using ACOG’s 2014 definitions, 40.6 weeks is “late-term”, rather than “post-term” (42 weeks and beyond). Mean Bishop score at enrollment was 2.9 to 4 (weighted mean of 3.6). The evidence is most applicable to women with a postterm pregnancy, or to prevent a postterm pregnancy. The findings were less applicable to women with comorbidities, such as diabetes or hypertension, including those that developed during pregnancy. Some studies did not restrict enrollment of women with these comorbidities, but the small percentages of such women ultimately enrolled did not allow for subgroup analyses to inform on the applicability of the findings to these subgroups. There was too little, or no, information on other important characteristics of pregnant women that were identified a priori, including race, ethnicity, body mass index (BMI), maternal pre-pregnancy health status (including mental health), and intangibles such as birth plan/philosophy and type of provider to understand implications for outcomes or applicability of the findings. Applicability to low-income patients, or those with low access to healthcare is unclear. Although a small number of studies required that women have had prenatal care, or that they live within 30 minutes of the hospital, and most of the studies did not have such restrictions, other factors, such as comorbidities (diabetes, hypertension) need to be considered. These
problems with applicability did not vary by the type of cervical ripening method. None of the studies reported other information relevant for assessing applicability, such as the description of the source of potential study participants and the number of women randomized relative to the number of women enrolled.

Intervention-related factors that may limit applicability include dose, route of administration, and re-administration schedule variation with medications, and balloon-fill volume variation with catheters. Across the studies, we found little variation in prostaglandin dose, within the specific drug and formulations or routes of administration. For example, the evidence for misoprostol applies to a 25 mcg intravaginal dose, or a 100 mcg oral dose, and dinoprostone doses and application were consistent with the specific product (i.e., the gel or the insert). All prostaglandin studies reported FHR monitoring following administration, but the duration ranged from 40 minutes to 2 hours. Reapplication criteria and schedules for prostaglandins were most often not well reported, and while a small number of studies limited to a single dose, most were either silent on reapplication or did not report the actual number of doses women received. Additionally, we noted that the applicability of the findings to misoprostol (PGE1) is limited, particularly for the inpatient versus outpatient comparison. Single- and double-balloon catheter fill volumes were consistent across the studies, according to the specific catheter, but monitoring procedures prior to discharge were inconsistently reported.

Outcomes related to time from admission to delivery often included cesarean as well as vaginal delivery. Inclusion of cesarean delivery reduces the applicability of the findings from individual studies, due to variation in clinical practice, policy, and preferences across provider, patient, health system or country, and across time (as temporal trends for cesarean delivery have changed).

Information relevant for assessing applicability of the care setting such as details on the type of outpatient setting (e.g., home, home birthing center) or inpatient setting (e.g., hospital, clinic) was poorly reported. Other limitations to our ability to assess applicability include the lack of information on variation in provider type (e.g., midwife, nurse, or generalist obstetrician), rural versus metropolitan, planned home birth versus planned inpatient birth, and country. While many studies reported that the interventions were applied by a midwife or a physician, none limited to a specific provider type or stratified results by this variable. Although 60 percent of studies were conducted in the United States, we were unable to assess the impact of country of study or other geographic location characteristics (e.g., rural, metropolitan) on the applicability of specific results.

A number of evidence gaps or limitations in evidence potentially impacted the applicability of our findings. Lack of evidence on misoprostol, double-balloon catheters, hygroscopic dilators, direct comparison of interventions, type of provider applying the intervention, reapplication criteria and schedules, and protocols for clinician contact/monitoring in outpatient cervical ripening may limit applicability to common OB practice. Inadequate reporting of maternal and fetal factors such as parity, maternal age, GBS status, diabetes, hypertension, fetal growth restriction, and gestational age made it difficult to access applicability of our findings to subgroups of patients with these factors.
Implications for Clinical Practice, Education, Research, or Health Policy

Considerations for Clinical Practice

The implications of the findings of this report include the ability to provide more guidance to clinicians and pregnant women on the relative benefits and harms of outpatient cervical ripening. This report finds low strength of evidence that outpatient cervical ripening with dinoprostone and single-balloon catheters does not impose increased risk of cesarean delivery and does not indicate that there are important signals of increased risk of fetal/neonatal and maternal harms, with the limitation that not all such harms were adequately studied or reported. The findings apply most directly to women under age 30, with singleton fetuses with vertex presentation, and no major comorbidities. The question of the characteristics of pregnant women and fetuses that will benefit most or have the lowest risk of harm is not addressed by this evidence. Similarly, there is less information to guide the use of double-balloon catheters, hygroscopic dilators, misoprostol, or to compare doses and routes of administration of prostaglandins.

Women’s preferences for the setting for cervical ripening (inpatient versus outpatient), and satisfaction with outpatient cervical ripening are important factors in applying the results of this systematic review. Our assessment of studies pertaining to these issues suggest a preference for outpatient cervical ripening and that women are willing to make tradeoffs based on their personal circumstances, however the decision-making process is complex. Commonly weighed factors included support at home, proximity to the hospital and perceived safety. Overall, women who had cervical ripening in the outpatient setting seemed to be more satisfied with their experience compared with women who had cervical ripening in an inpatient setting. These findings are based on qualitative studies of women undergoing cervical ripening in the US and Australia, or surveys of women participating in trials of outpatient cervical ripening in Canada and Australia. These studies included women with low-risk pregnancies. Most were being induced for postterm pregnancies and over half were nulliparous. These were identified as a “contextual question” for this report, and more detailed information on these studies is available upon request.

Research Recommendations

- Additional RCTs to corroborate findings for all comparisons. Additional RCTs are needed on outpatient misoprostol, and double-balloon catheters. Preference is for more studies in the US.
- Larger RCTs, with sample sizes large enough to evaluate important harm outcomes. Sample size calculations should be based on discussion with clinical experts on minimally important differences in risk for key rare harm outcomes.
- Quality improvement studies and registry studies, employing the power of electronic medical records, may be utilized to evaluate rare harms and explore process questions.
- RCTs with sufficient sample size to evaluate differential effectiveness and harms of outpatient cervical ripening in important subgroups of the population: parity, maternal age, Group B Streptococcus (GBS) status, diabetes (pre-gestational, gestational), hypertension (chronic, preeclampsia without severe features, gestational), fetal growth restriction, gestational age (<39 weeks, 39 to 41 weeks, >41 weeks).
Future studies should evaluate the effects of important subgroups (e.g. parity, race, maternal BMI) and additional factors not considered here (e.g. augmentation of labor with synthetic oxytocin, epidural anesthesia).

While we acknowledge that there can be challenges in enrolling pregnant women in RCTs, we do not recommend additional retrospective cohort studies. Large, prospective, cohort studies that use strong methods to control for variation in baseline risk of women studied (e.g., propensity score matching) may be helpful. Large, well-conducted case-control studies may be useful to evaluate the risk of rare harms.

Evidence comparing methods of fetal surveillance (e.g., intermittent heart rate auscultation versus electronic monitoring) during cervical ripening with prostaglandins is needed, as we found no studies of different approaches that might be useful in outpatient cervical ripening.

Conclusions

In women with low-risk pregnancies, incidence of cesarean delivery, fetal/neonatal, or maternal harms were not significantly different with dinoprostone or single-balloon catheter for cervical ripening in the outpatient versus inpatient setting, and compared with placebo, expectant management, or membrane sweeping in the outpatient setting. Evidence for these findings was low-strength; evidence for other comparisons or outcomes was insufficient.
References


### Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOG</td>
<td>American College of Obstetricians and Gynecologists</td>
</tr>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CQ</td>
<td>Contextual Question</td>
</tr>
<tr>
<td>DCE</td>
<td>discrete choice experiment</td>
</tr>
<tr>
<td>EHC</td>
<td>Effective Health Care</td>
</tr>
<tr>
<td>FHR</td>
<td>fetal heart rate</td>
</tr>
<tr>
<td>GA</td>
<td>gestational age</td>
</tr>
<tr>
<td>GBS</td>
<td>Group B Streptococcus</td>
</tr>
<tr>
<td>GDM</td>
<td>gestational diabetes</td>
</tr>
<tr>
<td>IOL</td>
<td>induction of labor</td>
</tr>
<tr>
<td>ITT</td>
<td>intent-to-treat</td>
</tr>
<tr>
<td>KQ</td>
<td>Key Question</td>
</tr>
<tr>
<td>L&amp;D</td>
<td>labor and delivery</td>
</tr>
<tr>
<td>MD</td>
<td>mean difference</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NICU</td>
<td>Neonatal Intensive Care Unit</td>
</tr>
<tr>
<td>NR</td>
<td>not reported</td>
</tr>
<tr>
<td>NS</td>
<td>not significant</td>
</tr>
<tr>
<td>OHSU</td>
<td>Oregon Health &amp; Science University</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>PCORI</td>
<td>Patient-Centered Outcomes Research Institute</td>
</tr>
<tr>
<td>PICOTs</td>
<td>Population, Intervention, Comparator, Outcome, Setting, Setting/Study Design</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic Reviews</td>
</tr>
<tr>
<td>PROM</td>
<td>premature rupture of membrane</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>ROM</td>
<td>rupture of membrane</td>
</tr>
<tr>
<td>RR</td>
<td>relative risk</td>
</tr>
<tr>
<td>SEADS</td>
<td>Supplemental Evidence and Data for Systematic review</td>
</tr>
<tr>
<td>SOE</td>
<td>strength of evidence</td>
</tr>
<tr>
<td>TEP</td>
<td>Technical Expert Panel</td>
</tr>
</tbody>
</table>