Evidence Table 1: Studies relevant to Key Question 1

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Alfirevic, Luckas, Walkin-shaw, et al., 1997	Design: RCT, randomization by sealed envelope Test(s) studied: 1) U/S measurement of amniotic fluid index (AFI) + computerized cardiotocography (CTG) using Oxford Sonicaid 8000 fetal monitor (n = 250) Protocol: If AFI < 7.3 cm (< 3rd percentile for 42-wk gestation) or if CTG abnormal (according to proprietary criteria), then labor induced. If AFI and CTG normal, then f/u visit arranged 3 days later, unless patient had reached 43 wks gestation (301 days), in which case labor induced regardless of test results. Labor induced with intravaginal prostaglandins (details NR). 2) U/S measurement of maximum pool depth (MPD) + computerized cardiotocography (CTG) using Oxford Sonicaid 8000 fetal monitor (n = 250) Protocol: If MPD < 1.8 cm (< 3rd percentile for 42-wk gestation) or if CTG abnormal (according to proprietary criteria), then labor induced. If MPD and CTG normal, then f/u visit arranged 3 days later, unless patient had reached 43 wks gestation (301 days), in which case labor induced regardless of test results.	No. of subjects at start: 500 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 500 Inclusion criteria: Uncomplicated singleton pregnancy; ≥ 40 wks gestation Exclusion criteria: Hypertension (≥ 140/95 mmHg); significant proteinuria (> 1+ on dipstick); history of antepartum hemorrhage; poor obstetric history; prior U/S suggesting IUGR Age (median, with interquartile [IQ] range): AFI + CTG: 28 (24-31); MPD + CTG: 28 (23-32) Race: NR Gestational age at entry (median, with IQ range): AFI + CTG: 290 days (289-291); MPD + CTG: 290 days (289-291) Dating criteria: 1) Certain LMP + U/S prior to 20 wks or 2) agree-ment within 1 wk between certain LMP and U/S after 20 wks Parity: AFI + CTG: 50% nulliparous; MPD + CTG: 50% nulliparous Bishop score: NR	1) Birthweight 2) Cord pH at delivery 3) Apgar < 7 at 5 minutes 4) Admission to NICU 5) Perinatal death 6) Cord base excess 7) Meconium 8) C-sections 9) Inductions	1) Birthweight (median, with IQ range): AFI + CTG: 3740 g (3417.5 to 3985) MPD + CTG: 3710 g (3390 to 4027.5) p = 0.89 2) Cord pH at delivery (median, with IQ range): AFI + CTG: 7.29 (7.25 to 7.34) MPD + CTG: 7.3 (7.25 to 7.34) p = 0.57 3) Apgar < 7 at 5 minutes: AFI + CTG: 5/250 (2%) MPD + CTG: 5/250 (2%) p = 1 4) Admission to NICU: AFI + CTG: 4/250 (1.6%) MPD + CTG: 4/250 (1.6%) p = 1 5) Perinatal death: AFI + CTG: 0/250 MPD + CTG: 0/250 p = 1 6) Cord base excess (median, with IQ range): AFI + CTG: -5.2 (-3.45 to -7.1) MPD + CTG: -5.4 (-3.9 to -7.2) p = 0.18 7) Meconium: AFI + CTG: 56/250 (22%) MPD + CTG: 56/250 (22%) p = 1 8) C-sections: *Overall:* AFI + CTG: 47/250 (19%) MPD + CTG: 33/250 (13%) p = 0.11	QUALITY SCORES: TESTING Reference standard: - Randomized: + Method of randomization: + Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: + MANAGEMENT Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Sample size estimates based on difference in C-section rates – power to detect differences in perinatal outcomes questionable.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Labor induced with intravaginal prostaglandins (details NR). Reference standard(s): None Dates: July 1994-July 1995 Location: Liverpool, UK Setting: University hospital Type(s) of providers: General OB/GYN, MFM, midwives (nonnurse) Length of follow-up: None			*For fetal distress:* AFI + CTG: 20/250 (8%) MPD + CTG: 10/250 (4%) p = 0.09 *For failure to progress:* AFI + CTG: 25/250 (10%) MPD + CTG: 21/250 (8%) p = 0.64 *For other indications:* AFI + CTG: 2/250 (0.8%) MPD + CTG: 2/250 (0.8%) p = 1 9) Inductions: *Overall:* AFI + CTG: 87/250 (35%) MPD + CTG: 77/250 (31%) p = 0.39 *For abnormal post-term monitoring:* AFI + CTG: 37/250 (15%) MPD + CTG: 21/250 (8%) p = 0.04 *Maternal request:* AFI + CTG: 24/250 (10%) MPD + CTG: 25/250 (10%) p = 1 *43 weeks' gestation:* AFI + CTG: 17/250 (7%) MPD + CTG: 21/250 (8%) p = 0.61 *For other indications:* AFI + CTG: 9/250 (4%) MPD + CTG: 10/250 (4%) p = 1	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Alfirevic and Walkinshaw, 1995	Design: RCT, randomization by sealed envelope Test(s) studied: 1) Simple monitoring = cardiotocography (CTG) + U/S measurement of maximum pool depth (MPD) (n = 73) Protocol: If CTG abnormal (< 2 accelerations [15 bpm lasting ≥ 15 sec] in 40 min or short-term variability ≤ 5 bpm with no decelerations) or MPD abnormal (< 2.1 cm), then labor induced. If both tests normal, then f/u visit arranged 3 days later, unless patient had reached 43 wks gestation, in which case labor induced regardless of test results. Labor induced with intravaginal prostaglandins (details NR). 2) Complex monitoring = modified biophysical profile (MBP) = computerized cardiotocography (using the Oxford Sonicaid 8000 fetal monitor) + U/S measurement of amniotic fluid index (AFI) + fetal breathing movements + fetal tone + fetal gross body measurements (last 3 all monitored by U/S) (n = 72) Protocol: If AFI < 7.3 cm (< 3rd percentile for 42 wks gestation), then labor induced. If MBP total score ≤ 6 of possible 10 (each component score 0 to 2, with 2 = normal), then labor induced. If AFI	No. of subjects at start: 145 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 145 Inclusion criteria: Uncomplicated singleton pregnancy; ≥ 41 wks gestation Exclusion criteria: Hypertension (≥ 140/95 mmHg); significant proteinuria (> 1+ on dipstick); history of antepartum hemorrhage; poor obstetric history; prior U/S suggesting IUGR Age (median, with interquartile [IQ] range): Simple, 28 (25-32); complex, 29 (25-31) Race: NR Gestational age at entry: NR; gestational age ≥ 41 weeks required for entry into study Dating criteria: Certain LMP or U/S prior to 20 weeks Parity: Simple, 33% nulliparous; complex: 40% nulliparous Bishop score: NR	1) Perinatal death 2) Admission to NICU 3) Apgar score < 7 at 5 minutes 4) Cord pH at delivery 5) Meconium 6) C-sections 7) Spontaneous labor 8) Inductions 9) Normal vaginal delivery 10) Abnormal CTG intrapartum	1) Perinatal death: Simple: 0/73 Complex: 1/72 (1%) (no p-value reported) 2) Admission to NICU: Simple: 2/73 (3%) Complex: 0/72 (no p-value reported) 3) Apgar score < 7 at 5 minutes: Simple: 0/73 Complex: 1/72 (1%) (no p-value reported) 4) Cord pH at delivery (median, with IQ range): Simple: 7.31 (7.26 to 7.35) Complex: 7.29 (7.25 to 7.33) p = 0.15 5) Meconium: Simple: 14/73 (19%) Complex: 20/72 (28%) p = 0.30 6) C-sections: *Overall:* Simple: 7/73 (10%) Complex: 13/72 (18%) p = 0.22 *For fetal distress:* Simple: 6/73 (8%) Complex: 8/72 (11%) p = 0.54 *For antepartum distress:* Simple: 2/73 (3%) Complex: 0/72 (no p-value reported)	QUALITY SCORES: TESTING Reference standard: - Randomized: + Method of randomization: + Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: + MANAGEMENT Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - No assessment of cervical ripeness – may explain high rate of meconium and C-section among those women with labor induced for abnormal MPD. Sample size estimates based on differences in cord pH. *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	normal and MBP normal, then f/u visit arranged 3 days later, unless patient had reached 43 wks gestation, in which case labor induced regardless of test results. Labor induced with intravaginal prostaglandins (details NR). Reference standard(s): None Dates: Jan-Dec 1973 Location: Liverpool, UK Setting: University hospital Type(s) of providers: Unspecified OB/GYN, MFM Length of follow-up: None			7) Spontaneous labor: Simple: 41/73 (56%) Complex: 29/72 (40%) p = 0.08 8) Inductions: *Overall:* Simple: 30/73 (41%) Complex: 43/72 (60%) p = 0.04 *For abnormal post-term monitoring:* Simple: 11/73 (15%) Complex: 28/72 (39%) p = 0.002 *43 weeks' gestation:* Simple: 12/73 (16%) Complex: 9/72 (13%) p = 0.66 *Maternal request:* Simple: 4/73 (5%) Complex: 2/72 (3%) p = 0.69 *Other:* Simple: 3/73 (4%) Complex: 4/72 (6%) p = 0.9 9) Normal vaginal delivery: Simple: 58/73 (79%) Complex: 50/72 (69%) p = 0.23 10) Abnormal CTG intrapartum: Simple: 29/73 (40%) Complex: 34/72 (47%) p = 0.36	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Arabin, Snyjders, Mohnhaupt, et al., 1993	Design: Case series, no controls Test(s) studied: Note: Tests 1) *and* 2) applied to all patients in the series (n = 110) 1) Traditional biophysical profile 2) Fetal assessment score consisting of 5 components: FHR pattern; uterine artery resistance by Doppler U/S; carotid artery resistance index by Doppler U/S; fetal tone (movements) by U/S; fetal reflexes (magnitude and speed of movements) by U/S Reference standard(s): Fetal distress (pathological FHR pattern resulting in operative delivery, Apgar score < 7 at 1 minute, or cord blood pH < 7.20) Dates: NR Location: Berlin, Germany Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 110 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 110 Inclusion criteria: Gestational age > 290 days; singleton pregnancy Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR; gestational age > 290 days required for entry into study; mean gestational age at delivery 295 days (range, 293-300) (all patients delivered within ≤ 3 days of assessment) Dating criteria: LMP confirmed by "early" U/S Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Cord pH < 7.20 4) C-sections due to fetal distress 5) Test performance	1) Apgar score < 7 at 1 minute: 10/110 (9%) 2) Apgar score < 7 at 5 minutes: 2/110 (2%) 3) Cord pH < 7.20: 9/110 (8%) 4) C-sections due to fetal distress: 38/110 (34.5%) 5) Test performance: Fetal assessment score provided better prediction of fetal distress and low Apgar score at 1 minute than did biophysical profile in ROC analysis ($p < 0.001$). No difference between the two tests for prediction of low pH. Stepwise discriminant analysis of individual components of biophysical profile showed that only FHR pattern and AFV contributed significantly to the diagnostic properties of the total score. Similar analysis of the new fetal assessment score showed that all components except fetal tone contributed significantly to the diagnostic properties of the total score.	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: + Fetal assessment score most superior to biophysical profile score in discriminating the relatively subjective outcome of "fetal distress."

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Arias, 1987	Design: RCT, randomization by last digit of year of birth Test(s) studied: 1) Nonstress test (NST) (n = 126) Protocol: Patients evaluated with weekly NST. NST considered reactive if 5 or more accelerations of ≥ 15 bpm lasting at least 15 sec, in association with fetal movements, in 20 minutes. If NST nonreactive, then oxytocin challenge test (OCT) performed. If OCT positive or suspicious, then labor induced. Method of induction not described. 2) U/S + NST (n = 117) Protocol: Weekly U/S evaluation, with assessment of fetal weight, AFV, and placenta. If placenta was grade III and there was decreased AFV, or if fetal weight ≥ 4000 g, then labor induced. Weekly NST as above, with same criteria for induction. Method of induction not described. Reference standard(s): Occurrence of abnormal outcomes (except those not predictable by NST) Dates: NR (15 months' duration) Location: St. Louis, MO Setting: Community hospital	No. of subjects at start: 287 Dropouts: 44 Loss to follow-up: NA No. of subjects at end: 243 Inclusion criteria: Excellent dates (based on LMP or U/S); > 40 wks gestation Exclusion criteria: Diabetes; hypertension; any medical complication of pregnancy Age (mean ± SD): NST: 25.6 ± 4.9; U/S + NST: 25.9 ± 4.9 Race: NR Gestational age at entry (mean ± SD): NST: 41.2 ± 0.7 weeks; U/S + NST: 41.2 ± 0.6 weeks Dating criteria: LMP or U/S during first 26 weeks Parity (mean ± SD): NST: 1.8 ± 1.1; U/S + NST: 1.8 ± 1.2 Bishop score: NR	1) Mean birthweight 2) Birthweight > 4000 g 3) Birthweight > 4500 g 4) Any complication 5) Shoulder dystocia 6) Meconium aspiration 7) Post-maturity syndrome 8) C-sections 9) C-sections due to fetal distress 10) 2 x 2 tables	P-values not reported for the outcomes listed here. 1) Mean birthweight (± SD): NST: 3742 ± 472 g U/S + NST: 3813 ± 482 g 2) Birthweight > 4000 g: NST: 45/126 (36%) U/S + NST: 27/117 (23%) 3) Birthweight > 4500 g: NST: 10/126 (8%) U/S + NST: 9/117 (8%) 4) Any complication: NST: 32/126 (25%) U/S + NST: 29/117 (25%) 5) Shoulder dystocia: NST: 6/126 (5%) U/S + NST: 2/117 (2%) 6) Meconium aspiration: NST: 5/126 (4%) U/S + NST: 3/117 (3%) 7) Post-maturity syndrome: NST: 5/126 (4%) U/S + NST: 4/117 (3%) 8) C-sections: NST: 32/126 (25%) U/S + NST: 33/117 (28%) 9) C-sections due to fetal distress: NST: 12/126 (9.5%) U/S + NST: 16/117 (14%)	QUALITY SCORES: TESTING Reference standard: + Randomized: + Method of randomization: - Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - MANAGEMENT Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Unspecified OB/GYN Length of follow-up: None			10) 2 x 2 tables: 2 x 2 Table 1: Reference standard = abnormal outcomes Screening test = NST Abnormal outcomes yes / no / Totals: NST + 6 / 8 / 14 NST - 12 / 86 / 98 Totals: 18 / 94 / 112 2 x 2 Table 2: Reference standard = abnormal outcomes Screening test = U/S + NST Abnormal outcomes yes / no / Totals: NST + 15 / 15 / 30 NST - 26 / 49 / 75 Totals: 41 / 64 / 105	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Battaglia, Larocca, Lanzani, et al., 1991	Design: Case series (prospective), no controls Test(s) studied: 1) Nonstress test (NST) + amnioscopy + amniotic fluid volume (AFV) + Doppler velocimetry of the uterine, umbilical, descending thoracic aorta, renal, and middle cerebral arteries + hPL + estriol + hematocrit + platelets + mean platelet volume + uric acid Protocol: NST, amnioscopy, AFV, and Doppler velocimetry performed every other day; remaining tests performed every 3 days. Time-averaged mean velocity in the descending thoracic aorta calculated using mean value of three consecutive waveforms. Reference standard(s): 1) "Poor condition" 2) Oligohydramnios 3) Meconium staining 4) NST 5) C-sections (overall) 6) C-sections for fetal distress Dates: Jan - Dec 1989 Location: Modena, Italy Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 82 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 82 Inclusion criteria: Gestational age ≥ 287 days; singleton fetus; cephalic presentation Exclusion criteria: Medical or obstetric complications Age (mean, with range): 27.9 (19-39) Race: NR Gestational age at entry (mean): 292.4 days Dating criteria: LMP + U/S before 24 weeks Parity: 0: 58/82 (71%) 1: 18/82 (22%) > 1: 6/82 (7%) Bishop score: NR	1) Birthweight 2) Macrosomia (birthweight > 4000 g) 3) "Poor condition" (both 1- and 5-minute Apgar scores < 7 *or* infant admitted to NICU for asphyxia and/or meconium aspiration syndrome) 4) Oligohydramnios (largest pocket < 2 cm) 5) Meconium staining 6) C-sections 7) 2 x 2 tables	1) Birthweight (mean): 3655.5 g 2) Macrosomia: 18/82 (22%) 3) "Poor condition": 1/82 (1%) 4) Oligohydramnios: 25/82 (30%) 5) Meconium staining: 24/82 (29%) 6) C-sections: 24/82 (29%) 7) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = "Poor condition" (as defined at left) Screening test = Time-averaged mean velocity of the descending thoracic aorta ("normal" defined as > 25 cm/sec) |	Poor condition yes	no	Totals:
---	---	---	---
Velocity abnormal	1	23	24
Velocity normal	0	58	58
Totals:	1	81	82 | Sensitivity: 100% Specificity: 71% <u>2 x 2 table 2:</u> Reference standard = Oligohydramnios Screening test = Time-averaged mean velocity of the descending thoracic aorta ("normal" defined as > 25 cm/sec) |	Oligohydramnios yes	no	Totals:
---	---	---	---
Velocity abnormal	16	8	24
Velocity normal	9	49	58
Totals:	25	57	82	 *(continued on next page)*	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: -

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Sensitivity: 64% Specificity: 86% <u>2 x 2 table 3:</u> Reference standard = Meconium staining Screening test = Time-averaged mean velocity of the descending thoracic aorta (“normal” defined as > 25 cm/sec) Meconium yes / no / Totals: Velocity abnormal: 22 / 2 / 24 Velocity normal: 2 / 56 / 58 Totals: 24 / 58 / 82 Sensitivity: 92% Specificity: 97% <u>2 x 2 table 4:</u> Reference standard = NST Screening test = Time-averaged mean velocity of the descending thoracic aorta (“normal” defined as > 25 cm/sec) NST abn / nl / Totals: Velocity abnormal: 13 / 11 / 24 Velocity normal: 0 / 58 / 58 Totals: 13 / 69 / 82 Sensitivity: 100% Specificity: 84% <u>2 x 2 table 5:</u> Reference standard = C-sections (overall) Screening test = Time-averaged mean velocity of the descending thoracic aorta (“normal” defined as > 25 cm/sec)	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				C-section yes / no / Totals: Velocity abnormal: 14 / 10 / 24 Velocity normal: 10 / 48 / 58 Totals: 24 / 58 / 82 Sensitivity: 58% Specificity: 50% 2 x 2 table 6: Reference standard = C-section for fetal distress Screening test = Time-averaged mean velocity of the descending thoracic aorta (“normal” defined as > 25 cm/sec) C-section/fetal distress yes / no / Totals: Velocity abnormal: 8 / 16 / 24 Velocity normal: 2 / 56 / 58 Totals: 10 / 72 / 82 Sensitivity: 80% Specificity: 78%	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Bochner, Medearis, Ross, et al., 1987	Design: Cohort study Test(s) studied: 1) Antepartum testing, including amniotic fluid assessment, NST, and, when necessary, contraction stress testing (CST). Uterine contractions, FHR, and fetal movements also assessed. Protocol: Testing performed twice weekly. Abnormal testing, leading to induction, included decreased amniotic fluid; repetitive variable or late decelerations during the NST or CST; and a nonreactive NST in a patient with an inducible cervix. Patients with a nonreactive NST and an unfavorable cervix had a repeat NST 2 hours later. CST done if the NST was again nonreactive. If the CST negative, then patients re-tested in 3-4 days. Reference standard(s): 1) Meconium aspiration 2) Low birthweight (< 10th percentile) 3) Perinatal mortality or morbidity 4) C-section for fetal distress 5) Apgar < 7 at 1 minute 6) Apgar < 7 at 5 minutes Dates: Jan 1983 - Jan 1986 Location: Los Angeles, CA Setting: University hospital Type(s) of providers:	No. of subjects at start: 845 Dropouts: 6 Loss to follow-up: NA No. of subjects at end: 839 Inclusion criteria: Gestational age of 41-42 completed weeks; referred for post-term fetal assessment Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR (gestational age of 41-42 completed weeks required for entry into study) Dating criteria: Combinations of early dating criteria, including LMP, initial uterine exam, 1st or 2nd trimester U/S, and timing of initial fetal heart tones by Doppler or fetoscopic auscultation Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Meconium aspiration 4) Mortality 5) Low birthweight (< 10th percentile) 6) C-section for fetal distress 7) 2 x 2 tables	1) Apgar score < 7 at 1 minute: 56/83 (6.7%) 2) Apgar score < 7 at 5 minutes: 13/839 (1.5%) 3) Meconium aspiration: 3/839 (0.4%) 4) Mortality: 0/839 5) Low birthweight (< 10th percentile): 7/839 (0.8%) 6) C-section for fetal distress: 52/839 (6.2%) 7) 2 x 2 tables (for patients with heavy meconium at rupture of the membranes only [n = 62]): <u>2 x 2 table 1:</u> Reference standard = Meconium aspiration Screening test = Antepartum testing |	Meconium aspiration yes	no	Totals:
---	---	---	---
Antepartum testing abn	1	13	14
Antepartum testing nl	2	46	48
Totals:	3	59	62 | <u>2 x 2 table 2:</u> Reference standard = Low birthweight (defined as < 10th percentile) Screening test = Antepartum testing |	Low birthweight yes	no	Totals:
---	---	---	---
Antepartum testing abn	2	12	14
Antepartum testing nl	5	43	48	 *(continued on next page)*	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: -

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Unspecified OB/GYN Length of follow-up: None			Totals: 7 55 62 2 x 2 table 3: Reference standard = Perinatal mortality or morbidity Screening test = Antepartum testing Mortality/ morbidity yes no Totals: Antepartum testing abn 0 14 14 Antepartum testing nl 0 48 48 Totals: 0 62 62 2 x 2 table 4: Reference standard = C-section for fetal distress Screening test = Antepartum testing C-section yes no Totals: Antepartum testing abn 11 3 14 Antepartum testing nl 2 46 48 Totals: 13 49 62 2 x 2 table 5: Reference standard = Apgar score at 1 minute Screening test = Antepartum testing Apgar at 1 min < 7 ≥ 7 Totals: Antepartum testing abn 6 8 14 Antepartum testing nl 18 30 48 Totals: 24 38 62	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 6: Reference standard = Apgar score at 5 minutes Screening test = Antepartum testing	

	Apgar at 5 min		
	< 7	≥ 7	Totals:
Antepartum testing abn	1	13	14
Antepartum testing nl	0	48	48
Totals:	1	61	62

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Bochner, Williams III, Castro, et al., 1988	Design: Case series, concomitant controls Test(s) studied: 1) Antenatal testing beginning at 41 (n = 908) or 42 (n = 352) weeks Protocol: Testing performed twice weekly. Standard fetal monitor recorded uterine contractions, fetal heart rate, and fetal movements. U/S evaluated AFV (< 3 cm abnormal). Nonstress test (NST) also performed. If NST nonreactive *and* AFV normal *and* cervix unfavorable for induction, then NST repeated in 2 hours; if second NST nonreactive, then contraction stress test (CST) performed. If CST negative, then patient re-tested in 3-4 days. Criteria for induction: Decreased AFV (< 3 cm); *or* bradycardia or repetitive variable or late decelerations during NST or CST; *or* nonreactive NST and inducible cervix. Method of induction not described. 2) No antenatal testing (n = 1807 controls). Management protocol not described. Reference standard(s): Intra-partum fetal distress, defined as: a) repetitive late decelerations; b) repetitive moderate or severe variable decelerations with pH < 7.2 or decreased variability; or c)	No. of subjects at start: 1260 subjects, 1807 controls Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 1260 subjects, 1807 controls Inclusion criteria: Uncomplicated post-term pregnancy (> 41 wks); first seen before 20 wks; trial of labor; delivery within 4 days of antepartum testing Exclusion criteria: High risk factors; suspected fetal growth retardation Age: NR Race: NR Gestational age at entry: NR Dating criteria: Accurate LMP; or 1st trimester uterine exam; or 1st or 2nd trimester U/S; or timing of initial auscultated fetal heart tones Parity: NR Bishop score: NR	1) Apgar scores < 7 at 1 minute 2) Apgar scores < 7 at 5 minutes 3) Meconium aspiration 4) Low birthweight 5) Stillbirth 6) Neonatal death 7) Major neonatal morbidity 8) Elective induction 9) C-sections 10) Total adverse outcomes 11) 2 x 2 tables 12) Predictive values	Outcomes 1-11 reported for subjects who delivered between 41 and 42 weeks (n = 512) and for controls, all of whom (n = 1807) delivered between 41 and 42 weeks. 1) Apgar scores < 7 at 1 minute: Testing: 24/512 (4.7%) No testing: 92/1807 (5.1%) p = not significant 2) Apgar scores < 7 at 5 minutes: Testing: 3/512 (0.6%) No testing: 16/1807 (0.9%) p = not significant 3) Meconium aspiration: Testing: 0/512 No testing: 3/1807 (0.2%) p = not significant 4) Low birthweight (< 10th percentile): Testing: 37/512 (7.2%) No testing: 123/1807 (6.8%) p = not significant 5) Stillbirth: Testing: 0/512 No testing: 3/1807 (0.2%) p = not significant 6) Neonatal death: Testing: 0/512 No testing: 0/1807 p = not significant 7) Major neonatal morbidity: Testing: 0/512 No testing: 7/1807 (0.4%) p = not significant 8) Elective induction: Testing: 62/512 (12%) No testing: 282/1807 (16%)	QUALITY SCORES: TESTING Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: + MANAGEMENT Randomized: - Method of randomization: NA Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: - Dating criteria: + Bishop score: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	prolonged bradycardia Dates: Jan 1984 – Jan 1987 Location: Los Angeles, CA Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None			p = not significant 9) C-sections: *Overall:* Testing: 115/512 (22%) No testing: 396/1807 (22%) p = not significant *For fetal distress:* Testing: 14/512 (2.7%) No testing: 60/1807 (3.3%) p = 0.07 *For other indications:* Testing: 101/512 (20%) No testing: 336/1807 (19%) p = not significant 10) Total number of adverse outcomes: Testing: 0/512 No testing: 13/1807 (0.7%) p < 0.05 11) 2 x 2 tables: 2 x 2 Table 1 (n = 908 subjects who started testing at 41 weeks): Reference standard = Intrapartum fetal distress Screening test = Testing Fetal distress yes / no / Totals: Screen test abn: 16 / 119 / 135 Screen test nl: 7 / 766 / 773 Totals: 23 / 885 / 908 2 x 2 Table 2 (n = 352 subjects who started testing at 42 weeks): Reference standard = Intrapartum fetal distress Screening test = Testing	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Fetal distress yes / no / Totals: Screen test abn: 17 / 60 / 77 Screen test nl: 4 / 271 / 275 Totals: 21 / 331 / 352 12) Predictive values of testing: Positive predictive value significantly higher for testing at 42 weeks than for testing at 41 weeks (21.1% vs. 11.9%, respectively). Negative predictive value significantly lower for testing at 42 weeks than for testing at 41 weeks (98.5% vs. 99.1%, respectively).	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Brar, Horenstein, Medearis, et al., 1989	Design: Case series (prospective), no controls Test(s) studied: 1) Nonstress test (NST) + amniotic fluid volume (AFV) assessment + vascular resistance as measured by Doppler U/S (n = 45) Protocol: NST and AFV performed twice weekly. Reactive NST defined as two accelerations in a 10-minute moving window or an acceleration of 15 beats by 15 seconds. AFV > 5 cm considered normal. Flow velocity waveforms of the left and right uterine artery and the umbilical artery obtained with a continuous wave Doppler U/S. Peak systolic (S) to end-diastolic (D) ratios computed over three different cardiac cycles; mean value calculated and used for analysis. Umbilical artery S:D ratio > 3 considered abnormal, as was any diastolic notching. Uterine artery S:D ratio > 2.6 considered abnormal. Reference standard(s): 1) C-section for fetal distress 2) Meconium 3) Apgar score at 5 minutes 4) Admission to NICU 5) Dysmature Dates: NR Location: Los Angeles, CA Setting: University hospital	No. of subjects at start: 45 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 45 Inclusion criteria: Gestational age ≥ 287 days Exclusion criteria: Medical or obstetric complication Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 287 days required for entry into study) Dating criteria: LMP confirmed by one of following: early pregnancy test; 1st trimester exam; U/S prior to 24 weeks; or fetal heart tones by fetoscopy at 18-20 weeks Parity: NR Bishop score: NR	1) Apgar score < 7 at 5 minutes 2) Meconium 3) Admission to NICU 4) Dysmature 5) C-section for fetal distress 6) 2 x 2 tables 7) Other test performance results	1) Apgar score < 7 at 5 minutes: 8/45 (18%) 2) Meconium: 11/45 (24%) 3) Admission to NICU: 6/45 (13%) 4) Dysmature: 3/45 (7%) 5) C-section for fetal distress: 13/45 (29%) 6) 2 x 2 tables: 2 x 2 table 1: Reference standard = C-section for fetal distress Screening test = Antepartum testing (APT) (NST and AFV) C-section yes / no / Totals: APT abnormal: 9 / 10 / 19 APT normal: 4 / 22 / 26 Totals: 13 / 32 / 45 2 x 2 table 2: Reference standard = Meconium Screening test = Antepartum testing (APT) (NST and AFV) Meconium yes / no / Totals: APT abnormal: 10 / 9 / 19 APT normal: 1 / 25 / 26 Totals: 11 / 34 / 45	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: - Sample size: - Statistical tests: - Relationship between Doppler studies and fetal outcomes not reported.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Not specified Length of follow-up: None			2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = Antepartum testing (APT) (NST and AFV) Apgar at 5 min < 7 / ≥ 7 / Totals: APT abnormal: 7 / 12 / 19 APT normal: 1 / 25 / 26 Totals: 8 / 37 / 45 2 x 2 table 4: Reference standard = Admission to NICU Screening test = Antepartum testing (APT) (NST and AFV) NICU admission yes / no / Totals: APT abnormal: 5 / 14 / 19 APT normal: 1 / 25 / 26 Totals: 6 / 39 / 45 2 x 2 table 5: Reference standard = Dysmature Screening test = Antepartum testing (APT) (NST and AFV) Dysmature yes / no / Totals: APT abnormal: 2 / 17 / 19 APT normal: 1 / 25 / 26 Totals: 3 / 42 / 45 7) Other test performance results: Umbilical and uterine artery S:D ratios were not significantly different between patients with normal and abnormal	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				antepartum test results. Cerebral S:D and cerebral placental resistance ratios were significantly lower in patients with abnormal antepartum test results.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chauhan, Sullivan, Lutton, et al., 1995	Design: Case series, no controls Test(s) studied: 1) Maternal estimation of birthweight (n = 70) Protocol: Patients interviewed as follows: "With your previous deliveries you looked and felt a certain way, and the newborn(s) weighed X amount. Based solely on those experiences, how much do you think this newborn will weigh?" 2) Clinical estimation of birthweight (n = 40) Protocol: Performed by obstetrician or midwife using Leopold's maneuvers alone (no computations or formulas). Reference standard(s): Actual birthweight Dates: NR; study conducted over a 3-year period Location: NR Setting: 3 unspecified hospitals Type(s) of providers: Unspecified OB/GYNs (n = 3); unspecified midwives (n = 2) Length of follow-up: None	No. of subjects at start: 70, all of whom provided maternal estimation of birthweight, and 40 of whom also received clinical estimation of birthweight Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 70 Inclusion criteria: Gestational age ≥ 41 weeks; parous; in early active labor with singleton gestation; vertex presentation; no evidence of fetal distress Exclusion criteria: None specified Age (mean ± SD, with range): 26.1 ± 4.5 (range, 17-38) Race: NR Gestational age at entry (mean ± SD, with range): 41.5 ± 0.6 weeks (range, 41-43 weeks) Dating criteria: LMP plus early obstetric examination *or* U/S before 20 weeks Parity (mean ± SD, with range): 1.4 ± 0.6 (range, 1-4) Bishop score: NR	1) Absolute error of birthweight estimate (absolute value of estimate - actual birthweight) 2) Standardized error of birthweight estimate (absolute error [g]/actual birthweight [kg]) 3) Percentage of estimates within ± 10% of actual birthweight 4) Sensitivity, specificity, and positive and negative predictive values of estimates ≥ 4000 g for predicting actual birthweight ≥ 4000g 5) Incidence of macrosomia (birthweight ≥ 4000 g)	1) Absolute error of birthweight estimate (mean ± SD; n = 40 women with both maternal and clinical estimates): Clinical estimate: 278 ± 232 g Maternal estimate: 349 ± 331 g p = not significant 2) Standardized error of birthweight estimate (mean ± SD; n = 40 women with both maternal and clinical estimates): Clinical estimate: 75 ± 71 g Maternal estimate: 92 ± 81 g p = not significant 3) Percentage of estimates within ± 10 of actual birthweight (mean ± SD; n = 40 women with both maternal and clinical estimates): Clinical estimate: 65.0% Maternal estimate: 67.5% p = not significant 4) Sensitivity, specificity, and positive and negative predictive values of estimates ≥ 4000 g for predicting actual birthweight ≥ 4000g: Maternal estimates (n = 70): Sensitivity: 56% Specificity: 94% + predictive value: 77% - predictive value: 86% Clinical estimates (n = 40): Sensitivity: 62% Specificity: 92% + predictive value: 70% - predictive value: 82% 5) Incidence of macrosomia: 18/70 (25.7%)	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: ? Testing protocol described: + Sample size: - Statistical tests: + Differential sample size – 70 for maternal estimates vs. 40 for clinical estimates.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chauhan, Sullivan, Magann, et al., 1994	Design: Case series (prospective), no controls Test(s) studied: Note: Birthweight estimated for each participant using both of the following methods: 1) Clinical estimate of birthweight Protocol: Estimated in early labor by clinician using Leopold maneuvers. 2) Sonographic estimate of birthweight Protocol: Same clinician obtained standard sonographic measurements of transverse abdominal diameter, anteroposterior abdominal diameter, and femur length, also in early labor. Reference standard(s): 1) Actual birthweight Dates: NR; study conducted over a 2-year period Location: Jackson, MS Setting: Community hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 84 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 84 Inclusion criteria: Gestational age ≥ 41 weeks Exclusion criteria: None specified Age (mean ± SD): 25.9 ± 4.7 Race: NR Gestational age at entry: NR (gestational age ≥ 41 weeks required for entry into study) Dating criteria: LMP + physical exam in 1st trimester or U/S at 20 weeks or earlier Parity (mean ± SD): 0.6 ± 0.7 Bishop score: NR	1) Mean absolute error of the two methods of estimating birthweight 2) Mean percentage absolute error 3) Percentage of estimates within 10% of actual birthweight 4) 2 x 2 tables	1) Mean absolute error of the two methods of estimating birthweight (± SD): Clinical: 322 ± 253 g Sonographic: 547 ± 425 g p < 0.001 2) Mean percentage absolute error (± SD): Clinical: 8.9 ± 7.1 g/kg Sonographic: 14.8 ± 11.0 g/kg p < 0.001 3) Percentage of estimates within 10% of actual birthweight: Clinical: 65.4% Sonographic: 42.8% p < 0.005 4) 2 x 2 tables: 2 x 2 table 1: Reference standard = Actual birthweight Screening test = Clinical estimate of birthweight Actual birthweight ≥ 4000 g / < 4000 g / Totals: Clin est ≥ 4000 g: 10 / 2 / 12 Clin est < 4000 g: 10 / 62 / 72 Totals: 20 / 64 / 84 2 x 2 table 2: Reference standard = Actual birthweight Screening test = Sonographic estimate of birthweight Actual birthweight ≥ 4000 g / < 4000 g / Totals: Sonog est ≥ 4000 g: 11 / 6 / 17 Sonog est < 4000 g: 9 / 58 / 67 Totals: 20 / 64 / 84	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: +

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chervenak, Divon, Hirsch, et al., 1989	Design: Case series (not specified if prospective or retrospective), with concomitant controls Test(s) studied: 1) Nonstress test (NST) + amniotic fluid volume (AFV) assessment + U/S estimation of fetal weight (n = 317 cases) Protocol: NST and AFV performed twice weekly. Fetal weight estimated (timing not specified) by biparietal diameter, femur length, and abdominal circumference. Estimated weight did not determine management. Reference standard(s): 1) Actual birthweight Dates: Jan 1987- June 1988 Location: NR Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up:	No. of subjects at start: 317 cases; 100 controls (consecutive patients between 38 and 40 weeks gestational age with no antepartum complications) Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 317 cases; 100 controls Inclusion criteria: Singleton, uncomplicated pregnancy; intact membranes; gestational age > 41 weeks Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry (mean ± SD): Cases, 42 ± 0.6 weeks; controls, 39.8 ± 0.5 weeks Dating criteria: LMP plus early first examination and U/S at < 20 weeks Parity: NR Bishop score: NR	1) Birthweight (mean) 2) Birthweight > 4000 g 3) C-sections 4) 2 x 2 table 5) Other test performance results	1) Birthweight (mean ± SD): Among study patients who delivered at 41 completed weeks (n = 172): 3710 ± 452 g Among study patients who delivered at ≥ 42 completed weeks (n = 145): 3705 ± 454 g Among control patients (n = 100): 3339 ± 360 g No p-values reported 2) Birthweight > 4000 g: Study patients: 81/317 (25.6%) Controls: 6/100 (6%) p < 0.05 3) C-sections: *Overall:* Study patients: 76/317 (24.0%) Controls: 4/100 (4%) p < 0.05 *Primary and repeat C-sections (study patients only):* Primary C-sections: 72/317 (22.7%) Repeat C-sections: 4/317 (1.3%) *C-sections for arrest or protraction disorders (study patients only):* Birthweights > 4000 g: 18/81 (22%) Birthweights < 4000 g: 23/235 (10%) p < 0.01 4) 2 x 2 table: Reference standard = Actual birthweight Screening test = Estimated birthweight (EBW) Actual birthweight > 4000 g / < 4000 g / Totals: EBW > 4000 g: 49 / 22 / 71 EBW < 4000 g: 32 / 214 / 246 Totals: 81 / 236 / 317	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - Unclear whether estimated fetal weight available to practitioner – possibility of bias in outcome of C-section. Morbidity related to macrosomia not reported. *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				5) Other test performance results: *Performance characteristics of estimated birthweight > 4000 g for predicting actual birthweight > 4000 g:* Sensitivity, 61%; specificity, 91%; positive predictive value, 70%; negative predictive value, 87% *Percentage of estimates within 15% of actual birthweight:* When based on biparietal diameter and abdominal circumference: 88% When based on biparietal diameter and femur length: 87% *Percentage of estimates within 10% of actual birthweight:* When based on biparietal diameter and abdominal circumference: 70% When based on biparietal diameter and femur length: 68% *Mean percentage error of estimates (± SD):* 7.5% ± 6.4%	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Crowley, O'Herlihy, and Boylan, 1984	Design: Cohort study Test(s) studied: 1) U/S assessment of AFV Protocol: AFV assessed at 42 weeks and every 4 days thereafter until delivery. If AFV reduced (no vertical pool measuring > 3 cm), then labor induced by amniotomy and oxytocin 24 hours later, if needed. Reference standard(s): 1) Meconium staining 2) C-section for fetal distress 3) Low birthweight (< 10th percentile) 4) Admission to NICU Dates: NR Location: Dublin, Ireland Setting: Unspecified hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 335 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 335 Inclusion criteria: Singleton pregnancy at 42 weeks Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: 42 weeks Dating criteria: Certain LMP or early U/S Parity: 138/335 (41%) primigravidae; 197/335 (59%) multigravidae Bishop score: NR	1) Meconium, grade I 2) Meconium, grade II or III 3) Low birthweight (< 10th percentile) 4) Admission to NICU 5) Convulsions 6) Abnormal tone and primitive reflexes 7) C-sections 8) 2 x 2 tables 9) Other test performance results	1) Meconium, grade I: 24/335 (7%) 2) Meconium, grade II or III: 24/335 (7%) 3) Low birthweight (< 10th percentile): 37/335 (11%) 4) Admission to NICU: 24/335 (7%) 5) Convulsions: 0/335 6) Abnormal tone and primitive reflexes: 2/335 (< 1%) 7) C-sections: 26/335 (8%) Overall: 26/335 (8%) For fetal distress: 9/335 (3%) For dystocia: 8/335 (2%) For failed induction: 3/335 (< 1%) Elective: 6/335 (2%) 8) 2 x 2 tables: <u>2 x 2 Table 1:</u> Reference standard = C-section for fetal distress Screening test = AFV (abn < 3 cm; nl > 3 cm) |	C-section	No C-section	Totals:
---	---	---	---
AFV abn	7	58	65
AFV nl	2	268	270
Totals:	9	326	335	 <u>2 x 2 Table 2:</u> Reference standard = Low birthweight (BW) (< 10th percentile) Screening test = AFV (abn < 3 cm; nl > 3 cm)	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				BW BW low / not low / Totals: AFV abn: 17 / 48 / 65 AFV nl: 20 / 250 / 270 Totals: 37 / 298 / 335 2 x 2 Table 3: Reference standard = Admission to NICU Screening test = AFV (abn < 3 cm; nl > 3 cm) NICU yes / NICU no / Totals: AFV abn: 9 / 56 / 65 AFV nl: 15 / 255 / 270 Totals: 24 / 311 / 335 9) Other test performance results: Clinical assessment of AFV by abdominal palpation showed a false positive rate of 25% and a false negative rate of 43% for detecting "significant meconium staining or absent amniotic fluid." Sensitivity, 75%; specificity, 57%.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Devoe and Sholl, 1983	Design: Case series, no controls Test(s) studied: 1) Maternal estriol + fetal heart rate tests (NST and CST) Protocol: Serial maternal urinary or plasma estriol tests performed biweekly. NST performed weekly and considered reactive if 3 or more accelerations of > 15 bpm amplitude and 15-second duration occurred, with fetal movements, in 30 minutes. If NST nonreactive, then CST performed. CST considered positive if at least 30% of contractions, occurring at a rate of 3/10 min, were followed by late decelerations in a 30-min period. CST equivocal if fewer late decelerations occurred and negative if no late decelerations occurred. Labor induced "either for elective reasons or because of abnormal fetal test results." Method of induction not described. Reference standard(s): 1) Apgar score at 5 minutes 2) Intrapartum fetal distress Dates: July 1977-June 1981 Location: Chicago, IL Setting: University hospital	No. of subjects at start: 248 Dropouts: NR Loss to follow-up: NA No. of subjects at end: 248 (if no dropouts) Inclusion criteria: Singleton pregnancy; unripe cervix at 40 weeks Exclusion criteria: Significant medical or OB complications Age: NR Race: NR Gestational age at entry: 40 weeks Dating criteria: Known LMP confirmed by OB milestones, early clinical exam, or U/S Parity: NR Bishop score: NR	1) Meconium staining 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Perinatal mortality 5) Intrauterine growth retardation (IUGR) 6) Post-maturity syndrome 7) Intrapartum fetal distress – defined as presence of 2 or more of the following: (a) persistent fetal tachycardia or bradycardia; (b) loss of beat-to-beat variability; (c) severe variable or late decelerations; (d) passage of thick, fresh meconium; or (e) scalp pH < 7.22 8) C-sections 9) 2 x 2 tables	1) Meconium staining: 74/248 (30%) 2) Apgar score < 7 at 5 minutes: 7/248 (3%) 3) Birthweight (mean ± SD): 3418 ± 443 g 4) Perinatal mortality: 2/248 (<1%) 5) IUGR: 7/248 (3%) 6) Post-maturity syndrome: 13/248 (5%) 7) Intrapartum fetal distress: 43/248 (17%) 8) C-sections: 34/248 (14%) 9) 2 x 2 tables: 2 x 2 Table 1: Reference standard = Apgar score at 5 minutes Screening test = FHR tests (NST and CST) |	Apgar < 7	Apgar ≥ 7	Totals:
---	---	---	---
NST non-r CST pos	0	22	22
NST non-r CST neg	0	17	17
NST r (no CST)	7	202	209
Totals:	7	241	248		QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Unspecified OB/GYN Length of follow-up: None			2 x 2 Table 2: Reference standard = Fetal distress (yes/no) Screening test = FHR tests (NST and CST) Distress yes / Distress no / Totals: NST non-r CST pos: 6 / 16 / 22 NST non-r CST neg: 6 / 11 / 17 NST r (no CST): 31 / 178 / 209 Totals: 43 / 205 / 248 2 x 2 Table 3: Reference standard = Apgar score at 5 minutes Screening test = Maternal estriol ("low" = below the 10th percentile for gestational age; "falling" = drop of more than 40% from mean of the 3 highest preceding values) Apgar < 7 / Apgar ≥ 7 / Totals: Estriol low or falling: 0 / 46 / 46 Estriol nl: 6 / 166 / 172 Totals: 6 / 212 / 218 2 x 2 Table 4: Reference standard = Fetal distress (yes/no) Screening test = Maternal estriol (as above) Distress yes / Distress no / Totals: Estriol low or falling: 4 / 42 / 46 Estriol nl: 31 / 141 / 172 Totals: 35 / 183 / 218	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Eden, Gergely, Schifrin, et al., 1982	Design: Case series (prospective), no controls Test(s) studied: 1) NST + CST (n = 78) Protocol: Weekly NST. If NST nonreactive, then CST. If CST negative, then repeat NST in 1 week. If CST suspicious, then repeat NST in 1 day. If CST positive, then deliver. 2) NST + modified biophysical profile (MBP) (n = 398) Protocol: Semi-weekly NST. If NST nonreactive, then MBP performed. If MBP normal, then NST repeated semi-weekly. If MBP abnormal, then deliver. 3) NST + AFV + MBP (n = 109) Protocol: Semi-weekly NST + weekly AFV. If AFV decreased, then deliver. If NST nonreactive and AFV normal, then perform MBP. If MBP normal, then resume semi-weekly NST and weekly AFV. If MBP abnormal, then deliver. Reference standard(s): 1) Apgar scores at 1 minute 2) Apgar scores at 5 minutes 3) Meconium aspiration 4) Resuscitation 5) C-section Dates: Nov 1978 – Aug 1981 Location: Los Angeles, CA	No. of subjects at start: 585 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 585 Inclusion criteria: 42 weeks of gestation; prenatal care for ≥ 20 weeks Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR Dating criteria: LMP with consistent exams, *or* sequential U/S exams Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Meconium aspiration 4) Resuscitation 5) Fetal distress requiring intervention (persistent abnormal FHR patterns) 6) Morbidity (defined as presence of any of following: fetal distress requiring intervention, 5-minute Apgar score < 7, neonatal resuscitation, postmaturity syndrome, meconium aspiration) 7) C-sections 8) 2 x 2 tables	1) Apgar score < 7 at 1 minute: NST + CST: 15.4% NST + MBP: 13.1% NST + AFV + MBP: 7.3% no significant differences 2) Apgar score < 7 at 5 minutes: NST + CST: 10.3% NST + MBP: 2.3% NST + AFV + MBP: 0 1 vs. 2, p < 0.05 1 vs. 3, p < 0.05 3) Meconium aspiration: NST + CST: 6.4% NST + MBP: 1.3% NST + AFV + MBP: 0 1 vs. 2, p < 0.05 4) Resuscitation: NST + CST: 12.8% NST + MBP: 10.1% NST + AFV + MBP: 0 1 vs. 2, p < 0.05 2 vs. 3, p < 0.05 5) Fetal distress: NST + CST: 21.8% NST + MBP: 4.5% NST + AFV + MBP: 5.5% 1 vs. 2, p < 0.05 1 vs. 3, p < 0.05 6) Morbidity: NST + CST: 25.6% NST + MBP: 14.3% NST + AFV + MBP: 5.5% 1 vs. 2, p < 0.05 1 vs. 3, p < 0.05 2 vs. 3, p < 0.05 7) C-sections: NST + CST: 11.5% NST + MBP: 29.9%	QUALITY SCORES: TESTING Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: - MANAGEMENT Randomized: - Method of randomization: NA Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: - Gestational age: - Dating criteria: + Bishop score: - Women with complications of pregnancy (e.g., preeclampsia, diabetes, previous stillbirth) NOT excluded. *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None			NST + AFV + MBP: 29.4% 1 vs. 2, $p < 0.05$ 1 vs. 3, $p < 0.05$ 8) 2 x 2 tables: <u>2 x 2 Table 1:</u> Reference standard = Apgar score at 1 minute Screening test = AFV For patients in NST + CST group only (n = 78) Apgar at 1 min < 7 / ≥ 7 / Totals: AFV decreased 7 20 27 AFV nl 6 45 51 Totals: 13 65 78 <u>2 x 2 Table 2:</u> Reference standard = Apgar score at 1 minute Screening test = AFV For patients in NST + MBP group only (n = 109) Apgar at 1 min < 7 / ≥ 7 / Totals: AFV decreased 4 22 26 AFV nl 4 79 83 Totals: 8 101 109 <u>2 x 2 Table 3</u>: Reference standard = Apgar score at 5 minutes Screening test = AFV For patients in NST + CST group only (n = 78) Apgar at 5 min < 7 / ≥ 7 / Totals: AFV decreased 7 20 27 AFV nl 1 50 51 Totals: 8 70 78 *(continued on next page)*	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 Table 4: Reference standard = Apgar score at 5 minutes Screening test = AFV For patients in NST + MBP group only (n = 109) Apgar at 5 min < 7 / ≥ 7 / Totals: AFV decreased: 0 / 26 / 26 AFV nl: 0 / 83 / 83 Totals: 0 / 109 / 109 2 x 2 Table 5: Reference standard = Meconium aspiration Screening test = AFV For patients in NST + CST group only (n = 78) Meconium aspiration yes / no / Totals: AFV decreased: 4 / 23 / 27 AFV nl: 1 / 50 / 51 Totals: 5 / 73 / 78 2 x 2 Table 6: Reference standard = Meconium aspiration Screening test = AFV For patients in NST + MBP group only (n = 109) Meconium aspiration yes / no / Totals: AFV decreased: 0 / 26 / 26 AFV nl: 0 / 83 / 83 Totals: 0 / 109 / 109	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 Table 7: Reference standard = Resuscitation Screening test = AFV For patients in NST + CST group only (n = 78) Resuscitation yes / no / Totals: AFV decreased: 6 / 21 / 27 AFV nl: 2 / 49 / 51 Totals: 8 / 70 / 78 2 x 2 Table 8: Reference standard = C-section Screening test = AFV For patients in NST + CST group only (n = 78) C-section yes / no / Totals: AFV decreased: 9 / 18 / 27 AFV nl: 10 / 41 / 51 Totals: 19 / 59 / 78 2 x 2 Table 9: Reference standard = Apgar score at 1 minute Screening test = FHR decelerations For patients in NST + CST group only (n = 78) Apgar at 1 min < 7 / ≥ 7 / Totals: FHR dec present: 5 / 5 / 10 FHR dec absent: 7 / 61 / 68 Totals: 12 / 66 / 78 2 x 2 Table 10: Reference standard = Apgar score at 5 minutes Screening test = FHR decelerations For patients in NST + CST group only (n = 78)	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Apgar at 5 min < 7 / ≥ 7 / Totals: FHR dec present: 4 / 6 / 10 FHR dec absent: 4 / 64 / 68 Totals: 8 / 70 / 78 2 x 2 Table 11: Reference standard = Meconium aspiration Screening test = FHR decelerations For patients in NST + CST group only (n = 78) Meconium aspiration yes / no / Totals: FHR dec present: 1 / 9 / 10 FHR dec absent: 5 / 63 / 68 Totals: 6 / 72 / 78 2 x 2 Table 12: Reference standard = Resuscitation Screening test = FHR decelerations For patients in NST + CST group only (n = 78) Resuscitation yes / no / Totals: FHR dec present: 5 / 5 / 10 FHR dec absent: 6 / 62 / 68 Totals: 11 / 67 / 78 2 x 2 Table 13: Reference standard = C-section Screening test = FHR decelerations For patients in NST + CST group only (n = 78)	*(continued on next page*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes

	C-section		
	yes	no	Totals:
FHR dec present	2	8	10
FHR dec absent	7	61	68
Totals:	9	69	78

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Farmakides, Schulman, Winter, et al., 1988	Design: Case series, no controls Test(s) studied: 1) Nonstress testing (NST) plus Doppler velocimetry Protocol: Testing interval not specified. Management based on NST, but not Doppler velocimetry. Precise management protocols not described. Reference standard(s): 1) C-section for fetal distress 2) Admission to NICU 3) Small for gestational age Dates: "During 1985" Location: Stony Brook, NY Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 140 (46 of whom were "post-dates") Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 140 Inclusion criteria: Women referred for pre-natal testing for a variety of indications Exclusion criteria: None Age: NR Race: NR Gestational age at entry: NR Dating criteria: NR Parity: NR Bishop score: NR Other: Indications for prenatal testing: Post-dates: 46 (33%) Hypertension: 33 (24%) Diabetes: 14 (10%) Suspected IUGR: 10 (7%) Congenital anomaly: 4 (3%) Other: 33 (24%)	1) Fetal distress (not defined) 2) Small for gestational age (not defined) 3) Admission to NICU 4) C-section for fetal distress 5) 2 x 2 tables	1) Fetal distress: 41/140 (29%). "Most" of the cases of fetal distress came from the post-dates subgoup. 2) Small for gestational age: 15/140 (11%) 3) Admission to NICU: 24/140 (17%) 4) C-section for fetal distress: 39/140 (28%). In the group with abnormal NST, but normal velocimetry, there were significantly more women undergoing C-sections for fetal distress. Again, the majority of these women were in the post-dates subgroup. 5) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress Screening test = Nonstress test (NST) C-section yes / no / Totals: NST abn: 26 / 34 / 60 NST nl: 13 / 67 / 80 Totals: 39 / 101 / 140 <u>2 x 2 table 2:</u> Reference standard = NICU admission Screening test = Nonstress test (NST) NICU admission yes / no / Totals: NST abn: 13 / 47 / 60 NST nl: 11 / 69 / 80 Totals: 24 / 116 / 140 <u>2 x 2 table 3:</u> Reference standard = Small for gestational age (SGA) (not defined) Screening test = Nonstress test (NST) *(continued on next page)*	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: - Sample size: - Statistical tests: - Results not reported separately for subgroup of patients referred for pre-natal testing for "post-date" pregnancy (33% of total study population).

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes

	SGA	Not SGA	Totals:
NST abn	9	51	60
NST nl	6	74	80
Totals:	15	125	140

2 x 2 table 4:
Reference standard = NST
Screening test = Velocimetry

	NST abn	NST nl	Totals:
Velocimetry abnormal	16	28	44
Velocimetry normal	44	52	96
Totals:	60	80	140

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Fleischer, Schulman, Farmakides, et al., 1985	Design: Cohort study (retrospective) Test(s) studied: 1) Nonstress testing (NST) (n = 228) Protocol: NST started at 41 weeks' gestation. If score normal (7-10), then NST repeated weekly. If score inconclusive (5-6), then test repeated within 24 hours or followed by a contraction stress test. Patients with abnormal scores on NST (1-4) were evaluated for delivery. 2) No monitoring (n = 30) Protocol: No antenatal monitoring or NST within 7 days of delivery. Reference standard(s): 1) C-section for fetal distress 2) Apgar ≤ 6 at 1 minute 3) Apgar ≤ 6 at 5 minutes 4) Admission to NICU 5) Neonatal death 6) Stillbirth Dates: Jan 1980 - June 1981 Location: Bronx, NY Setting: University hospital Type(s) of providers: General OB/GYN; MFM Length of follow-up: None	No. of subjects at start: 258 (228 of whom received NST within 7 days of delivery) Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 258 Inclusion criteria: Gestational age ≥ 42 weeks at time of delivery Exclusion criteria: Twin gestation; breech presentation; congenital anomalies; chorioamnionitis Age: NR Race: NR Gestational age at entry: NR (NST initiated at 41 weeks) Dating criteria: Consistency between uterine size and gestational age by LMP Parity: NR Bishop score: NR	1) Apgar score ≤ 6 at 1 minute 2) Apgar score ≤ 6 at 5 minutes 3) Admission to NICU 4) Neonatal death 5) Stillbirth 6) C-section for fetal distress 7) 2 x 2 tables	1) Apgar score ≤ 6 at 1 minute: NST: 15/228 (7%) No monitoring: 15/30 (50%) p < 0.01 2) Apgar score ≤ 6 at 5 minutes: NST: 7/228 (3%) No monitoring: 6/30 (20%) p < 0.05 3) Admission to NICU: NST: 7/228 (3%) No monitoring: 5/30 (17%) p < 0.05 4) Neonatal death: NST: 3/228 (1%) No monitoring: 1/30 (3%) p = not significant 5) Stillbirth: NST: 2/228 (1%) No monitoring: 4/30 (13%) p < 0.05 6) C-section for fetal distress: NST: 26/228 (11%) No monitoring: 19/30 (63%) p < 0.001 7) 2 x 2 tables (for patients in the NST group only, n = 228) <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress Screening test = Nonstress test (NST) C-section yes / no / Totals: NST 1-4: 6 / 4 / 10 NST 5-6: 5 / 23 / 28 NST 7-10: 15 / 175 / 190 Totals: 26 / 202 / 228	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 2: Reference standard = Apgar score at 1 minute Screening test = Nonstress test (NST) Apgar at 1 min ≤ 6 / > 6 / Totals: NST 1-4: 4 / 6 / 10 NST 5-6: 2 / 26 / 28 NST 7-10: 9 / 181 / 190 Totals: 15 / 213 / 228 2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = Nonstress test (NST) Apgar at 5 min ≤ 6 / > 6 / Totals: NST 1-4: 1 / 9 / 10 NST 5-6: 1 / 27 / 28 NST 7-10: 5 / 185 / 190 Totals: 7 / 221 / 228 2 x 2 table 4: Reference standard = Admission to NICU Screening test = Nonstress test (NST) NICU admission yes / no / Totals: NST 1-4: 4 / 6 / 10 NST 5-6: 0 / 28 / 28 NST 7-10: 3 / 187 / 190 Totals: 7 / 221 / 228 2 x 2 table 5: Reference standard = Neonatal death Screening test = Nonstress test (NST)	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Neonatal death yes / no / Totals: NST 1-4: 3 / 7 / 10 NST 5-6: 0 / 28 / 28 NST 7-10: 0 / 190 / 190 Totals: 3 / 225 / 228 2 x 2 table 6: Reference standard = Stillbirth Screening test = Nonstress test (NST) Stillbirth yes / no / Totals: NST 1-4: 0 / 10 / 10 NST 5-6: 0 / 28 / 28 NST 7-10: 2 / 188 / 190 Totals: 2 / 226 / 228	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Gilby, Williams, and Spellacy, 2000	Design: Cohort study Test(s) studied: U/S within 7 days of delivery to measure abdominal circumference Reference standard(s): Macrosomia (defined using two different thresholds, 4000 g and 4500 g) Dates: 1992-1997 Location: Tampa, FL Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 1996 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 1996 Inclusion criteria: Singleton pregnancies with U/S within 7 days of delivery Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR Dating criteria: NR Parity: NR Bishop score: NR	1) 2 x 2 tables 2) Other test performance results	1) 2 x 2 tables: Article includes ROC curves for performance of different abdominal circumference cutoff points (35 cm and 38 cm) for predicting macrosomia at two thresholds, 4000 g and 4500 g. 2 x 2 tables could be constructed only for the 4500 g macrosomia cutoff point. <u>2 x 2 Table 1:</u> Reference standard = Macrosomia (birthweight ≥ 4500 g) Screening test = Abdominal circumference (Abd C), cutoff point at 35 cm |	BW ≥ 4500 g	BW < 4500 g	Totals:
---	---	---	---
Abd C ≥ 35 cm	68	683	751
Abd C < 35 cm	1	1244	1245
Totals:	69	1927	1996 | <u>2 x 2 Table 2:</u> Reference standard = Macrosomia (birthweight ≥ 4500 g) Screening test = Abd C, cutoff point at 38 cm |	BW ≥ 4500 g	BW < 4500 g	Totals:
---	---	---	---
Abd C ≥ 38 cm	37	62	99
Abd C < 38 cm	32	1865	1897
Totals:	69	1927	1996	 2) Other test performance results: Abdominal circumference ≥ 35 cm had the following test performance characteristics: Sensitivity, 98.5%; specificity, 64.5%; negative predictive value, 64.5%	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: -

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Gilson, O'Brien, Vera, et al., 1988	Design: Case series (not specified if prospective or retrospective), no controls Test(s) studied: 1) Nonstress test (NST) + biophysical profile (BP) (n = 128) Protocol: Testing started when patient "almost" 42 weeks. NST performed twice weekly, BP weekly at first and twice weekly after 43 weeks. Cervix examined at each visit. If BP score 8-10, then patient given another NST in 3-4 days and a repeat BP in 7 days. If BP score 5-7, then BP repeated in 24 hours; if still abnormal, then patient transferred to hospital for induction. If oligohydramnios, spontaneous decelerations on NST, or score < 4, then patient induced. Patients with BP scores of 8-10 allowed to deliver in birthing center if NST reactive and no indication of fetal distress or failure to progress. Otherwise transferred to hospital for labor and delivery. Reference standard(s): 1) Apgar scores at 1 and 5 minutes 2) Post-maturity syndrome 3) Fetal distress 4) C-section for fetal distress Dates: Jan 1984 - Feb 1986 Location: Brownsville, TX Setting: Freestanding birthing center	No. of subjects at start: 178 Dropouts: 50 (delivered before biophysical profile score assessed) Loss to follow-up: NA No. of subjects at end: 128 Inclusion criteria: Gestational age 42 completed weeks; otherwise low risk; biophysical profile score recorded within 1 week of delivery Exclusion criteria: None specified Age: NR Race: 100% Hispanic Gestational age at entry: NR (gestational age of 42 completed weeks required for entry into study) Dating criteria: Clinical sizing (LMP supported by appropriate fundal heights), stethoscope fetal heart tones (for more than 22 weeks), or 2nd trimester U/S Parity: NR Bishop score: NR	1) 2 x 2 tables	1) 2 x 2 tables: 2 x 2 Table 1: Reference standard = Apgar score at 1 minute Screening test = Biophysical profile score (BPS) Apgar < 7 / Apgar ≥ 7 / Totals: BPS < 8: 2 / 24 / 26 BPS 8-10: 12 / 90 / 102 Totals: 14 / 114 / 128 2 x 2 Table 2: Reference standard = Apgar score at 5 minutes Screening test = BPS Apgar < 7 / Apgar ≥ 7 / Totals: BPS < 8: 0 / 26 / 26 BPS 8-10: 0 / 102 / 102 Totals: 0 / 128 / 128 2 x 2 Table 3: Reference standard = Post-maturity syndrome Screening test = BPS Post-maturity: yes / no / Totals: BPS < 8: 7 / 19 / 26 BPS 8-10: 6 / 96 / 102 Totals: 13 / 115 / 128 2 x 2 Table 4: Reference standard = Fetal distress Screening test = BPS Fetal distress: yes / no / Totals: BPS < 8: 4 / 22 / 26 BPS 8-10: 11 / 91 / 102 Totals: 15 / 113 / 128	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: ? Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - Study underpowered to detect differences in categorical variables. *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Unspecified OB/GYN; nurse midwives Length of follow-up: None			2 x 2 Table 5: Reference standard = C-section for fetal distress Screening test = BPS C-section for fetal distress yes / no / Totals: BPS < 8: 2 / 24 / 26 BPS 8-10: 1 / 101 / 102 Totals: 3 / 125 / 128	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Hann, McArdle, and Sachs, 1987	Design: Case series, no controls Test(s) studied: Biophysical Profile Score (BPS). Included 6 components: 1) NST; 2) fetal breathing movements; 3) fetal movements; 4) fetal tone; 5) amniotic fluid volume (AFV); and 6) placental grading. Score of 0-2 given to each variable. Abnormal score defined as < 6. Patients with scores of 4-6 managed "on an individualized basis"; those with scores < 4 delivered immediately. Reference standard(s): "Poor neonatal outcome," which included neonatal distress requiring admission to the NICU, endotracheal intubation, use of positive pressure oxygen for more than 6 hours, and persistent fetal circulation Dates: NR Location: Boston, MA Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 131 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 131 Inclusion criteria: Gestational age ≥ 41 completed weeks; singleton pregnancy; no congenital anomalies Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 41 completed weeks required for entry into study) Dating criteria: U/S early in pregnancy or reliable menstrual dates and serial physical exams Parity: NR Bishop score: NR	1) Meconium aspiration 2) Admission to NICU 3) Seizure 4) 2 x 2 table 5) Predictive values	1) Meconium aspiration: 5/131 (4%) 2) Admission to NICU: 5/131 (4%) 3) Seizure: 1/131 (< 1%) 4) 2 x 2 table: Reference standard = Poor neonatal outcome Screening test = Biophysical Profile Score (BPS) Neonatal outcome poor / normal / Totals: BPS abn (< 6): 1 / 7 / 8 BPS nl (≥ 6): 6 / 117 / 123 Totals: 7 / 124 / 131 5) Predictive values: *Positive predictive values:* Total BPS: 14% Amniotic fluid volume: 17% Placental grading: 4% Fetal breathing movements: 5% Fetal tone/movements: 40% NST: 14% *Negative predictive values:* Total BPS: 94% Amniotic fluid volume: 95% Placental grading: 91% Fetal breathing movements: 94% Fetal tone/movements: 95% NST: 94%	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: - Testing protocol described: - Sample size: - Statistical tests: +

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Imai, Tani, Saito, et al., 2001	Design: Cohort study Test(s) studied: 1) Fetal fibronectin obtained from posterior vaginal fornix. Collected once between 29 and 35 weeks, then weekly from 36 weeks until parturition. 2) Cytokines Interleukin-1, beta, IL-6, IL8, and tumor necrosis factor alpha. Collected from endocervix at same intervals as above. Reference standard(s): Delivery within 7 days of sampling Dates: NR Location: Kanagawa, Japan Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 122 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 120 (2 excluded for no labor) Inclusion criteria: Singleton pregnancy; vertex presentation Exclusion criteria: Maternal or obstetric complications that might cause premature delivery, premature rupture of membranes, vaginal bleeding, or fetal anomalies Age: Mean, 30; range, 20-45 Race: NR Gestational age at entry: NR (gestational age between 29 and 35 weeks required for entry into study) Dating criteria: LMP, confirmed by ultrasound prior to 20 weeks Parity: 71% nulliparous Bishop score: NR	1) Fetal fibronectin 2) IL-1 beta 3) 2x2 tables	1) Fetal fibronectin: At threshold of > 50 ng/ml: Sensitivity: 90% Specificity: 51% Positive predictive value: 75% Negative predictive value: 75% 2) IL-1 beta: At threshold of 100 pg/ml: Sensitivity: 55% Specificity: 76% Positive predictive value: 79% Negative predictive value: 50% 3) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = Delivery within 7 days Screening test = Fetal Fibronectin (fFN) Time to delivery ≤ 7 days / > 7 days / Totals: fFN > 50: 120 / 39 / 159 fFN ≤ 50: 13 / 40 / 53 Totals: 133 / 79 / 212 <u>2 x 2 table 2:</u> Reference standard = Delivery within 7 days Screening test = IL-2 beta Time to delivery ≤ 7 days / > 7 days / Totals: IL-2 > 100: 73 / 19 / 92 IL-2 ≤ 100: 60 / 60 / 120 Totals: 133 / 79 / 212	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - Reported sensitivity/specificity was reversed in tables and text of article; values from text used here. Any variations by gestational age within the 36-42 week gestational range not reported.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Jazayeri, Heffron, Phillips, et al., 1999	Design: Case series (retrospective), concomitant controls Test(s) studied: 1) U/S measuring estimated fetal weight, abdominal circumference, biparietal diameter, and femur length Protocol: Measurements taken within 2 weeks of delivery Reference standard(s): 1) Macrosomia Dates: Jan-Dec 1996 Location: Tampa, FL Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 168 (84 with macrosomic infants; 84 with nonmacrosomic infants) Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 168 Inclusion criteria: Women with macrosomic infants (≥ 4000 g) and U/S within 2 weeks prior to delivery; these women compared with group of women with non-macrosomic infants and recent U/S Exclusion criteria: None specified Age (mean ± SD): Macrosomic, 25.9 ± 6; nonmacrosomic, 24.4 ± 5 Race: Macrosomic, 45% White, 25% Black, 30% Hispanic; non-macrosomic, 40% White, 30% Black, 30% Hispanic Gestational age at entry (mean ± SD): Macrosomic, 40.1 ± 1.5 weeks; nonmacrosomic, 37.1 ± 3.6 weeks (p = 0.001) Dating criteria: NR Gravidity (median, with range): Macrosomic, 3 ± 2; non-macrosomic, 2 ± 1 Bishop score: NR	1) Shoulder dystocia 2) C-section for fetal distress 3) 2 x 2 table 4) Other test performance results	1) Shoulder dystocia: Macrosomic: 13/84 (15%) Nonmacrosomic: 0/84 p = 0.001 In macrosomic newborns, labor induction was associated with a 22% rate of should dystocia, whereas augmentation and spontaneous labor were each associated with an 8% rate of shoulder dystocia (odds ratio, 3.4; 95% CI, 1.4 to 8.2; p < 0.01). In a multivariable model controlling for birthweight, induction alone was associated with shoulder dystocia (odds ratio, 2.97; 95% CI, 1.24 to 7.4; p < 0.015). 2) C-section for fetal distress: Macrosomic: 25/84 (30%) Nonmacrosomic: 19/84 (23%) p = not significant 3) 2 x 2 table: Reference standard = Macrosomia (≥ 4000 g) Screening test = U/S measurement of abdominal circumference (AC) Macrosomia yes / no / Totals: AC ≥ 35 cm: 75 / 6 / 81 AC < 35 cm: 9 / 78 / 87 Totals: 84 / 84 / 168 4) Other test performance results: Multiple regression analysis showed abdominal circumference to be the best predictor of birthweight in macrosomic infants.	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: - Sample size: - Statistical tests: +

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Knox, Huddleston, and Flowers, 1979	Design: RCT, allocation to group by last digit of hospital number Test(s) studied: 1) Amniocentesis (n = 90) Protocol: If no meconium discovered and fluid obtained, then amniocentesis repeated in 1 week. If meconium discovered or no fluid obtained, then labor induced. Labor induced with IV oxytocin, with direct FHR and intrauterine pressure monitoring. 2) Oxytocin challenge test (OCT) (n = 90) Protocol: Initial amniocentesis followed by OCT. If meconium present or no fluid discovered on amniocentesis, then labor induced. If OCT negative, the repeated in 1 week. If OCT positive, the labor induced. Reference standard(s): 1) Low birthweight 2) Neonatal morbidity 3) Perinatal death 4) C-sections 5) Apgar scores at 1 minute 6) Apgar scores at 5 minutes Dates: Aug 1975 - July 1976 Location: Birmingham, AL Setting: University hospital Type(s) of providers: Unspecified OB/GYN	No. of subjects at start: 187 Dropouts: 7 (excluded due to complications) Loss to follow-up: NA No. of subjects at end: 180 Inclusion criteria: Gestational age ≥ 42 weeks Exclusion criteria: Any obstetric complication Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 42 weeks required for entry into study) Dating criteria: Either a) reliable LMP confirmed by pelvic exam prior to 12 weeks, U/S at 20-30 weeks, or auscultation of unamplified fetal heart tones for at least 22 weeks; or b) if LMP unreliable, then 2 of above 3 assessments consistent with 42 weeks' gestation Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Low birthweight (< 10th percentile) 4) Neonatal morbidity 5) Perinatal death 6) Meconium 7) C-sections 8) Induction 9) Abnormal labor (prolonged latent phase, primary dysfunctional labor, secondary arrest of dilatation, or arrest of descent) 10) 2 x 2 tables 11) Other test performance results	1) Apgar score < 7 at 1 minute: Amniocentesis: 19/90 (21%) OCT: 12/90 (13%) p = not significant 2) Apgar score < 7 at 5 minutes: Amniocentesis: 6/90 (7%) OCT: 2/90 (2%) p = not significant 3) Low birthweight (< 10th percentile): Amniocentesis: 3/90 (3%) OCT: 4/90 (4%) p = not significant 4) Neonatal morbidity: Amniocentesis: 6/90 (7%) OCT: 7/90 (8%) p = not significant 5) Perinatal death: Amniocentesis: 3/90 (3%) OCT: 1/90 (1%) p = not significant 6) Meconium (overall only): On initial amniocentesis: 22% At delivery: 44% 7) C-sections: Amniocentesis: 11/90 (12%) OCT: 8/90 (9%) p = not significant 8) Induction: Amniocentesis: 29/90 (32%) OCT: 11/90 (12%) p < 0.005 9) Abnormal labor: Amniocentesis: 13/90 (14%) OCT: 12/90 (13%) p = not significant	QUALITY SCORES: TESTING Reference standard: + Randomized: + Method of randomization: - Verification bias: + Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: + MANAGEMENT Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: - Dating criteria: + Bishop score: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None			10) 2 x 2 tables: 2 x 2 table 1: Reference standard = Low birthweight (< 10th percentile) Screening test = Meconium at initial amniocentesis Low birthweight yes / no / Totals: Meconium present: 2 / 77 / 79 Meconium absent: 5 / 96 / 101 Totals: 7 / 173 / 180 2 x 2 table 2: Reference standard = Neonatal morbidity Screening test = Meconium at initial amniocentesis Morbidity yes / no / Totals: Meconium present: 6 / 73 / 79 Meconium absent: 7 / 94 / 101 Totals: 13 / 167 / 180 2 x 2 table 3: Reference standard = Perinatal death Screening test = Meconium at initial amniocentesis Death yes / no / Totals: Meconium present: 4 / 75 / 79 Meconium absent: 0 / 101 / 101 Totals: 4 / 176 / 180 2 x 2 table 4: Reference standard = C-sections Screening test = Meconium at initial amniocentesis	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				C-section yes / no / Totals: Meconium present: 11 / 68 / 79 Meconium absent: 8 / 93 / 101 Totals: 19 / 161 / 180 2 x 2 table 5: Reference standard = Apgar score at 1 minute Screening test = Meconium at initial amniocentesis Apgar at 1 min < 7 / ≥ 7 / Totals: Meconium present: 23 / 56 / 79 Meconium absent: 8 / 93 / 101 Totals: 31 / 149 / 180 2 x 2 table 6: Reference standard = Apgar score at 5 minutes Screening test = Meconium at initial amniocentesis Apgar at 5 min < 7 / ≥ 7 / Totals: Meconium present: 8 / 71 / 79 Meconium absent: 0 / 101 / 101 Totals: 8 / 172 / 180 11) Other test performance results: In subset of patients with meconium present, there were no significant differences between the two groups for any outcome.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Leveno, Quirk, Cunningham, et al., 1984	Design: Cohort study (prospective) Test(s) studied: 1) Amniotic fluid volume (AFV) assessment Protocol: AFV assessed weekly. Oligohydramnios defined as two or fewer 1-cm pockets of amniotic fluid. If any of the following occurred, then labor was induced using oxytocin followed by amniotomy: a) certain completion of 43 weeks' gestation; b) absence of amniotic fluid on physical exam; c) markedly diminished fetal activity; or d) development of pregnancy-induced hypertension. Intrapartum electronic FHR monitoring used. Reference standard(s): 1) C-section for fetal distress 2) Small for gestational age 3) Stillbirth or meconium aspiration Dates: July 1980 - July 1982 Location: Dallas, TX Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 727 (of whom 213 underwent U/S assessment of AFV) Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 727 Inclusion criteria: Gestational age ≥ 41 completed weeks Exclusion criteria: Obstetric or medical complications Age: 55% were age 20-30 Race: 39% White, 39% Black, 22% Hispanic Gestational age at entry: 42-43 weeks (certain): 16% 43-44 weeks (certain): 8% > 44 weeks (certain): 1% Uncertain prolonged pregnancy: 75% Dating criteria: LMP corroborated by a) fetal heart auscultation between 17 and 20 weeks; *or* b) fundal height measurements between 20 and 30 weeks; *or* c) U/S before 26 weeks Parity: "Approximately half" were nulliparous Bishop score: NR	1) C-sections 2) 2 x 2 tables	1) C-sections: Overall: 196/727 (27%) For cephalopelvic disproportion: 114/727 (16%) For fetal distress:* 59/727 (8%) For abnormal presentation: 16/727 (2%) For other reasons: 7/727 (1%) *"Fetal distress" diagnosed when one or more of the following were identified on intrapartum FHR monitoring: a) repetitive late decelerations; b) severe variable decelerations of < 60 bpm for ≥ 1 minute; c) prolonged decelerations lasting ≥ 2 minutes; or d) unexplained abnormal baseline heart rate or diminished beat-to-beat variability, especially when either accompanied by meconium staining. 2) 2 x 2 tables (for women undergoing AFV assessment only, n = 213) <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress (as defined above) Screening test = Amniotic fluid volume (AFV) assessment C-section yes / no / Totals: AFV decreased: 11 / 73 / 84 AFV normal: 7 / 122 / 129 Totals: 18 / 195 / 213 <u>2 x 2 table 2:</u> Reference standard = Small for gestational age (SGA) Screening test = AFV	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes

	SGA yes	SGA no	Totals:
AFV decreased	8	76	84
AFV normal	8	121	129
Totals:	16	197	213

2 x 2 table 3:
Reference standard = Stillbirth or meconium aspiration
Screening test = AFV

	Stillbirth/ meconium yes	no	Totals:
AFV decreased	2	82	84
AFV normal	0	129	129
Totals:	2	211	213

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Monaghan, O'Herlihy, and Boylan, 1987	Design: Cohort study (not specified if prospective or retrospective) Test(s) studied: 1) Ultrasound used to measure deepest amniotic fluid pool and to grade placental echogenic changes (n = 200) Protocol: U/S scans performed every 3-5 days beginning at 42 weeks. Used to measure deepest vertical amniotic fluid pool. If no pool exceeded 30 mm, then oligohydramnios diagnosed and labor induced. U/S also used to grade echogenic characteristics of placenta from 0 (homogeneous placenta with smooth chorionic plate) to III (placenta completely divided into compartments by indentation of the chorionic plate extending all the way to the basal layer). Placental grading not used to make management decisions. Reference standard(s): 1) Fetal acidosis 2) C-section for fetal distress 3) Low birthweight 4) Admission to NICU 5) Perinatal death Dates: NR Location: Dublin, Ireland Setting: Unspecified hospital Type(s) of providers:	No. of subjects at start: 225 Dropouts: 25 (excluded because of uncertain gestational age) Loss to follow-up: NA No. of subjects at end: 200 Inclusion criteria: Gestational age ≥ 42 weeks; singleton pregnancy Exclusion criteria: Uncertain gestational age Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 42 weeks required for entry into study) Dating criteria: Certain LMP or early U/S Parity: 41% primiparous Bishop score: NR	1) Fetal acidosis (pH < 7.25) 2) Low birthweight (< 10th percentile) 3) Admission to NICU 4) Perinatal death 5) Inductions 6) C-sections 7) 2 x 2 tables 8) Other test performance results	1) Fetal acidosis: 13/200 (7%) 2) Low birthweight: 23/200 (12%) 3) Admission to NICU: 18/200 (9%) 4) Perinatal death: 2/200 (1%) 5) Inductions: 69/200 (35%) Labor induced in 32 cases because of oligohydramnios, and in 37 cases with favorable cervical status and normal amniotic fluid estimates. 6) C-sections: Overall: 12/200 (6%) For fetal distress: 3/200 (2%) 7) 2 x 2 tables: <u>2 x 2 Table 1:</u> Reference standard = Fetal acidosis (pH < 7.25) Screening test = Amniotic fluid index (AFI) ("low" if no pool exceeded 30 mm) Fetal acidosis yes / no / Totals: AFI low: 3 / 29 / 32 AFI normal: 10 / 158 / 168 Totals: 13 / 187 / 200 <u>2 x 2 Table 2:</u> Reference standard = C-section for fetal distress Screening test = AFI (as above) C-section yes / no / Totals: AFI low: 1 / 31 / 32 AFI normal: 2 / 166 / 168 Totals: 3 / 197 / 200	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Unspecified OB/GYN Length of follow-up: None			2 x 2 Table 3: Reference standard = Low birthweight (< 10th percentile) Screening test = AFI (as above) Low birthweight yes / no / Totals: AFI low: 11 / 21 / 32 AFI normal: 12 / 156 / 168 Totals: 23 / 177 / 200 2 x 2 Table 4: Reference standard = Admission to NICU Screening test = AFI (as above) NICU admission yes / no / Totals: AFI low: 3 / 29 / 32 AFI normal: 15 / 153 / 168 Totals: 18 / 182 / 200 2 x 2 Table 5: Reference standard = Perinatal death Screening test = AFI (as above) Perinatal death yes / no / Totals: AFI low: 0 / 32 / 32 AFI normal: 2 / 166 / 168 Totals: 2 / 198 / 200 8) Other test performance results: Ultimate placental grading was associated with an increased incidence of C-section. The increased incidence associated with grade III placenta was related to mothers with coincident oligohydramnios. The frequency of meconium staining and no amniotic fluid after amniotomy was higher in patients with oligohydramnios. There were no differences in acidosis or	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				NICU admission between pregnancies with normal versus reduced amniotic fluid, or grade 1-11 versus grade III placentas. The incidence of low birthweight was significantly higher in patients with oligohydramnios than in patients with grade III placentas.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Montan and Malcus, 1995	Design: Cohort study (prospective) Test(s) studied: 1) Amniotic fluid index (AFI) and FHR pattern Protocol: AFI and FHR pattern measured at 2-day intervals from 42 weeks until delivery. Labor induced (by oxytocin or artificial rupture of the membranes) for abnormal fetal or maternal findings. Reference standard(s): 1) C-section 2) Apgar < 7 at 1 minute 3) Apgar < 7 at 5 minutes Dates: 1992-93 Location: Ängelholm, Sweden Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 116 women delivered at ≥ 42 weeks gestation; 88 of them had AFI measured at least once before onset of labor Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 116 Inclusion criteria: Gestational age ≥ 42 completed weeks Exclusion criteria: None specified Age (mean, with range): 28 (17-46) Race: NR Gestational age at entry: NR (gestational age required to be ≥ 42 completed weeks for entry into study) Dating criteria: U/S (biparietal diameter and femur length) in weeks 16-19 Parity: 49% primigravida Bishop score: NR	1) 2 x 2 tables 2) Other test performance results	1) 2 x 2 tables 2 x 2 table 1: Reference standard = C-section Screening test = AFI C-section yes / no / Totals: AFI < 5 cm: 1 / 10 / 11 AFI ≥ 5 cm: 11 / 66 / 77 No AFI: 7 / 21 / 28 Totals: 19 / 97 / 116 2 x 2 table 2: Reference standard = Apgar score at 1 minute Screening test = AFI Apgar at 1 min < 7 / ≥ 8 / Totals: AFI < 5 cm: 0 / 11 / 11 AFI ≥ 5 cm: 3 / 74 / 77 No AFI: 1 / 27 / 28 Totals: 4 / 112 / 116 2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = AFI Apgar at 5 min < 7 / ≥ 8 / Totals: AFI < 5 cm: 0 / 11 / 11 AFI ≥ 5 cm: 2 / 75 / 77 No AFI: 0 / 28 / 28 Totals: 2 / 114 / 116 2) Other test performance results: There was no association between low AFI (< 5 cm) and signs of fetal distress expressed as abnormal FHR pattern, meconium staining, Apgar scores < 7, or C-section.	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: + The definition of low AFI used in this study (< 5 cm) is more liberal than that used in many studies (3 cm or 1 cm) and may explain the lack of association between low AFI and fetal compromise reported here.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Mouw, Egberts, Kragt, et al., 1998	Design: Case series (prospective), no controls Test(s) studied: 1) Fetal fibronectin concentration (fFN) Protocol: Fetal fibronectin concentration measured in cervicovaginal secretions obtained in sterile speculum examination at 41 weeks. Concentrations of < 50 ng/ml were interpreted as negative, ≥ 50 ng/ml as positive. Pregnancies were managed expectantly, and induction was performed only for obstetric "or sometimes psychological" reasons. Reference standard(s): 1) Birth within 3 days of fFN testing Dates: NR Location: Leiden and Voorburg, The Netherlands Setting: 2 university hospitals Type(s) of providers: General OB/GYN Length of follow-up: None	No. of subjects at start: 80 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 80 Inclusion criteria: Gestational age ≥ 41 weeks Exclusion criteria: In labor; clinical evidence of ruptured membranes Age (mean ± SD): 31 ± 6 Race: NR Gestational age at entry: Range, 287-304 days Dating criteria: NR Parity (mean ± SD): 1 ± 1 Bishop score: NR	1) 2 x 2 table 2) Other test performance results	1) 2 x 2 table: Reference standard = Birth within 3 days of fFN testing Screening test = fFN |	Birth within 3 days		
	yes	no	Totals:
fFN ≥ 50 ng/ml	30	15	45
fFN< 50 ng/ml	12	27	39
Totals:	42	42	84	 2) Other test performance results: A positive fFN test (≥ 50 ng/ml) had sensitivity of 0.71 (95% CI, 0.58 to 0.86) and specificity of 0.64 (95% CI, 0.48 to 0.78) for predicting birth within 3 days. The change from negative to positive fFN values often occurred between 1 and 4 days before birth in women with a spontaneous onset of labor. The mean interval between positive test and birth was 2.5 ± 2.5 days (range, 0-11). fFN was moderately correlated with Bishop score. Bishop score > 5 had sensitivity 0.67 (95% CI, 0.48 to 0.82) and specificity 0.77 (95% CI, 0.54 to 0.92) for predicting birth within 3 days. (Only 74% of study participants had Bishop scores recorded.)	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: + Sensitivity/specificity results include some repeat tests.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
O'Reilly-Green and Divon, 1996	Design: Cohort study (retrospective) Test(s) studied: 1) Sonographic estimate of fetal weight (EFW) plus nonstress test (NST) and amniotic fluid index (AFI) Protocol: Sonographic EFW done at initial appointment. NST and AFI performed twice weekly. If AFI ≤ 5 cm, then patient delivered within 24 hours, even if all other testing parameters were normal. Reference standard(s): 1) Apgar score at 1 minute 2) Apgar score at 5 minutes 3) Any complication Dates: July 1991- Sep 1992 Location: Bronx, NY Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 449 Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 449 Inclusion criteria: Prolonged pregnancy (defined as 1 or more weeks beyond expected date of delivery) Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR Dating criteria: Nagle's rule or sonographic criteria Parity: NR Bishop score: NR	1) Apgar score < 8 at 1 minute 2) Apgar score < 9 at 5 minutes 3) 2 x 2 tables 4) Other test performance results	1) Apgar score < 8 at 1 minute: 66/449 (15%) 2) Apgar score < 9 at 5 minutes: 24/449 (5%) 3) 2 x 2 tables: 2 x 2 Table 1: Reference standard = Apgar score at 1 minute Screening test = Amniotic fluid index (AFI) Apgar at 1 min < 8 / ≥ 8 / Totals: AFI ≤ 5: 5 / 45 / 50 AFI > 5: 61 / 337 / 398 Totals: 66 / 382 / 448 2 x 2 Table 2: Reference standard = Apgar score at 5 minutes Screening test = Amniotic fluid index (AFI) Apgar at 5 min < 9 / ≥ 9 / Totals: AFI ≤ 5: 2 / 48 / 50 AFI > 5: 22 / 376 / 398 Totals: 24 / 424 / 448 2 x 2 Table 3: Reference standard = Any complication Screening test = Amniotic fluid index (AFI) Complication yes / no / Totals: AFI ≤ 5: 4 / 46 / 50 AFI > 5: 25 / 372 / 397 Totals: 29 / 418 / 447 4) Other test performance results: Additional analyses showed significant association between AFI ≤ 5 and clinical oligohydramnios.	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: - Same population as in O'Reilly-Green and Divon, 1997, below.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
O'Reilly-Green and Divon, 1997	Design: Cohort study (retrospective) Test(s) studied: 1) Sonographic estimate of fetal weight Protocol: Estimate made ≤ 21 days before admission (≤ 22 days before delivery). Estimated fetal weight (EFW) calculated using formulas at the discretion of the clinician interpreting the study. An adjusted EFW was calculated by adding 12.7 g to the EFW for each day that elapsed between the sonographic measurements and delivery. Reference standard(s): 1) Actual birthweight Dates: July 1991 - Sep 1992 Location: Bronx, NY Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None	No. of subjects at start: 445 Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 445 Inclusion criteria: Prolonged pregnancy (defined as 4 or more days beyond expected date of delivery) Exclusion criteria: Diabetes Age: NR Race: NR Gestational age at entry (mean ± SD): 291 ± 6.7 days Dating criteria: Naegele's rule or sonographic criteria Parity: NR Bishop score: NR	1) 2 x 2 tables 2) Other test performance results	1) 2 x 2 tables: 2 x 2 Table 1: Reference standard = Actual birthweight Screening test = Estimated fetal weight (EFW) Birthweight ≥ 4000 g / < 4000 g / Totals: EFW ≥ 3711 g: 91 / 94 / 185 EFW < 3711 g: 16 / 244 / 260 Totals: 107 / 338 / 445 2 x 2 Table 2: Reference standard = Actual birthweight Screening test = EFW Birthweight ≥ 4500 g / < 4500 g / Totals: EFW ≥ 4192 g: 15 / 35 / 50 EFW < 4192 g: 3 / 392 / 395 Totals: 18 / 427 / 445 2 x 2 Table 3: Reference standard = Actual birthweight Screening test = EFW Birthweight ≥ 4000 g / < 4000 g / Totals: EFW ≥ 4000 g: 60 / 29 / 89 EFW < 4000 g: 47 / 309 / 356 Totals: 107 / 338 / 445 2 x 2 Table 4: Reference standard = Actual birthweight Screening test = EFW	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: + Same population as in O'Reilly-Green and Divon, 1996, above.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Birthweight ≥ 4500 g / < 4500 g / Totals: EFW ≥ 4500 g: 4 / 5 / 9 EFW < 4500 g: 14 / 422 / 436 Totals: 18 / 427 / 445 2) Other test performance results: EFW ≥ 3711 g had sensitivity 0.85 and specificity 0.72 for predicting birthweight ≥ 4000 g. EFW ≥ 4000 g had sensitivity 0.56 and specificity 0.91 for predicting birthweight ≥ 4000 g. The area under ROC curve for EFW within 4 days of delivery as a predictor of birthweight ≥ 4000 g was 0.85; for 5-22 days, 0.85; and for 0-22 days, 0.85. EFW ≥ 4192 g had sensitivity 0.83 and specificity 0.92 for predicting birthweight ≥ 4500 g. EFW ≥ 4500 g had sensitivity 0.22 and specificity 0.99 for predicting birthweight ≥ 4500 g. The area under ROC curve for EFW within 4 days of delivery as a predictor of birthweight ≥ 4500 g was 0.93; for 5-22 days, 0.95; and for 0-22 days, 0.95. The area under ROC curve for the adjusted EFW within 4 days of delivery as a predictor of birthweight ≥ 4500 g was 0.93; for 5-22 days, 0.95; and for 0-22 days, 0.95.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Phelan, Platt, Yeh, et al., 1984	Design: Case series (retrospective), no controls Test(s) studied: 1) Nonstress test (NST) (n = 239) Protocol: Last NST conducted within 7 days of delivery. NST considered reactive if ≥ 2 FHR accelerations of > 15 bpm, lasting 15 seconds, in a 20-min period. Reactive NSTs repeated in a week (or sooner if serum estriol was low). NST considered nonreactive if there were not 2 acceptable FHR accelerations in any 20-min period of observation totaling 40 minutes. If test nonreactive, then patient re-tested in afternoon. If afternoon test nonreactive, then CST performed (or, if CST contraindicated, then biophysical profile done). If CST negative, then repeated in 24 hours. Reference standard(s): 1) C-section for fetal distress 2) Meconium aspiration 3) Apgar score at 1 minute 4) Apgar score at 5 minutes 5) Macrosomia 6) Post-maturity syndrome Dates: July 1980 - June 1981 Location: Los Angeles, CA Setting: University hospital Type(s) of providers: General OB/GYN; specially trained antepartum nurses	No. of subjects at start: 239 Dropouts: NA (retrospective analysis) Loss to follow-up: NA No. of subjects at end: 239 Inclusion criteria: Post-dates (> 294 days); underwent NST within 7 days of delivery Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR; gestational age > 294 days required for inclusion in study Dating criteria: LMP Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Meconium staining 4) Meconium aspiration 5) Macrosomia (birthweight ≥ 4000 g) 6) Post-maturity syndrome 7) C-sections 8) 2 x 2 tables 9) Other test performance results	1) Apgar score < 7 at 1 minute: 47/239 (20%) 2) Apgar score < 7 at 5 minutes: 6/239 (3%) 3) Meconium staining: 99/239 (41%) 4) Meconium aspiration: 19/239 (8%) 5) Macrosomia (birthweight ≥ 4000 g): 52/239 (22%) 6) Post-maturity syndrome: 40/239 (17%) 7) C-sections: Overall: 42/239 (18%) For fetal distress: 13/239 (5%) 8) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress Screening test = Nonstress test (NST) C-section yes / no / Totals: NST nonreactive: 4 / 28 / 32 NST reactive: 9 / 198 / 207 Totals: 13 / 226 / 239 <u>2 x 2 table 2:</u> Reference standard = Meconium aspiration Screening test = NST Meconium aspiration yes / no / Totals: NST nonreactive: 5 / 27 / 32	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - Same patient population as Phelan, Platt, Yeh, et al. 1985, below.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None			NST reactive 14 193 207 Totals: 19 220 239 2 x 2 table 3: Reference standard = Apgar score at 1 minute Screening test = NST Apgar at 1 min < 7 ≥ 7 Totals: NST nonreactive 11 21 32 NST reactive 36 171 207 Totals: 47 192 239 2 x 2 table 4: Reference standard = Apgar score at 5 minutes Screening test = NST Apgar at 5 min < 7 ≥ 7 Totals: NST nonreactive 2 30 32 NST reactive 4 203 207 Totals: 6 233 239 2 x 2 table 5: Reference standard = Macrosomia (birthweight ≥ 4000 g) Screening test = NST Macrosomia yes no Totals: NST nonreactive 4 28 32 NST reactive 48 159 207 Totals: 52 187 239 *(continued on next page)*	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 6: Reference standard = Post-maturity syndrome Screening test = NST Post-maturity yes / no / Totals: NST nonreactive: 4 / 28 / 32 NST reactive: 36 / 171 / 207 Totals: 40 / 199 / 239 9) Other test performance results: Among patients with reactive NSTs, those with decelerations had significant increases in C-sections for fetal distress, meconium passage, and Apgar scores < 7 at 5 minutes compared to those without decelerations.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Phelan, Platt, Yeh, et al., 1985	Design: Case series (retrospective), no controls Test(s) studied: 1) Nonstress test (NST), biophysical profile, and amniotic fluid volume (AFV) Protocol: Testing schedule not described (though referenced). Patients with FHR bradycardia revealed on the NST were evaluated for delivery. AFV considered "adequate" if largest pocket > 1 cm in vertical diameter; "decreased" if largest pocket ≤ 1 cm; and "adequate, but decreased" if largest pocket > 1 cm, but overall impression of sonographer was that fluid was decreased. Reference standard(s): 1) C-section for fetal distress 2) Apgar score at 1 minute 3) Apgar score at 5 minutes 4) Birthweight Dates: July 1980 - June 1981 Location: Los Angeles, CA Setting: University hospital Type(s) of providers: General OB/GYN; specially trained antepartum nurses Length of follow-up: None	No. of subjects at start: 236 Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 236 Inclusion criteria: Post-dates; underwent biophysical testing within 7 days of delivery Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR Dating criteria: NR Parity: NR Bishop score: NR	1) Macrosomia (birthweight > 4000 g) 2) Apgar score < 7 at 1 minute 3) Apgar score < 7 at 5 minutes 4) Post-maturity syndrome 5) Meconium staining 6) Meconium aspiration 7) Deceleration or bradycardia 8) Fetal death 9) C-sections 10) 2 x 2 tables	1) Macrosomia (birthweight > 4000 g): 52/236 (22%) 2) Apgar score < 7 at 1 minute: 49/236 (21%) 3) Apgar score < 7 at 5 minutes: 8/236 (3%) 4) Post-maturity syndrome: 40/236 (17%) 5) Meconium staining: 99/236 (42%) 6) Meconium aspiration: 19/236 (8%) 7) Deceleration or bradycardia: 62/236 (26%) 8) Fetal death: 2/236 (< 1%) 9) C-sections: Overall: 45/236 (19%) For fetal distress: 13/236 (6%) 10) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress Screening test = Amniotic fluid volume (AFV) C-section yes / no / Totals: AFV decreased: 3 / 4 / 7 AFV adequate/ decreased: 6 / 32 / 38 AFV adequate: 4 / 187 / 191 Totals: 13 / 223 / 236	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: + Same patient population as in Phelan, Platt, Yeh, et al., 1984, above.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 2: Reference standard = Apgar score at 1 minute Screening test = AFV Apgar at 1 min < 7 / ≥ 7 / Totals: AFV decreased: 6 / 1 / 7 AFV adequate/ decreased: 12 / 26 / 38 AFV adequate: 31 / 160 / 191 Totals: 49 / 187 / 236 2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = AFV Apgar at 5 min < 7 / ≥ 7 / Totals: AFV decreased: 2 / 5 / 7 AFV adequate/ decreased: 1 / 37 / 38 AFV adequate: 5 / 186 / 191 Totals: 8 / 228 / 236 2 x 2 table 4: Reference standard = Birthweight Screening test = AFV Birthweight > 4000 g / ≤ 4000 g / Totals: AFV decreased: 0 / 7 / 7 AFV adequate/ decreased: 6 / 32 / 38 AFV adequate: 46 / 145 / 191 Totals: 52 / 184 / 236	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Pollack, Hauer-Pollack, and Divon, 1992	Design: Case series (retrospective), no controls Test(s) studied: 1) Ultrasound examination to estimate fetal weight Protocol: Exam performed within 1 week of delivery. Estimate of fetal weight based on biparietal diameter, abdominal circumference, and femur length. Reference standard(s): 1) Macrosomia (defined using two different thresholds) Dates: Jan 1989 - Sep 1990 Location: Bronx, NY Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 519 Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 519 Inclusion criteria: Gestational age ≥ 41 weeks; singleton pregnancy; U/S estimation of fetal weight within 1 week of delivery Exclusion criteria: Any complications of pregnancy Age: NR Race: NR Gestational age at entry: NR; gestational age ≥ 41 weeks required for inclusion in study Dating criteria: LMP and early U/S, when available; U/S dates preferred when there was a discrepancy of > 10 days between menstrual dates and U/S Parity: NR Bishop score: NR	1) 2 x 2 tables 2) Other test performance results	1) 2 x 2 tables: 2 x 2 table 1: Reference standard = Macrosomia (defined as birthweight > 4000 g) Screening test = Estimated fetal weight (EFW) Birthweight > 4000 g / ≤ 4000 g / Totals: EFW ≥ 4000 g: 67 / 36 / 103 EFW < 4000 g: 52 / 364 / 416 Totals: 119 / 400 / 519 2 x 2 table 2: Reference standard = Macrosomia (defined as birthweight > 4500 g) Screening test = EFW Birthweight > 4500 g / ≤ 4500 g / Totals: EFW ≥ 4500 g: 3 / 6 / 9 EFW < 4500 g: 18 / 492 / 510 Totals: 21 / 498 / 519 2) Other test performance results: *EFW > 4000 g as a predictor of macrosomia (> 4000 g):* Sensitivity: 0.56 Specificity: 0.91 Positive predictive value: 0.64 Negative predictive value: 0.87 *EFW > 4500 g as a predictor of macrosomia (> 4500 g):* Sensitivity: 0.15 Specificity: 0.99 Positive predictive value: 0.81 Negative predictive value: 0.80	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: - Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: Sample size: - Statistical tests: +

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Ramreker-singh-White, Farkas, Chard, et al., 1993	Design: Case series, no controls Test(s) studied: 1) Blood pressure, urine analysis, maternal weight, fetal movements, cardiotocography, and Doppler U/S velocimetry of utero-placental and umbilical blood flow (n = 167) Protocol: Above-mentioned tests performed twice weekly Reference standard(s): 1) Meconium staining Dates: 1991 Location: London, UK Setting: Unspecified hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 167 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 167 Inclusion criteria: Gestational age ≥ 280 days; uncomplicated pregnancy Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 280 days required for entry into study) Dating criteria: LMP and U/S at 16 weeks Parity: NR Bishop score: NR	1) Meconium staining 2) Fetal distress (defined as a cardiotocographic abnormality significant enough to lead to operative delivery) 3) Stillbirth 4) 2 x 2 table 5) Other test performance results	1) Meconium staining: 15/167 (9%) 2) Fetal distress: 16/167 (10%) 3) Stillbirth: 1/167 (< 1%) 4) 2 x 2 table: Reference standard = Meconium staining Screening test = Fetal distress (defined at left) |	Meconium yes	no	Totals:
---	---	---	---
Fetal distress	5	11	16
No fetal distress	10	141	151
Totals:	15	152	167	 5) Other test performance results: There were no differences in mean Doppler indices (resistance index for right and left arcuate arteries, resistance and pulsatility indices for umbilical artery) between the 16 women with fetal distress and the remaining 151 women. No quantitative data reported.	QUALITY SCORE: Reference standard: - Randomized: - Method of randomization: NA Verification bias: + Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: + Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: -

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Rayburn, Motley, Stempel, et al., 1982	Design: Cohort study (prospective) Test(s) studied: 1) Nonstress test (NST) + fetal movement charting + urine estrogen-to-creatinine ratio. Protocol: Above-mentioned tests performed semi-weekly or weekly. If NST reactive (≥ 2 adequate accelerations of baseline FHR during a 20- to 40-minute period), then repeated on the next visit. If NST nonreactive, then test either repeated or a CST given the same day. Reference standard(s): 1) Post-maturity syndrome Dates: July 1979 - Apr 1981 Location: Columbus, OH Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 147 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 147 Inclusion criteria: Gestational age ≥ 42 weeks; scheduled to undergo NST Exclusion criteria: None specified Age: 20% ≤ 19; 69% 20-29; 11% ≥ 30 Race: 66% White, 34% Black Gestational age at entry: 42-43 weeks: 69% 43-44 weeks: 22% ≥ 44 weeks: 9% Dating criteria: LMP + either physical exam before 12th week or U/S before 20th week Parity: 46% primiparous Bishop score: NR Other: Cervical dilation: > 2 cm: 24% ≤ 2 cm: 76%	1) Post-maturity syndrome 2) Admission to NICU 3) Meconium aspiration 4) Birth asphyxia 5) Death 6) 2 x 2 tables 7) Other test performance results	1) Post-maturity syndrome: 32/147 (22%) 2) Admission to NICU: 7/147 (5%) 3) Meconium aspiration: 3/147 (2%) 4) Birth asphyxia: 1/147 (1%) 5) Death: 1/147 (1%) 6) 2 x 2 tables: 2 x 2 table 1: Reference standard = Post-maturity syndrome Screening test = Antepartum FHR monitoring |	Post-maturity yes	no	Totals:
---	---	---	---
FHR abnormal	3	0	3
FHR normal	29	115	144
Totals:	32	115	147 | 2 x 2 table 2: Reference standard = Post-maturity syndrome Screening test = Urine estrogen-to-creatinine ratio (E:C) |	Post-maturity yes	no	Totals:
---	---	---	---
E:C subnormal	12	0	12
E:C normal	3	50	53
Totals:	15	50	65	 2 x 2 table 3: Reference standard = Post-maturity syndrome Screening test = Fetal movement (FM) charting	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: + Placenta grading was not possible in 70/147 cases (48%) because ultrasonic visualization was too poor.

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes

	Post-maturity yes	no	Totals:
FM inactive	0	0	0
FM active	32	115	147
Totals:	32	115	147

2 x 2 table 4:
Reference standard = Post-maturity syndrome
Screening test = Amniotic fluid volume (AFV)

	Post-maturity yes	no	Totals:
AFV adequate	24	5	29
AFV pockets	5	48	53
Oligo-hydramnios	3	62	65
Totals:	32	115	147

2 x 2 table 5:
Reference standard = Post-maturity syndrome
Screening test = Fetal motion (FM) on U/S

	Post-maturity yes	no	Totals:
FM absent	11	6	17
FM present	21	109	130
Totals:	32	115	147

2 x 2 table 6:
Reference standard = Post-maturity syndrome
Screening test = Fetal breathing (FB) on U/S

	Post-maturity yes	no	Totals:
FB absent	15	32	47
FB present	17	83	100
Totals:	32	115	147

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				7) Other test performance results: For predicting post-maturity syndrome:	

Test	Sensitivity	Specificity
Oligo-hydramnios	75%	96%
FHR testing	9%	100%
Fetal movement charting	0%	100%
E:C ratio	80%	100%
Grade 3 placenta	100%	11%
Gross fetal body motion	34%	95%
Fetal breathing	47%	72%

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sarkar and Duthie, 1997	Design: Cohort study (retrospective) Test(s) studied: 1) Cardiotocography and amniotic fluid index (AFI) (n = 184) Protocol: Cardiotocography and AFI performed twice weekly. Protocol not specified; presumably if AFI reduced, labor induced and continuous FHR monitoring used. Reference standard(s): 1) Birthweight 2) Apgar score at 5 minutes 3) Intubation 4) Admission to NICU 5) Emergency C-section Dates: Jan 1993 - Dec 1994 Location: Chester, UK Setting: Unspecified hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 184 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 184 Inclusion criteria: Gestational age ≥ 42 completed weeks; uncomplicated singleton pregnancy Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 42 weeks required for entry into study) Dating criteria: LMP and U/S dates within 10 days of one another Parity: NR Bishop score: NR	1) Meconium staining 2) Low birthweight (< 5th percentile) 3) Apgar score < 7 at 5 minutes 4) Intubation 5) Admission to NICU 6) Abnormal FHR tracings 7) Emergency C-section (for fetal distress) 8) 2 x 2 tables	1) Meconium staining: 18/184 (9.8%) 2) Low birthweight (< 5th percentile): 2/184 (1%) 3) Apgar score < 7 at 5 minutes: 9/184 (4.9%) 4) Intubation: 5/184 (2.7%) 5) Admission to NICU: 1/184 (0.5%) 6) Abnormal FHR tracings: 47/184 (25.5%) 7) Emergency C-section (for fetal distress): 36/184 (19.6%) 8) 2 x 2 tables: 2 x 2 table 1: Reference standard = Birthweight ("low" defined as < 5th percentile) Screening test = AFI Birthweight low / normal / Totals: AFI decreased: 2 / 16 / 18 AFI normal: 0 / 166 / 166 Totals: 2 / 182 / 184 2 x 2 table 2: Reference standard = Apgar score at 5 minutes Screening test = AFI Apgar at 5 min < 7 / ≥ 7 / Totals: AFI decreased: 0 / 18 / 18 AFI normal: 9 / 157 / 166 Totals: 9 / 175 / 184	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: + Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 3: Reference standard = Intubation Screening test = AFI Intubation yes / no / Totals: AFI decreased: 0 / 18 / 18 AFI normal: 5 / 161 / 166 Totals: 5 / 179 / 184 2 x 2 table 4: Reference standard = Admission to NICU Screening test = AFI Admission to NICU yes / no / Totals: AFI decreased: 1 / 17 / 18 AFI normal: 0 / 166 / 166 Totals: 1 / 183 / 184 2 x 2 table 5: Reference standard = Emergency C-section Screening test = AFI C-section yes / no / Totals: AFI decreased: 6 / 12 / 18 AFI normal: 30 / 136 / 166 Totals: 36 / 148 / 184	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Schreyer, Bar-Natan, Sherman, et al., 1991	Design: Case series (prospective), no controls Test(s) studied: 1) Fetal breathing movements (n = 65) Protocol: Fetal breathing movements were measured by U/S immediately before elective induction for reactive NST. Fetal breathing was considered to be present (+) when sustained for ≥ 20 seconds, and absent (-) when no sustained movement could be detected over a 45-minute period. Bishop score was assessed. Patients with Bishop score 0-2 were eliminated from the study and treated expectantly or by intracervical PGE_2 gel application. Labor was induced with oxytocin at 2 mIU/min, increasing by 1 mIU/min every 30 minutes until 3 contractions per 10 minutes. When cervix effaced and dilated 2-3 cm, membranes were artificially ruptured and internal cardiotocography initiated. Reference standard(s): 1) C-section for fetal distress 2) Apgar score at 5 minutes 3) Macrosomia Dates: June 1988 - June 1989 Location: Tel Aviv, Israel Setting: University hospital	No. of subjects at start: NR Dropouts: NR Loss to follow-up: NA No. of subjects at end: 65 Inclusion criteria: Gestational age 287-294 days Exclusion criteria: Pregnancy-induced hypertension; diabetes mellitus; previous C-section; IUGR; estimated fetal weight > 4300 g; malpresentation Age (mean): 27.0 Race: NR Gestational age at entry (mean): 291.4 days Dating criteria: LMP and either a) 1st trimester U/S or b) two 2nd trimester U/S Parity: 29% primiparous Bishop score: 41.5% > 6; 58.5% 3-6	1) Apgar score < 7 at 5 minutes 2) Macrosomia (> 4000 g) 3) C-sections 4) 2 x 2 tables 5) Other test performance results	1) Apgar score < 7 at 5 minutes: 1/65 (1.5%) 2) Macrosomia (> 4000 g): 10/65 (15.4%) 3) C-sections: Overall: 4/65 (6.2%) For fetal distress: 1/65 (1.5%) 4) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress (not defined) Screening test = Fetal breathing movements (FBM) by U/S C-section yes / no / Totals: FBM - 0 / 24 / 24 FBM + 1 / 40 / 41 Totals: 1 / 64 / 65 <u>2 x 2 table 2:</u> Reference standard = Apgar score at 5 minutes Screening test = FBM Apgar at 5 min < 7 / ≥ 7 / Totals: FBM - 0 / 24 / 24 FBM + 1 / 40 / 41 Totals: 1 / 64 / 65 <u>2 x 2 table 3:</u> Reference standard = Macrosomia (birthweight > 4000 g) Screening test = FBM Macrosomia yes / no / Totals: FBM - 4 / 20 / 24 FBM + 6 / 35 / 41 Totals: 10 / 55 / 65	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: NR Length of follow-up: None			5) Other test performance results: Presence of fetal breath movements (FBM+) was associated with: a) No difference in birthweight (3608 ± 671 g vs. 3719 ± 710 g; p = not significant) b) Longer total induction time (648.5 ± 354 min vs. 319.3 ± 137 min; p < 0.001) c) Higher oxytocin requirement (2708 ± 1727 mIU vs. 1134 ± 709 mIU; p < 0.001)	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sherer, Onyeije, Binder, et al., 1998	Design: Cohort/nested case-control study (retrospective) Test(s) studied: 1) FHR assessed for baseline fetal tachycardia (≥ 160 bpm) or bradycardia (≤ 120 bpm) Protocol: Baseline FHR assessed at post-term evaluation. Reference standard(s): 1) Apgar score at 5 minutes 2) Meconium aspiration 3) Admission to NICU Dates: July 1985 - June 1995 Location: Bronx, NY Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 107 cases and 283 controls: 31 patients with baseline tachycardia, plus 66 matched controls; 76 patients with baseline brady-cardia, plus 217 matched controls Dropouts: NA (retrospective study) Loss to follow-up: NA No. of subjects at end: 107 cases and 283 controls Inclusion criteria: Singleton pregnancy; gestational age ≥ 41 weeks; not in labor; afebrile; normal fetal anatomy; reactive NST; intact membranes; no evidence of chorioamnionitis Exclusion criteria: Fetal tachy- or brady-arrhythmias; FHR decelerations; loss of short-term beat-to-beat variability Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 41 weeks required for entry into study) Dating criteria: LMP and U/S before 20 weeks Parity: NR Bishop score: NR	1) Apgar score < 7 at 5 minutes 2) Meconium staining 3) Meconium aspiration 4) Admission to NICU 5) Fetal growth restriction (< 10th percentile for gestational age) 6) C-sections 7) 2 x 2 tables	1) Apgar score < 7 at 5 minutes: Tachycardia: 7/31 (23%) Matched controls: 10/66 (15%) p = 0.369 Bradycardia: 14/76 (18%) Matched controls: 54/217 (25%) p = 0.25 2) Meconium staining: Tachycardia: 13/31 (42%) Matched controls: 28/66 (42%) p = 0.964 Bradycardia: 26/76 (34%) Matched controls: 63/217 (29%) p = 0.398 3) Meconium aspiration: Tachycardia: 2/31 (7%) Matched controls: 1/66 (2%) p = 0.190 Bradycardia: 4/76 (5%) Matched controls: 12/217 (6%) p = 0.929 4) Admission to NICU: Tachycardia: 2/31 (7%) Matched controls: 3/66 (5%) p = 0.692 Bradycardia: 11/76 (15%) Matched controls: 20/217 (9%) p = 0.199 5) Fetal growth restriction: Tachycardia: 1/31 (3%) Matched controls: 10/66 (15%) p = 0.084 Bradycardia: 8/76 (11%) Matched controls: 15/217 (7%) p = 0.313	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: - *(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				6) C-sections: Tachycardia: 8/31 (25%) Matched controls: 11/66 (29%) p = 0.29 Bradycardia: 11/76 (15%) Matched controls: 52 217 (24%) p = 0.083 7) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = Apgar score at 5 minutes Screening test = Baseline bradycardia (BB) (≤ 120 bpm) Apgar at 5 min < 7 / ≥ 7 / Totals: BB yes: 14 / 62 / 76 BB no: 54 / 163 / 217 Totals: 68 / 225 / 293 <u>2 x 2 table 2:</u> Reference standard = Meconium aspiration Screening test = BB Meconium aspiration yes / no / Totals: BB yes: 4 / 72 / 76 BB no: 12 / 205 / 217 Totals: 16 / 277 / 293 <u>2 x 2 table 3:</u> Reference standard = Admission to NICU Screening test = BB NICU admission yes / no / Totals: BB yes: 11 / 65 / 76 BB no: 20 / 197 / 217 Totals: 31 / 262 / 293	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 4: Reference standard = Apgar score at 5 minutes Screening test = Baseline tachycardia (BT) (≥ 160 bpm) Apgar at 5 min < 7 / ≥ 7 / Totals: BT yes: 7 / 24 / 31 BT no: 10 / 56 / 66 Totals: 17 / 80 / 97 2 x 2 table 5: Reference standard = Meconium aspiration Screening test = BT Meconium aspiration yes / no / Totals: BT yes: 2 / 29 / 31 BT no: 1 / 65 / 66 Totals: 3 / 94 / 97 2 x 2 table 6: Reference standard = Admission to NICU Screening test = BT NICU admission yes / no / Totals: BT yes: 2 / 29 / 31 BT no: 3 / 63 / 66 Totals: 5 / 92 / 97	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Small, Phelan, Smith, et al., 1987	Design: Case series (retrospective), historical controls Test(s) studied: 1) Nonstress test (n = 470) Protocol: For patients with good dates (U/S before 28 weeks or multiple 1st and 2nd trimester exams), NST performed twice weekly. If cervix favorable (Bishop score ≥ 9), then labor induced. For patients with unreliable dates (LMP only), NST performed weekly. NST considered reactive whenever ≥ 2 FHR accelerations observed within 10 minutes. Accelerations had to rise 15 bpm and last 15 seconds. Labor induced for FHR deceleration of any type; persistent nonreactive NST; oligohydramnios (< 1 cm) on U/S; positive contraction stress test (CST); or biophysical profile score ≤ 4. Reference standard(s): 1) C-section for fetal distress 2) Apgar score at 1 minute 3) Apgar score at 5 minutes 4) Macrosomia 5) Post-maturity Dates: Jan - Dec 1984 (study group); 1980 (controls) Location: Los Angeles, CA Setting: University hospital Type(s) of providers: General OB/GYN	No. of subjects at start: 476 cases (met inclusion criteria); 239 historical controls Dropouts: 6 cases (excluded due to incomplete delivery information) Loss to follow-up: NA No. of subjects at end: 470 cases; 239 historical controls Inclusion criteria: Gestational age > 294 days/42 weeks; antepartum FHR testing within 7 days of delivery Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry: NR (gestational age > 42 weeks required for entry into study) Dating criteria: LMP Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Meconium staining 4) Macrosomia (> 4000 g) 5) Post-maturity 6) Perinatal death 7) C-sections 8) 2 x 2 tables 9) Comparisons with historical controls	1) Apgar score < 7 at 1 minute: 86/470 (18%) 2) Apgar score < 7 at 5 minutes: 9/470 (2%) 3) Meconium staining: 126/470 (27%) 4) Macrosomia: 98/470 (21%) 5) Post-maturity: 32/470 (7%) 6) Perinatal death: 3/470 (< 1%) 7) C-sections: Overall: 79/470 (17%) For fetal distress: 19/470 (4%) 8) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = C-section for fetal distress (not defined) Screening test = Nonstress test (NST) ("reactive" whenever ≥ 2 FHR accelerations observed within 10 minutes; accelerations had to rise 15 bpm and last 15 seconds) C-section for fetal distress yes / no / Totals: NST nonreactive: 4 / 46 / 50 NST reactive: 15 / 405 / 420 Totals: 19 / 451 / 470 <u>2 x 2 table 2:</u> Reference standard = Apgar score at 1 minute Screening test = NST (as above)	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: - Dating criteria: + Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: - Sample size: - Statistical tests: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None			Apgar at 1 min < 7 / ≥ 7 / Totals: NST nonreactive 11 39 50 NST reactive 75 345 420 Totals: 86 384 470 2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = NST (as above) Apgar at 5 min < 7 / ≥ 7 / Totals: NST nonreactive 1 49 50 NST reactive 8 412 420 Totals: 9 461 470 2 x 2 table 4: Reference standard = Macrosomia (birthweight > 4000 g) Screening test = NST (as above) Macrosomia yes / no / Totals: NST nonreactive 5 45 50 NST reactive 93 327 420 Totals: 98 372 470 2 x 2 table 5: Reference standard = Post-maturity Screening test = NST (as above) Post-maturity yes / no / Totals: NST nonreactive 4 46 50 NST reactive 28 392 420 Totals: 32 438 470 9) Comparisons with historical controls: Compared to controls from 1980 (n = 239), post-dates patients from 1984	*(continued on next page)*

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				(n = 470) were significantly less likely to have meconium. Compared to controls, 1984 post-dates patients with reactive NST and decelerations were significantly less likely to have C-section for fetal distress, meconium, or birthweight > 4000 g. (Reactive NST with decelerations included among criteria for induction in 1984.)	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Tam, Tai, and Rogers, 1999	Design: Cohort study (not specified if prospective) Test(s) studied: 1) Fetal fibronectin (fFN) testing, followed by induction using PGE_2 pessaries (n = 58) Protocol: Cervico-vaginal secretion tested for presence of fetal fibronectin prior to cervical ripening/induction. Labor induced with PGE_2 pessary (3 mg). Cervical status reassessed 4-6 hours later. If Bishop score < 5, then second dose given. If Bishop score ≥ 5, then artificial rupture of membranes performed. Oxytocin begun at 2.5 mU/min of 1 mU/min for nulliparous and multiparous women, respectively, with dose increased every 15 minutes. Reference standard(s): 1) C-section Dates: Apr 1996 - Feb 1997 Location: Hong Kong Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 58 (30 negative for fetal fibronectin [fFN-]; 28 positive [fFN+]) Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 58 Inclusion criteria: Term or post-term pregnancy; documented indication for induction Exclusion criteria: Bishop score ≥ 5; ruptured membranes Age (mean, with range): fFN-: 30 (27-33) fFN+: 28 (24-34) Race: NR Gestational age at entry (median, with range): fFN-: 281 days (272-294) fFN+: 294 days (280-294) p = 0.10 Dating criteria: NR Parity (median, with range): fFN-: 1 (0-1) fFN+: 1 (0-2) Bishop score (median, with range): fFN-: 3 (1-4) fFN+: 3 (1-4)	1) 2 x 2 table 2) Interval from induction to delivery	1) 2 x 2 table: Reference standard = C-section Screening test = Fetal fibronectin (fFN) status C-section yes / no / Totals: fFN+: 3 / 16 / 19 fFN-: 11 / 28 / 39 Totals: 14 / 44 / 58 2) Interval from induction to delivery (median, with range): fFN+: 760 minutes (540-1375) fFN-: 1285 minutes (692-2266) p = 0.04	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: - Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: + Results not stratified by indication for induction.

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Tongsong and Srisom-boon, 1993	Design: Cohort study (prospective) Test(s) studied: 1) Nonstress test (NST) + amniotic fluid volume (AFV) (n = 252) Protocol: Above-mentioned tests performed twice weekly. If NST or AFV abnormal, then contraction stress test (CST) performed. If CST negative, then patient re-tested in 3-4 days (uncertain if repeat test was NST+AFV or repeat CST). If CST positive, then labor induced. If cervix favorable, then labor induced. Reference standard(s): 1) Fetal distress/obstetric intervention 2) Apgar score at 1 minute 3) Apgar score at 5 minutes Dates: June 1989 - May 1992 Location: Chiang Mai, Thailand Setting: Unspecified hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 252 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 252 Inclusion criteria: Singleton pregnancy; attended antenatal clinic in 1st trimester; delivery after 42 weeks' gestation Exclusion criteria: Any medical or obstetric complication; congenital abnormalities of fetus Age: NR Race: NR Gestational age at entry: NR (delivery after 42 weeks required for entry into study) Dating criteria: LMP + 1st trimester clinical exam Parity: NR Bishop score: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Fetal distress 4) Meconium staining 5) Obstetric intervention for fetal distress 6) 2 x 2 tables 7) Other test performance results	1) Apgar score < 7 at 1 minute: 17/252 (7%) 2) Apgar score < 7 at 5 minutes: 6/252 (2%) 3) Fetal distress: 11/252 (4%) 4) Meconium staining: 87/252 (35%) 5) Obstetric intervention for fetal distress: 11/252 (4%) 6) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = Obstetric intervention for fetal distress (defined as repetitive late decelerations, repetitive moderate to severe variable decelera-tions, or prolonged bradycardia) Screening test = Amniotic fluid volume (AFV) ("abnormal" if largest vertical pocket < 3 cm) Fetal distress yes / no / Totals: AFV abnormal: 8 / 22 / 30 AFV normal: 3 / 219 / 222 Totals: 11 / 241 / 252 <u>2 x 2 table 2:</u> Reference standard = Apgar score at 1 minute Screening test = AFV (as above) Apgar at 1 min < 7 / ≥ 7 / Totals: AFV abnormal: 8 / 22 / 30 AFV normal: 9 / 213 / 222 Totals: 17 / 235 / 252	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: + Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				2 x 2 table 3: Reference standard = Apgar score at 5 minutes Screening test = AFV (as above) Apgar at 5 min < 7 / ≥ 7 / Totals: AFV abnormal 2 28 30 AFV normal 4 218 222 Totals: 6 246 252 2 x 2 table 4: Reference standard = Obstetric intervention for fetal distress (as above) Screening test = Nonstress test (NST) (“abnormal” if nonreactive or reactive with variable or late decelerations) Fetal distress yes / no / Totals: NST abnormal 7 44 51 NST normal 4 197 201 Totals: 11 241 252 2 x 2 table 5: Reference standard = Apgar score at 1 minute Screening test = NST (as above) Apgar at 1 min < 7 / ≥ 7 / Totals: NST abnormal 7 44 51 NST normal 10 191 201 Totals: 17 235 252 2 x 2 table 6: Reference standard = Apgar score at 5 minutes Screening test = NST (as above)	

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Apgar at 5 min < 7 / ≥ 7 / Totals: NST abnormal: 2 / 49 / 51 NST normal: 4 / 197 / 201 Totals: 6 / 246 / 252 7) Other test performance results: AFV was more accurate than NST in predicting intrapartum fetal distress ($p < 0.05$). AFV sensitivity, 0.73; specificity, 0.91; positive predictive value, 0.27; negative predictive value, 0.99. NST sensitivity, 0.64; specificity, 0.82; positive predictive value, 0.14; negative predictive value, 0.98.	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Weiner, Farmakides, Schulman, et al., 1994	Design: Cohort study (prospective) Test(s) studied: 1) Nonstress test (NST) with computerized analysis of fetal heart rate (FHR) variation + Doppler examination of umbilical artery + biophysical profile (n = 337) Protocol: Above-mentioned tests performed every 2-4 days beginning at 41 weeks. Labor induced (using oxytocin infusion and amniotomy) if FHR variation reduced (< 30 msec), FHR decelerations appeared, or amniotic fluid index (AFI) ≤ 5, and after 42 weeks if Bishop score > 7. Reference standard(s): 1) Fetal distress 2) Acidosis 3) Neonatal death Dates: June 1991 - May 1993 Location: Mineola, NY Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 337 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 337 Inclusion criteria: Delivery at > 41 weeks' gestation; uncomplicated pregnancy Exclusion criteria: None specified Age (mean ± SD): 29 ± 4.6 Race: NR Gestational age at entry: NR (delivery at > 41 weeks required for entry into study) Dating criteria: U/S before 22 weeks Parity: NR Bishop score: NR	1) Fetal distress 2) Acidosis 3) Neonatal death 4) C-sections 5) 2 x 2 tables 6) Other test performance results	1) Fetal distress: 37/337 (11%) 2) Acidosis: 10/337 (3%) 3) Neonatal death: 2/337 (0.6%) 4) C-sections: Overall: 101/337 (30%) For fetal distress: 33/337 (10%) 5) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = Fetal distress (definition included presence of FHR late decelerations, severe FHR variable decelerations, and reduced beat-to-beat variability) Screening test = FHR variation Fetal distress yes / no / Totals: FHR variation < 30 msec: 11 / 1 / 12 FHR variation ≥ 30 msec: 28 / 297 / 325 Totals: 39 / 298 / 337 <u>2 x 2 table 2:</u> Reference standard = Fetal acidosis (umbilical artery pH < 7.2) Screening test = FHR variation Acidosis yes / no / Totals: FHR variation < 30 msec: 7 / 5 / 12 FHR variation ≥ 30 msec: 3 / 322 / 325 Totals: 10 / 327 / 337 <u>2 x 2 table 3:</u> Reference standard = Neonatal death Screening test = FHR variation	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: + Statistical tests: +

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Neonatal death yes / no / Totals: FHR variation < 30 msec: 1 / 11 / 12 FHR variation ≥ 30 msec: 1 / 324 / 325 Totals: 2 / 335 / 337 6) Other test performance results: For predicting intrapartum fetal distress and acidosis at delivery, FHR variations showed higher area under the ROC curve than did amniotic fluid index or umbilical S:D ratio. Nonreactive NST and presence of decelerations were also predictive of distress in labor, and decelerations were predictive of acidosis at delivery ($p < 0.001$).	

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Weiner, Reichler, Zlozover, et al., 1993	Design: Cohort study (prospective) Test(s) studied: 1) Nonstress test (NST) + amniotic fluid volume (AFV) + Doppler velocimetry of umbilical and uterine arteries (n = 142) Protocol: Above-mentioned tests performed every 3 days. Labor induced if abnormal NST, oligohydramnios, or favorable cervix (Bishop score > 7) after 42 weeks gestation. Reference standard(s): 1) Fetal outcome Dates: NR; data collected over a 1-year period Location: Haifa, Israel Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None	No. of subjects at start: 142 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 142 Inclusion criteria: Gestational age > 287 days Exclusion criteria: Pregnancy complications (e.g., hypertension, gestational diabetes) Age (mean ± SD): 27.3 ± 5.6 Race: NR Gestational age at entry: NR (gestational age > 287 days required for entry into study) Dating criteria: "Early fetal biometry" Parity: 31% primiparous Bishop score: NR	1) Apgar score < 7 at 5 minutes 2) Admission to NICU 3) C-sections 4) 2 x 2 tables 5) Other test performance results	1) Apgar score < 7 at 5 minutes: 2/142 (1.4%) 2) Admission to NICU: 1/142 (0.7%) 3) C-sections: Overall: 13/142 (9.2%) For fetal distress: 7/142 (4.9%) 4) 2 x 2 tables: <u>2 x 2 table 1:</u> Reference standard = Fetal outcome ("abnormal" defined as 5-minute Apgar score < 7, admission to NICU, C-section for fetal distress, or birthweight < 5th percentile) Screening test = NST Fetal outcome abn / nl / Totals: NST abnormal: 1 / 6 / 7 NST normal: 11 / 124 / 135 Totals: 12 / 130 / 142 <u>2 x 2 table 2:</u> Reference standard = Fetal outcome (as above) Screening test = AFV ("low" defined as < 5 cm) Fetal outcome abn / nl / Totals: AFV low: 3 / 8 / 11 AFV normal: 9 / 122 / 131 Totals: 12 / 130 / 142 <u>2 x 2 table 3:</u> Reference standard = Fetal outcome (as above) Screening test = NST, AFV, and umbilical and uterine artery resistance index	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: + Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: + Testing protocol described: + Sample size: - Statistical tests: -

(continued on next page)

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes

	Fetal outcome		
	abn	nl	Totals:
At least one test abn	8	18	26
All tests normal	4	112	116
Totals:	12	130	142

5) Other test performance results:
For predicting abnormal fetal outcome (as defined above), screening tests had the following performance characteristics:

	Sensitivity	Specificity
NST	0.08	0.95
AFV	0.25	0.94
Resistance index	0.17	0.96
Any test abnormal	0.67	0.88

Evidence Table 1: Studies relevant to Key Question 1 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Witter and Weitz, 1989	Design: Case series (prospective), no controls Test(s) studied: 1) Cervical exam + induction by oxytocin infusion and amniotomy (n = 76) Protocol: At 42 completed weeks, cervical exam performed prior to induction of labor. Oxytocin infusion started at 7:00 AM with 1 mU/min and increased by 1 mU/min every 10 min until a dose of 30 mU/min reached or a regular pattern of adequate uterine contractions established. Amniotomy performed as soon as possible, but always after oxytocin had established regular contractions. If patient had intact membranes and was not in active phase labor by evening, the induction was stopped and the patient was rested overnight. The induction was restarted in the morning. If the patient failed to enter the active phase of labor by 20 hours of induction, then C-section performed. Reference standard(s): 1) C-section Dates: NR Location: Baltimore, MD Setting: University hospital Type(s) of providers: MFM Length of follow-up: None	No. of subjects at start: 103 (see Notes) Dropouts: 27 (did not have cervical exam) Loss to follow-up: NA No. of subjects at end: 76 Inclusion criteria: Gestational age ≥ 42 weeks Exclusion criteria: Previous C-section Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 42 weeks required for entry into study) Dating criteria: 2 or more of the following: certain LMP; basal body temperature indicating ovulation temperature shift for the present pregnancy; positive urinary pregnancy test at 6 weeks from LMP; fetal heart tones heard with DeLee stethoscope at 18-20 weeks; fundal height at the umbilicus at 20 weeks; fundal height in cm equal to gestational age in weeks within 2 cm from 20-34 weeks; early registration with dates equal to exam prior to 13 weeks; U/S dating by crown-rump length between 6 and 14 weeks or by biparietal diameter prior to 26 weeks Parity: NR Bishop score: NR	1) C-sections 2) 2 x 2 tables	1) C-sections: 26/76 (34%) 2) 2 x 2 tables: 2 x 2 table 1: Reference standard = C-section Screening test = Cervical dilation C-section yes / no / Totals: Dilation 0 cm: 20 / 11 / 31 Dilation > 0 cm: 6 / 39 / 45 Totals: 26 / 50 / 76 2 x 2 table 2: Reference standard = C-section Screening test = Cervical effacement C-section yes / no / Totals: Effacement 0%: 12 / 6 / 18 Effacement > 0%: 14 / 44 / 58 Totals: 26 / 50 / 76 2 x 2 table 3: Reference standard = C-section Screening test = Cervical station C-section yes / no / Totals: Station ≥ -3: 19 / 22 / 41 Station < -3: 7 / 28 / 35 Totals: 26 / 50 / 76	QUALITY SCORE: Reference standard: + Randomized: - Method of randomization: NA Verification bias: - Test reliability/variability: - Gestational age: + Dating criteria: + Other risk factors absent: - Similar to likely pt pop: - Testing protocol described: + Sample size: - Statistical tests: + Study population was subgroup (76/103) of patients randomized to induction in Witter and Weitz, 1987 (see Evidence Table: Key Question 3).

Evidence Table 2: Studies relevant to Key Question 2

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Augensen, Bergsjø, Eikeland, et al., 1987	Design: RCT, randomization by random numbers list Interventions: 1) Immediate induction at time of referral/admission into study (n = 214) Protocol: 5 IU oxytocin given intravenously, with dose rates increased stepwise according to response. Amniotomy performed once labor established or, in exceptional cases, at the start of induction. If no labor after 6-8 hours, then induction considered unsuccessful, and patient managed according to postponed induction protocol. 2) Delayed induction after monitoring for 1 wk (n = 195) Protocol: NST on day of referral/admission into study and again on day 3 or 4 if still undelivered. If birth had not occurred by day 7, then labor induced as above. If this induction attempt failed, then management "left to clinical judgement." Dates: Jan 1982 - June 1985 Location: Bergen, Norway Setting: University hospital Type(s) of providers: NR Length of follow-up: NA	No. of subjects at start: 409 Dropouts: 0 (see notes) Loss to follow-up: NA No. of subjects at end: 409 Inclusion criteria: Healthy women with normal pregnancies; singleton fetus; cephalic presentation; gestational age 290-297 days; reliable dates Exclusion criteria: Use of OCPs during two months before LMP; hypertension; IUGR; other medical conditions; geographical and social considerations (not specified) Age: NR Race: NR Gestational age at entry: NS; gestational age of 290-297 days required for entry into study Dating criteria: LMP ("clear recollection") Parity: Immediate induction, 46% nulliparous; delayed induction, 42% Bishop score: Immediate induction, 36% < 6; delayed induction, 35% < 6	1) Meconium 2) Admission to NICU 3) Length of stay in NICU 4) Hyperbilirubinemia 5) Difficult shoulder delivery 6) C-sections 7) Number of days in hospital 8) Courses of induction	1) Meconium: Immediate: 37/214 (17%) Delayed: 32/195 (16%) (no p-value reported) 2) Admission to NICU: Immediate: 12/214 (5.6%) Delayed: 15/195 (7.7%) (no p-value reported) 3) Length of stay in NICU (mean): Immediate: 4.3 days Delayed: 9.7 days (one patient stayed in NICU 93 days) (no p-value reported) 4) Hyperbilirubinemia: Immediate: 10/214 (4.7%) Delayed: 1/195 (0.51%) 0.01 > p > 0.005 5) Difficult shoulder delivery: Immediate: 1/214 (0.5%) Delayed: 0/195 (no p-value reported) 6) C-sections: Immediate: 14/214 (6.5%) Delayed: 15/195 (7.7%) (no p-value reported) 7) Number of days in hospital (mean ± SD): Immediate: 7.05 ± 1.67 days Delayed: 6.69 ± 1.37 days p = 0.02 8) Courses of induction (mean): Immediate: 1.09 Delayed: 0.34 (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Four patients randomized to immediate induction delivered spontaneously before being induced; these patients were included in the analysis. Results not stratified by parity.

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Bergsø, Huang, Yu, et al., 1989	Design: RCT, randomization by list of random numbers Interventions: 1) Induction (n = 94) Protocol: Labor induced at or shortly after 42 weeks by stripping of the membranes, followed by oxytocin infusion (5 IU in 500 ml solution). Infusion rate regulated according to response. Membranes ruptured artificially if cervix dilated ≥ 3 cm. 2) Monitoring (n = 94) Protocol: Patients admitted to hospital to undergo "close daily clinical surveillance." Fetal movement tests, atropine tests, U/S, and urinary estriol excretion tests also employed. Labor induced as above at ≥ 43 weeks "according to clinical judgement." Dates: July 1982 - sometime in 1984 Location: Wuhan, China Setting: Community hospital which also serves as regional referral center for high-risk obstetrics Type(s) of providers: Unspecified OB/GYN Length of follow-up: NA	No. of subjects at start: 188 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 188 Inclusion criteria: Gestational age ≥ 42 weeks (294 days); not in labor; intact membranes; normal pregnancy without significant risk factors; normal menstrual cycle (28 ± 4 days) with accurate recall of LMP Exclusion criteria: None specified Age (mean): Induction, 26.1; monitoring, 27.8 Race: 100% Chinese Gestational age at entry: NR; gestational age of ≥ 42 weeks required for entry into study Dating criteria: LMP Parity: Induction, 6% nulliparous; monitoring, 13% nulliparous Bishop score: NR	1) Apgar scores 2) Fetal distress 3) Hyperbilirubinemia 4) Respiratory distress syndrome 5) Aspiration pneumonia 6) Total operative deliveries (C-sections, forceps-assisted deliveries, and vacuum extractions) 7) C-sections 8) Forceps-assisted deliveries 9) Vacuum extractions 10) Length of hospital stay	1) Apgar scores: No quantitative data reported. Authors stated only that "Apgar score distributions were almost equal between the groups." 2) Fetal distress (not defined): Induction: 17/94 (18.1%) Monitoring: 18/94 (19.1%) p = not significant 3) Hyperbilirubinemia: Induction: 6/94 (6.4%) Monitoring: 3/94 (3.2%) p = not significant 4) Respiratory distress syndrome: Induction: 4/94 (4.3%) Monitoring: 8/94 (8.5%) p = not significant 5) Aspiration pneumonia: Induction: 4/94 (4.3%) Monitoring: 8/94 (8.5%) p = not significant 6) Total operative deliveries: Induction: 48/94 (51.1%) Monitoring: 64/94 (68.1%) p < 0.05 7) C-sections: Induction: 27/94 (28.7%) Monitoring: 39/94 (41.5 %) p = not significant 8) Forceps-assisted deliveries: Induction: 9/94 (9.6%) Monitoring: 11/94 (11.7%) p = not significant 9) Vacuum extractions: Induction: 12/94 (12.8%) Monitoring: 14/94 (14.9%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: - Mode of delivery: - Sample size: - Statistical tests: ?? Gestational age: + Dating criteria: + Bishop score: - Results not stratified by parity. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				p = not significant 10) Length of hospital stay (mean, with range) Induction: 7.9 days (1-28) Monitoring: 8.1 days (1-22 (no p-value reported)	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Cardozo, Fysh, and Pearce, 1986 **(Original intention-to-treat analysis)** ***and*** **Pearce and Cardozo, 1988** **(Sup-plementary analysis including only patients who actually received the treatment to which they were allocated)**	Design: RCT, randomization by chart number Interventions: 1) Planned induction (n = 195 [intention-to-treat analysis]; 125 [supplemental analysis]) Protocol: Labor induced between 40 weeks + 12 days and 40 weeks + 14 days (2-4 days after recruitment/ randomization). PGE_2 suppository (3 mg) inserted, followed 3 hours later by amniotomy and, if necessary, oxytocin infusion. 2) Expectant management (n = 207 [intention-to-treat analysis]; 156 [supplemental analysis]) Protocol: U/S exam given between 40 weeks + 12 days and 40 weeks + 16 days (2-6 days after recruitment/ randomization) to determine ratio of head circumference to abdominal circumference and to estimate amniotic fluid volume. Patients monitored with daily kick count charts and cardiotocography on alternate days. Labor induced for asymmetric IUGR with abnormal cardiotocogram; PROM; or onset of hypertension. Patients in both groups were permitted to request or decline induction of labor after 42 weeks' gestation.	No. of subjects at start: 402 Dropouts: 70 patients in the active group and 41 or 51 in the expectant management group. According to the original publication (Cardozo, Fysh, and Pearch, 1986), 49/70 dropouts from the active group went into labor spontaneously during the waiting period before the planned induction, while the other 21 asked to be induced. According to the supplementary analysis (Pearce and Cardozo, 1988), all 70 went into labor spontaneously. According to the original publication, 2/41 dropouts in the expectant management group had elective C-sections, while the remaining 39 were induced during the waiting period. According to the supplementary analysis, 41 women in the expectant management group went into spontaneous labor during the waiting period, and an additional 10 were induced during the waiting period. All these patients were included in the original intention-to-treat analysis, but were excluded from the later supplementary analysis. Demographic data below are for the intention-to-treat population. Loss to follow-up: NA No. of subjects at end: 402 (intention-to-treat analysis); 281 (supplemental analysis) Inclusion criteria: Uncomplicated pregnancy; gestational age 40	1) Apgar score < 5 at 1 minute 2) Apgar score < 5 at 5 minutes 3) Birthweight 4) Cord venous pH 5) Meconium aspiration syndrome 6) Major FHR tracing abnormality 7) Admission to NICU 6) Duration of 2nd stage of labor 7) Intervention during 2nd stage of labor 8) Forceps-assisted delivery 9) Emergency C-sections 10) Patient satisfaction	1) Apgar score < 5 at 1 minute: *Intention-to-treat analysis:* Induction: 30/195 (15%) Expectant mgmt: 25/207 (12%) p = not significant *Supplemental analysis:* Induction: 19/125 (15%) Expectant mgmt: 16/156 (10%) p = not significant 2) Apgar score < 5 at 5 minutes: *Intention-to-treat analysis:* Induction: 2/195 (1%) Expectant mgmt: 4/207 (2%) p = not significant *Supplemental analysis:* Induction: 1/125 (1%) Expectant mgmt: 2/156 (1%) p = not significant 3) Birthweight (mean ± SD): *Intention-to-treat analysis:* Induction: 3.69 ± 0.51 kg Expectant mgmt: 3.63 ± 0.43 kg p = not significant *Supplemental analysis:* Induction: 3670 ± 500 g Expectant mgmt: 3630 ± 400 g p = not significant 4) Cord venous pH (mean ± SD): *Intention-to-treat analysis:* Induction (n = 84): 7.29 ± 0.10 Expectant mgmt (n = 99): 7.32 ± 0.08 p < 0.05 *Supplemental analysis:* Induction: 7.28 ± 0.10 Expectant mgmt: 7.33 ± 0.08 p = 0.006	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Differences exist between the original and supplementary articles in reporting of the number of patients who went into spontaneous labor before the planned induction period. Original article: 49 (induction group) vs. 0 (expectant management group). Supplementary article: 70 (induction group) vs. 41 (expectant management group) (p < 0.05). Significant difference between two groups in racial distribution at baseline. Results not stratified by parity. No data on baseline Bishop scores. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Dates: NR (patients enrolled over a 21-month period) Location: London, England Setting: 2 hospitals of unspecified type Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	weeks + 10 days (290 days) Exclusion criteria: None specified Age: NR; authors stated only that two groups were "well matched" for maternal age Race: Induction, 73% White; expectant management, 83% White ($p < 0.05$) Gestational age at entry (mean ± SD): 290 days (inclusion criterion) Dating criteria: LMP and U/S performed before 20 weeks Parity: NR; authors stated only that two groups were "well matched" for parity Bishop score: Baseline scores NR		5) Meconium aspiration syndrome: *Intention-to-treat analysis:* Induction: 1/195 (0.5%) Expectant mgmt: 1/207 (0.5%) p = not significant *Supplemental analysis:* Induction: 4/125 (3%) Expectant mgmt: 5/156 (1%) p = not significant 6) Major FHR tracing abnormality: *Intention-to-treat analysis:* Induction: 27/195 (14%) Expectant mgmt: 17/207 (8%) p = not significant *Supplemental analysis:* Induction: 22/125 (14%) Expectant mgmt: 11/156 (7%) $p < 0.02$ 7) Admission to NICU: *Intention-to-treat analysis:* Induction: 6/195 (3%) Expectant mgmt: 3/207 (1.5%) p = not significant *Supplemental analysis:* Induction: 5/125 (4%) Expectant mgmt: 1/156 (1%) p = not significant 6) Duration of 2nd stage of labor (mean): *Intention-to-treat analysis:* Induction (n = 175): 72 minutes Expectant mgmt (n = 188): 77 minutes p = not significant *Supplemental analysis:* Induction (n = 108): 66.2 minutes Expectant mgmt (n = 141): 78.8 minutes p = not significant	*(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				7) Intervention during 2nd stage of labor: *Intention-to-treat analysis:* Induction (n = 175): 44/175 (25%) Expectant mgmt (n = 188): 54/188 (29%) p = not significant *Supplemental analysis:* Induction (n = 108): 31/108 (29%) Expectant mgmt (n = 141): 40/141 (28%) p = not significant 8) Forceps-assisted delivery: *Intention-to-treat analysis:* Induction: 39/195 (20%) Expectant mgmt: 54/207 (26%) p = not significant *Supplemental analysis:* Induction (n = 108): 28/108 (26%) Expectant mgmt (n = 141): 39/141 (28%) p = not significant 9) Emergency C-sections: *Intention-to-treat analysis:* Induction: 25/195 (13%) Expectant mgmt: 18/207 (9%) p = not significant *Supplemental analysis:* Induction (n = 108): 3/108 (3%) Expectant mgmt (n = 141): 1/141 (1%) p = not significant 10) Patient satisfaction: *Intention-to-treat analysis:* Induction / ExpMgmt Pleased 49% / 53% No comment 34% / 35% Disappointed 15% / 11% No response 3% / 1% p = not significant *Supplemental analysis:* Not reported	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Dyson, Miller, and Armstrong, 1987	Design: RCT, randomization according to table of random numbers and sealed envelopes Interventions: 1) Cervical ripening and induction (n = 152) Protocol: Patients underwent cervical ripening with PGE_2 gel (3 mg in initial phase of study, later changed to 0.5 mg), applied intravaginally on an outpatient basis. Patients monitored for ≥ 45 minutes. Those with regular contractions admitted to hospital for continued observation; others allowed to go home. If no labor the next morning (16-18 hours later), then patient admitted to hospital. Oxytocin induction begun if cervical score ≥ 5. If cervical score < 5, then second dose of PGE_2 gel administered and patient monitored for 4 hours. After 4 hours, oxytocin induction started regardless of cervical score. 2) Antepartum monitoring (n = 150) Protocol: NST performed twice weekly. Pelvic exam and determination of AFV performed weekly between 41 and 42 weeks gestation and twice weekly after 42 weeks. Labor induced if abnormal results on fetal testing or if cervical score became ≥ 6.	No. of subjects at start: 302 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 302 Inclusion criteria: Gestational age ≥ 287 days; low risk; unfavorable cervix Exclusion criteria: Risk factors known to increase perinatal mortality and morbidity (e.g., chronic hypertension, pre-eclampsia, diabetes, growth retardation, previous stillbirth); risk factors known to increase risk of induction (e.g., multiple gestation and polyhydramnios); risk factors know to affect C-section rate (e.g., breech presentation and previous C-section); favorable cervix (cervical score ≥ 6); nonreactive NST; variable deceleration on NST; oligohydramnios Age (mean ± SD): Induction, 24.8 ± 4.8; monitoring, 25.1 ± 5.0 Race: NR Gestational age at entry (mean ± SD): Induction, 290.8 ± 2.8 days; monitoring, 290.5 ± 2.6 days Dating criteria: 1) LMP confirmed by either a positive urine test within ≤ 6 weeks gestation or a 1st trimester pelvic exam or a 1st or 2nd trimester U/S; *or* 2) serial U/S exams, with the first performed before 24 weeks	1) Perinatal death 2) Apgar score < 7 at 1 minute 3) Apgar score < 7 at 5 minutes 4) Meconium staining 5) Meconium aspiration (meconium below the vocal cords on intubation, with admission to the NICU for oxygen administration) 6) Post-maturity syndrome 7) Fetal distress (abnormality of FHR tracing prompting C-section or midforceps delivery) 8) Birthweight 9) Macrosomia 10) C-sections 11) Maternal hospital stay 12) Infant hospital stay	1) Perinatal death: Induction: 0 Monitoring: 1/150 (< 1%) p = not significant 2) Apgar score < 7 at 1 minute: Induction: 17/152 (11.2%) Monitoring: 32/150 (21.3%) p < 0.02 3) Apgar score < 7 at 5 minutes: Induction: 2/152 (1.3%) Monitoring: 3/150 (2%) p = not significant 4) Meconium staining: Induction: 29/152 (19.1%) Monitoring: 70/150 (46.7%) p < 0.01 5) Meconium aspiration: Induction: 0 Monitoring: 6/150 (4.0%) p < 0.02 6) Post-maturity syndrome: Induction: 8/152 (5.3%) Monitoring: 22/150 (14.7%) p < 0.01 7) Fetal distress: Induction: 4/152 (2.6%) Monitoring: 27/150 (18.0%) p < 0.01 8) Birthweight (mean ± SD): Induction: 3696 ± 370 g Monitoring: 3766 + 428 p = not significant 9) Macrosomia: Induction: 29/152 (19.1%) Monitoring: 42/150 (28.2%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Dates: Jan 1983 - Dec 1985 Location: Santa Clara, CA Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	Parity (mean ± SD): Induction, 0.4 ± 0.7 (70% nulliparous); monitoring, 0.3 ± 0.6 (73% nulliparous) Bishop score: NR (though see inclusion and exclusion criteria)		10) C-sections: *Overall:* Induction: 22/152 (14.5%) Monitoring: 41/150 (27.3%) p < 0.01 *Among nulliparous women:* Induction: 21/106 (19.8%) Monitoring: 38/110 (34.6%) p < 0.02 *Among multiparous women:* Induction: 1/46 (2.2%) Monitoring: 3/40 (7.5%) p = not significant 11) Maternal hospital stay (mean ± SD): Induction: 3.2 ± 1.3 days Monitoring: 3.5 ± 1.2 days p < 0.04 12) Infant hospital stay (mean ± SD): Induction: 3.0 ± 1.2 days Monitoring: 3.3 ± 1.5 days p = not significant	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Egarter, Kofler, Fitz, et al., 1989	Design: RCT, method of randomization not described Interventions: 1) Induction of labor at due date by intravaginal PGE_2 tablets (3 mg) (n = 180) Protocol: 3 mg PGE_2 tablets applied vaginally. Dose repeated at 6 hours if labor did not start or contractions were inadequate. If patient still undelivered at 24 hours, but cervix ≥ 3 cm, then another treatment course given. If cervix < 3 cm, no further induction attempt performed. 2) "Watchful waiting" (n = 165) Protocol: Cardiotocographic evaluation of fetal well-being performed at 2- to 3-day intervals. Labor induced as above at completion of 42 weeks of amenorrhea. Dates: NR Location: Vienna, Austria Setting: University hospital Type(s) of providers: NR Length of follow-up: None	No. of subjects at start: 356 Dropouts: 11 Loss to follow-up: NA No. of subjects at end: 345 Inclusion criteria: Singleton pregnancies in cephalic presentation reaching their estimated date of confinement; intact membranes; cervix favorable for induction (modified Bishop score > 4) Exclusion criteria: Any fetal or maternal risk factor Age: NR Race: NR Gestational age at entry: NR Dating criteria: "Early" U/S Parity: Induction, 55% nulliparous; watchful waiting, 53% nulliparous Bishop score: NR	1) Fetal death 2) Other fetal outcomes 3) C-sections 4) Forceps-assisted delivery 5) Time from initial visit to spontaneous onset of labor (watchful waiting group only) 6) Number of pregnancies undelivered at 294 days in watchful waiting group 7) Use of analgesic treatment during labor	No p-values reported for outcomes described below. 1) Fetal death: Induction: 0 Watchful waiting: 1/165 (< 1%) 2) Other fetal outcomes: No significant differences between the two groups for birthweight and length, meconium staining, low Apgar scores, or pH. No quantitative data reported for these outcomes. 3) C-sections: *Among primiparae:* Induction: 1/99 (1.0%) Watchful waiting: 3/88 (3.4%) *Among multiparae:* Induction: 1/81 (1.2%) Watchful waiting: 0/77 4) Forceps-assisted delivery: *Among primiparae:* Induction: 3/99 (3.0%) Watchful waiting: 3/88 (3.4%) *Among multiparae:* Induction: 1/81 (1.2%) Watchful waiting: 0/77 5) Time from initial visit to spontaneous onset of labor (mean ± SD) (watchful waiting group only): Among nulliparae (n = 81): 4.5 ± 3.7 days Among multiparae (n = 75): 3.9 ± 2 days 6) Number of pregnancies undelivered at 294 days in watchful waiting group: 7/165 pregnancies (4.2%). All 7 deliveries were "uneventful," though	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: - Dating criteria: + Bishop score: + 11 patients crossed over, but were dropped from analysis. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				umbilical artery pH was slightly low (7.23) in one case. 7) Use of analgesic treatment during labor: Induction: 35% Watchful waiting: 35%	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
El-Torkey and Grant, 1992	Design: RCT, randomization by random permuted blocks and sealed envelopes Interventions: 1) Sweeping of the membranes (n = 33) Protocol: Examination gloves lubricated with jelly or obstetric cream. As much of the membranes as possible were separated from the lower segment. If the cervix would not admit a finger, it was massaged vigorously to encourage prostaglandin release. Patients given date for formal induction of labor. 2) Monitoring (n = 32) Protocol: No form of vaginal examination given. No further details provided on management protocol. Patients given date for formal induction of labor. Dates: June 1990 - Mar 1991 Location: Bellshill, UK Setting: Community hospital Type(s) of providers: NR Length of follow-up: None	No. of subjects at start: 65 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 65 Inclusion criteria: Between 41 and 42 weeks gestation; preferred induction to monitoring when given choice Exclusion criteria: None specified Age (mean ± SD): Sweeping, 27.2 ± 4.7; monitoring, 25.3 ± 5.1 Race: NR Gestational age at entry (mean ± SD): Sweeping, 286.6 ± 2.8 days; monitoring, 286.3 ± 2.8 Dating criteria: NR Parity: Sweeping, 52% nulliparous; monitoring, 44% nulliparous Bishop score: NR	1) Apgar score < 6 at 1 minute 2) Apgar score < 6 at 5 minutes 3) Serious infection 4) Perinatal death 5) Maternal fever (axillary temperature > 37.1° C) 6) C-sections 7) Forceps-assisted delivery 8) Spontaneous delivery 9) Spontaneous labor	1) Apgar score < 6 at 1 minute: Sweeping: 2/33 (6%) Monitoring: 6/32 (19%) p = 0.12 2) Apgar score < 6 at 5 minutes: Sweeping: 1/33 (3%) Monitoring: 1/32 (3%) p = 0.98 3) Serious infection: Sweeping: 0 Monitoring: 0 4) Perinatal death: Sweeping: 0 Monitoring: 0 5) Maternal fever: Sweeping: 0 Monitoring: 4/32 (12.5%) p = 0.04 6) C-sections: Sweeping: 5/33 (15%) Monitoring: 4/32 (12.5%) p = 0.76 7) Forceps-assisted delivery: Sweeping: 2/33 (6%) Monitoring: 3/32 (9%) p =0.62 8) Spontaneous delivery: Sweeping: 26/33 (79%) Monitoring: 25/32 (78%) p =0.95 9) Spontaneous labor: Sweeping: 25/33 (76%) Monitoring: 12/32 (38%) p =0.002	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: - Patients in control group not informed that they were taking part in a randomized trial. Trial suspended before reaching n = 110 because of discrepancy in spontaneous labor rates (main outcome). No specific mention of use of Bishop score, except in reference to other studies that did not use it. Sample size estimates based on proportion of patients entering spontaneous labor.

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Gonen, Rosen, Dolfin, et al., 1997	Design: RCT, randomization by randomly generated numbers; method of concealment NR Interventions: 1) Induction of labor using oxytocin or prostaglandins, depending on Bishop score (criteria not specified) (n = 140) 2) Expectant management with NST/biophysical profile twice weekly and induction if no labor by 42 weeks (n = 144) Dates: Feb 1992 - Aug 1995 Location: Kfar-Saba, Israel Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 284 Dropouts: 11 Loss to follow-up: NA No. of subjects at end: 273 Inclusion criteria: Referral for ultrasound evaluation for potential macrosomia; completed 38 weeks; ultrasound EFW between 4,000 and 4,500 grams Exclusion criteria: Active labor; diabetes; prior cesarean delivery; nonvertex presentation; indications for induction other than macrosomia Age (mean ± SD): Induction, 30.8 ± 5.0; expectant, 29.5 ± 5.2 (p = 0.02) Race: NR Gestational age at entry (mean ± SD): Induction, 284.1 ± 6.4 days; expectant, 284.4 ± 5.7 days Dating criteria: NR Parity: Induction, 31% nulliparous; expectant, 29% nulliparous Bishop score: NR Other: Among nulliparous women, expectantly managed women younger (24.7 ± 3.0 vs. 27.6 ± 4.6; p = 0.001); no other differences	1) Time to delivery 2) Vaginal deliveries, stratified by parity 3) Instrumental deliveries, stratified by parity 4) C-section rates, stratified by parity 5) Umbilical artery pH 6) Shoulder dystocia 7) Cephalohematoma 8) Clavicular fracture 9) Brachial plexus palsy 10) Intraventricular hemorrhage	1) Time to delivery: Induction: 18.6 hours; range, 2-72 hours; 78% delivered within 24 hours Expectant: 4.1 ± 4.0 days Results similar in nulliparous and parous women 2) Vaginal deliveries, stratified by parity: *Overall:* Induction: 67.9% Expectant: 65.5% p = not significant *Nulliparous:* Induction: 35.7% Expectant: 50.0% p = not significant *Multiparous:* Induction: 82.6% Expectant: 71.7% p = not significant 3) Instrumental deliveries, stratified by parity: *Overall:* Induction: 12.7% Expectant: 12.9% p = not significant *Nulliparous:* Induction: 26.2% Expectant: 15.0% p = not significant *Multiparous:* Induction: 6.5% Expectant: 12.1% p = not significant 4) C-section rates, stratified by parity: *Overall:* Induction: 19.4% Expectant: 21.6%	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Study underpowered to detect differences in categorical variables and rare outcomes. Unclear if any women randomized to expectant management who were induced because of abnormal testing were excluded from analysis. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				p = not significant *Nulliparous:* Induction: 38.1% Expectant: 35.0% p = not significant *Multiparous:* Induction: 10.9% Expectant: 16.2% p = not significant 5) Umbilical artery pH: Induction: 7.32 ± 0.07 Expectant: 7.33 ± 0.06 No differences when stratified by parity 6) Shoulder dystocia: Induction: 5/108 Expectant: 6/109 p = not significant 7) Cephalohematoma: Induction: 6/134 (5 instrumental deliveries) Expectant: 3/139 (1 instrumental delivery) 8) Clavicular fracture: Induction: 0/134 Expectant: 2/139 9) Brachial plexus palsy: Induction: 0/134 Expectant: 2/139 10) Intraventricular hemorrhage: Induction: 44/134 had ultrasound; confirmed in 3 Expectant: 31/139 had ultrasound; confirmed in 2	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Hannah, Hannah, Hellmann, et al., 1992 ***and*** **Goeree, Hannah, Hewson, 1995** **(cost-effective-ness analysis)**	Design: RCT, randomization stratified according to center, parity, and duration of gestation Interventions: 1) Induction of labor (n = 1701) Protocol: Subjects enrolled as outpatients. Labor to be induced within 4 days of randomization. If cervix < 3 cm dilated and < 50% effaced, and FHR normal, then patient given PGE_2 gel (0.5 mg) intracervically. Fetus monitored for minimum of 1 hour. Up to 3 doses of gel could be given at 6-hour intervals. If gel not used or did not induce labor, then labor induced by IV oxytocin, amniotomy, or both. Oxytocin infusion not started until 12 hours after last dose of gel. 2) Monitoring (n = 1706) Protocol: Subjects enrolled as outpatients and asked to do "kick counts" over 2-hour period each day, undergo NST 3 times per week, and undergo U/S assessments of AFV 2-3 times per week. If kick count < 6, then patients to contact physician and have NST within 12 hours. If NST nonreactive or showed deceleration in FHR, if AFV low (a pocket of < 3 cm), if obstetrical complications developed, or if gestational age reached 44 weeks, then fetus to be delivered either by	No. of subjects at start: 3418 Dropouts: 11 Loss to follow-up: NA No. of subjects at end: 3407 (Note: 7 of these 3407 women had infants with major congenital anomalies and were excluded from the analysis of perinatal and neonatal outcomes, as were 2 stillborns) Inclusion criteria: Live singleton fetus; ≥ 41 weeks gestation Exclusion criteria: Cervical dilatation ≥ 3 cm; gestational age ≥ 44 weeks; noncephalic presentation; lethal congential anomaly; diabetes mellitus; preeclampsia; intrauterine growth retardation; pre-labor rupture of membranes; need for urgent delivery; contraindications to vaginal delivery Age: Induction Monitoring < 20 4% 3% 20-35 86% 87% > 35 10% 10% Race: Induction Monitoring White 93% 92% Black 3% 3% Asian 2% 2% Other/ Unknown 2% 3% Gestational age at entry (in weeks): Induction Monitoring 40 3% 3%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight > 4500 g 4) Shoulder dystocia 5) Meconium aspiration 6) Cord pH < 7.10 7) Admission to NICU 8) Stillbirths 9) Neonatal death 10) C-sections 11) Instrumental delivery 12) Length of stay 13) Hospital costs per patient 14) Professional fees per patient	1) Apgar score < 7 at 1 minute: Induction: 216/1700 (12.7%) Monitoring: 216/1698 (12.7%) p = not significant 2) Apgar score < 7 at 5 minutes: Induction: 18/1700 (1.1%) Monitoring: 20/1698 (1.2% p = not significant 3) Birthweight > 4500 g: Induction: 78/1700 (4.6%) Monitoring: 94/1698 (5.5%) p = not significant 4) Shoulder dystocia: Induction: 24/1701 (1.4%) Monitoring: 28/1706 (1.6%) p = not significant 5) Meconium aspiration: Induction: 96/1700 (5.7%) Monitoring: 95/1698 (5.6%) p = not significant 6) Cord pH < 7.10: Induction: 23/1700 (1.4%) Monitoring: 29/1698 (1.7%) p = not significant 7) Admission to NICU: Induction: 239/1700 (14.1%) Monitoring: 263/1698 (15.5%) p = not significant 8) Stillbirths: Induction: 0 Monitoring: 2 (no p-value reported) 9) Neonatal deaths: Induction: 0 Monitoring: 0	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Selection of mode of delivery was not standardized, but rather determined by the attending physician. For the cost analysis, minor costs were estimated from a sample of 129 charts. Sample size estimates based on reduction in incidence of Apgar score < 7 at 5 minutes. C-section rates higher among nulliparous women, older women, women with less dilatation at randomization, and women in "Black" and "Other" racial categories, independent of study group. Women induced in monitoring group less likely to receive prostaglandin for induction. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	inducing labor (using oxytocin or amniotomy) or by C-section. In every case, mode of delivery determined by attending physician. Dates: Nov 1985 - Dec 1990 Location: 22 sites "through-out Canada" (Canadian Multicentre Postterm Pregnancy Trial) Setting: 19 university hospitals and 3 community hospitals Type(s) of providers: Unspecified OB/GYN; radiologists Length of follow-up: None	41 88% 89% 42 9% 7% 43 < 1% < 1% Dating criteria: Either 1) LMP or known date of conception, confirmed by pregnancy test at < 6 weeks, physical exam at ≤ 20 weeks, or U/S at ≤ 26 weeks; *or* 2) U/S ≤ 26 weeks (if LMP uncertain); *or* 3) two consistent U/S at ≤ 26 weeks (if LMP unknown) Parity: 68% nulliparous (both groups) Bishop score: NR Other: Cervical dilatation before entry (in cm): Induction Monitoring 0 40% 40% 1-2 51% 49% 3-4 1% 1% Unknown 9% 10%		10) C-sections: *Overall:* Induction: 360 (21.2%) Monitoring: 418 (24.5%) p = 0.03 (controlled for parity, maternal age, cervical dilatation at time of randomization, and race) OR = 1.22 (95% CI, 1.02-1.45) *For fetal distress:* Induction: 97 (5.7%) Monitoring: 141 (8.3%) p = 0.003 11) Instrumental delivery: Induction: 473/1341 (35.3%) Monitoring: 449/1288 (34.9%) (no p-value reported) 12) Length of stay (mean): Induction: 3.9 days Monitoring: 4.0 days (no p-value reported) 13) Hospital costs (mean per patient in 1992 Canadian dollars): Induction: $2502 Monitoring: $2684 p < 0.0001 14) Professional fees (mean per patient in 1992 Canadian dollars): Induction: $437 Monitoring: $448 p = 0.025	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Hedén, Ingemarsson, Ahlström, et al., 1991	Design: RCT, randomization by "birth registration number" Interventions: 1) Induction (n = 109) Protocol: Labor induced on day of recruitment by amniotomy and oxytocin infusion. (No further details provided.) 2) Monitoring ("expectant management") (n = 129) Protocol: Every-other-day clinical exam, cervical exam, and NST + weekly U/S assessment of AFV. If NST "ominous," then labor induced. If NST nonreactive, but not ominous, then oxytocin stress test (OST) performed. If OST normal, then monitoring protocol continued. If OST "ominous," then labor induced. If no pocket of fluid measuring at least 2 x 2 cm detected on U/S, then labor induced. Dates: NR; study conducted over a 3-year period Location: Lund and Ängelholm, Sweden Setting: University hospital and community hospital (2 sites) Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 238 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 238 Inclusion criteria: 42 weeks gestation; no complications; singleton fetus in vertex presentation; intact membranes; cervix < 4 cm; no regular contractions; normal NST; normal AFV Exclusion criteria: Prior C-section Age (mean ± SD): Induction, 29.5 ± 5.4; monitoring, 28.4 ± 4.9 Race: NR Gestational age at entry: 42 weeks (both groups) Dating criteria: U/S during weeks 16-18 Parity: Induction, 37% nulliparous; monitoring, 48% nulliparous Bishop score (mean ± SD): Induction, 5.3 ± 1.7; monitoring, 5.0 ± 2.1	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Severe dysmaturity 5) Admission to NICU 6) Meconium staining 7) C-sections 8) Forceps/vacuum extraction	1) Apgar score < 7 at 1 minute: Induction: 5/109 (4.6%) Monitoring: 6/129 (4.7%) p = not significant 2) Apgar score < 7 at 5 minutes: Induction: 3/109 (2.8%) Monitoring:1/129 (0.8%) p = not significant 3) Birthweight (mean): Induction: 4000 g Monitoring: 3900 g p = not significant 4) Severe dysmaturity: Induction: 4/109 (3.7%) Monitoring: 3/129 (2.3%) p = not significant 5) Admission to NICU: Induction: 10/109 (9.2%) Monitoring:8/129 (6.2%) p = not significant 6) Meconium staining: Induction: 15.6% Monitoring: 24.8% p = not significant 7) C-sections: Induction: 10/109 (9.2%) Monitoring: 9/129 (7.0%) p = not significant 8) Forceps/vacuum extraction: *Total:* Induction: 3/109 (2.8%) Monitoring: 20/129 (15.5%) p < 0.01 *For secondary arrest:* Induction: 2/109 (1.8%) Monitoring: 17/129 (13.2%) (p < 0.01)	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: - Mode of delivery: - Sample size: - Statistical tests: - Gestational age: + Dating criteria: + Bishop score: + No sample size estimates. Unequal distribution of "semi-randomization" raises question of bias. Results not stratified by parity.

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Herabutya, Prasert-sawat, Tongyai, et al., 1992	Design: RCT, method of randomization not described Interventions: 1) Cervical ripening and induction (n = 57) Protocol: PGE_2 gel applied intracervically (6 tablets of 0.5 mg each mixed into 5 ml K-Y Jelly). Patient reassessed in 4-6 hours. If Bishop score > 6, then patient induced with amniotomy ± oxytocin (at discretion of obstetrician in charge of labor ward). If Bishop score < 6, then patient sent home, unless uterine contractions or "anticipated problem"; patients in latter categories kept in hospital and could receive 2nd dose after 6 hours if "urgent reasons" to repeat dose. Process repeated next morning, up to maximum of 3 doses. If Bishop score still < 6, then patient induced by amniotomy or oxytocin or both. 2) Monitoring (n = 51) Protocol: NST once weekly from 42-43 weeks and twice weekly after 43 weeks. Labor induced if NST abnormal, Bishop score > 6, or 44 weeks of gestation completed. For both groups, intrapartum management *not* dictated by study protocol. Dates: July 1987 - Jan 1991 Location: Bangkok, Thailand	No. of subjects at start: 108 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 108 Inclusion criteria: Gestational age ≥ 42 weeks; low risk Exclusion criteria: Bishop score > 6 Age (mean ± SD): Induction, 27.4 ± 4.1; monitoring, 27.1 ± 4.3 Race: 100% Thai Gestational age at entry: NR (required to be ≥ 42 weeks for entry into study) Dating criteria: LMP, with consistent obstetric exam at < 20 weeks Parity: Induction, 90% nulliparous; monitoring, 80% nulliparous Bishop score: NR (required to be ≤ 6 for entry into study)	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Meconium 4) Intubation required 5) Admission to NICU 6) Birthweight 7) Length of 1st stage of labor 8) C-sections 9) Instrumental deliveries	1) Apgar score < 7 at 1 minute: Induction: 15/57 (26.3%) Monitoring: 15/51 (29.4%) p = 0.89 2) Apgar score < 7 at 5 minutes: Induction: 1/57 (1.8%) Monitoring: 4/51 (7.8%) p = 0.19 3) Meconium: Induction: 8/57 (14.0%) Monitoring: 11/51 (21.6%) p = 0.44 4) Intubation required: Induction: 1/57 (1.8%) Monitoring: 4/51 (7.8%) p = 0.19 5) Admission to NICU: Induction: 1/57 (1.8%) Monitoring: 4/51 (7.8%) p = 0.19 6) Birthweight (mean ± SD): Induction: 3190 ± 429 g Monitoring: 3348 ± 421 g p = 0.06 7) Length of 1st stage of labor (mean ± SD): Induction: 8.15 ± 3.5 hours Monitoring: 9.15 ± 4.6 hours p =0.36 8) C-sections: *Overall:* Induction: 27/57 (47.4%) Monitoring: 24/51 (47.1%) p = 0.87 *For cephalopelvic disproportion:* Induction: 25/57 (43.9%)	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not stratified by parity. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None			Monitoring: 19/51 (37.3%) p = 0.62 *For fetal distress:* Induction: 2/57 (3.5%) Monitoring: 5/51 (9.8%) p = 0.26 9) Instrumental deliveries: Induction: 11/57 (19.3%) Monitoring: 9/51 (17.6%) p = 0.98	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Katz, Yemini, Lancet, et al., 1983	Design: RCT, assignment to group by even/odd chart number Interventions: 1) Induction at 294 days (n = 78) Protocol: Labor induced by amniotomy and oxytocin infusion at 294 days. 2) Monitoring (n = 78) Protocol: Patients instructed to count fetal movements at home twice daily and to report to labor and delivery ward if movements decline by more than 50% or fall below 10 per hour. Patients seen every 3 days for assessment of "pelvic score" (Burnett, 1966), amnioscopy to check for meconium, OCT, and assessment of fetal movement count. If pelvic score > 4 or any of other 3 indicators "pathologic," then patient induced. Dates: NR Location: Jerusalem, Israel Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 156 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 156 Inclusion criteria: 294 days amenorrhea; "pelvic score" (Burnett, 1966) ≤ 4; vertex presentation; no obstetric pathology; no uterine scars; clear amniotic fluid by amnioscopy; normal NST; regular fetal movement perceived by mother Exclusion criteria: None specified Age (mean ± SD): Induction, 26.3 ± 4.1; monitoring, 26.5 ± 4.2 Race: NR Gestational age at entry: Both groups, 294 days Dating criteria: Positive pregnancy test within 6 weeks of LMP or 4 weeks following ovulation; *or* palpation of the uterus during 1st trimester and/or U/S before 30th week Parity: Induction, 46% primiparae; Monitoring, 45% primiparae Bishop score: NR; "pelvic score" (Burnett, 1966) required to be ≤ 4 for entry into study Other: NA	1) Apgar scores at 5 minutes (mean) 2) Apgar score < 7 at 5 minutes 3) Meconium staining 4) Intrapartum changes in FHR 5) Post-maturity syndrome 6) Birthweight (mean) 7) Birthweight > 4000 g 8) Perinatal death 9) C-sections 10) Duration of labor	1) Apgar scores at 5 minutes (mean): Induction: 9.5 Monitoring: 9.7 p = not significant 2) Apgar score < 7 at 5 minutes: Induction: 3/78 (3.8%) Monitoring: 1/78 (1.3%) (no p-value reported) 3) Meconium staining: Induction: 11/78 (14.1%) Monitoring: 12/78 (15.4%) p = not significant 4) Intrapartum changes in FHR: Induction: 9/78 (11.5%) Monitoring: 5/78 (6.4%) p = not significant 5) Post-maturity syndrome: Induction: 5/78 (6.4%) Monitoring: 11/78 (14.1%) p = not significant 6) Birthweight (mean): Induction: 3380 g Monitoring: 3540 g p = not significant 7) Birthweight > 4000 g: Induction: 6/78 (7.9%) Monitoring: 23/78 (29.5%) p < 0.05 8) Perinatal death: Induction: 1/78 (1.3%) Monitoring: 1/78 (1.3%) p = not significant 9) C-sections: Induction: 16/78 (20.5%) Monitoring: 7/78 (8.8%) p < 0.05	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: - Gestational age: + Dating criteria: + Bishop score: + Burnett, 1966 = Burnett JE. Preinduction scoring: an objective approach to induction of labour. Obstet Gynecol 1966;28:479-83. Results not stratified by parity. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				10) Duration of labor (mean ± SD): Induction: 9.4 ± 5.9 hours Monitoring: 6.7 ± 4.1 hours $p < 0.01$	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Martin, Sessums, Howard, et al., 1989	Design: RCT, randomization by sealed envelope Interventions: 1) Induction (n = 12) Protocol: Patients admitted to hospital. Laminaria tent(s) inserted. Subsequently (usually the following morning), laminaria tents(s) removed, and labor induced by oxytocin infusion. Fetal heart tones monitored throughout labor. 2) Monitoring (n = 10) Protocol: Weekly monitoring, including U/S assessment of AFV, NST/CST, and cervical exam. Patients "admitted for delivery" if any monitoring test abnormal, or at the end of 43rd week of gestation. Dates: July 1987 - Jan 1988 Location: Jackson, MS Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 22 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 22 Inclusion criteria: Gestational age ≥ 41 weeks Exclusion criteria: Oligo-hydramnios (< 1 cm); nonreactive NST; positive CST; Bishop score > 5 Age (mean, with range): Induction, 23.3 (17-34); monitoring, 25.8 (18-37) Race: NR Gestational age at entry (mean, with range): Induction, 42 weeks (41-2/7 to 43-2/7); monitoring, 42 weeks (41-3/7 to 43-3/7) Dating criteria: LMP, 1st trimester pelvic exam, and/or U/S before 26 weeks Parity (mean): Induction, 0.76; monitoring, 0.58 Bishop score: NR	1) Apgar score at 1 minute 2) Apgar score at 5 minutes 3) Birthweight 4) Meconium 5) Complications 6) C-sections 7) Forceps-assisted deliveries 8) Length of labor 9) Maternal morbidity 10) Length of hospital stay	1) Apgar score at 1 minute (mean): Induction: 8.08 Monitoring: 8.4 p = not significant 2) Apgar score at 5 minutes (mean): Induction: 9.75 Monitoring: 9.7 p = not significant 3) Birthweight (mean, with range): Induction: 3560 g (2780-4110) Monitoring: 3472 g (2840-4180) p = not significant 4) Meconium: Induction: 1/12 (8%) Monitoring: 3/10 (30%) (no p-value reported) 5) Complications: Induction: 3/12 (25%) Monitoring: 1/10 (10%) (no p-value reported) 6) C-sections: Induction: 2/12 (17%) Monitoring: 1/10 (10%) p = not significant 7) Forceps-assisted deliveries: Induction: 3/12 (25%) Monitoring: 2/10 (25%) p = not significant 8) Length of labor (mean, with range): Induction: 6.33 hours (4-15) Monitoring: 8.3 hours (4-16) p = not significant 9) Maternal morbidity: Induction: 4/12 (33%) Monitoring: 2/10 (20%) (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Results not stratified by parity. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				10) Length of hospital stay (mean, with range): Induction: 3.41 days (2-5) Monitoring: 2.6 days (2-6) p = not significant	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
National Institute of Child Health and Human Develop-ment Network of Maternal-Fetal Medicine Units, 1994	Design: RCT, randomization by computer-generated random numbers Interventions: 1) PGE_2 gel + induction by oxytocin (n = 174) Protocol: PGE_2 gel (0.5 mg) inserted into intracervical canal within 24 hours of randomization. No repeat applications. FHR and uterine contractions monitored continuously for ≥ 4 hours. If no labor after 12 hours, then patient induced using amniotomy (where clinically feasible), followed by oxytocin infusion ("according to a uniform protocol"). If no active labor 24 hours after oxytocin infusion, then C-section performed or induction of labor continued. (Decision to perform C-section not dictated by study protocol.) 2) Placebo gel + induction by oxytocin (n = 91) Protocol: Same as in 1), above, except that placebo gel used instead of PGE_2 gel. 3) Monitoring (n = 175) Protocol: Weekly cervical exam + twice-weekly NST and U/S assessment of AFV. Spontaneous labor awaited, but labor could be induced if: Bishop score > 6; estimated fetal weight > 4500 g; medical or obstetric indication for delivery developed; largest pocket of amniotic fluid < 2	No. of subjects at start: 440 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 440 Inclusion criteria: Gestational age ≥ 287 days and < 301 days Exclusion criteria: Medical or obstetric complications requiring induction, C-section, or frequent monitoring; estimated fetal weight > 4500 g; Bishop score ≥ 7; non-reactive NST; amniotic fluid pocket < 2 cm Age (mean ± SD): PGE_2-oxytocin: 25.4 ± 5.7 Placebo-oxytocin: 25.4 ± 5.3 Monitoring: 26.1 ± 5.8 Race: PGE_2-oxytocin: 67% White, 32% Black, 1% not available Placebo-oxytocin: 63% White, 37% Black Monitoring: 60% White, 38% Black, 2% not available Gestational age at entry: PGE_2-oxytocin: 8l% 287-293 days; 19% 295-301 days Placebo-oxytocin: 79% 287-293 days; 21% 295-301 days Monitoring: 79% 287-293 days; 21% 295-301 days Dating criteria: Any one of following: 1) LMP + audible fetal heartbeat documented for ≥ 21 weeks by fetoscope or ≥ 30 weeks	1) Mechanical ventilation 2) Meconium aspiration 3) Nerve injury 4) Seizures 5) ≥ 1 adverse neonatal outcome 6) Apgar score < 4 at 5 minutes 7) Birthweight (mean) 8) Birthweight ≥ 4500 g 9) Time from randomization to delivery 10) Gestational age at delivery 11) Maternal infection 12) Maternal transfusion 13) Hyperstimulation 14) C-sections	1) Mechanical ventilation: PGE_2-oxytocin: 0 Placebo-oxytocin: 1/91 (1%) Monitoring: 1/175 (< 1%) (no p-value reported) 2) Meconium aspiration: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 1/91 (1%) Monitoring: 2/175 (1%) (no p-value reported) 3) Nerve injury: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 0 Monitoring: 0 (no p-value reported) 4) Seizures: PGE_2-oxytocin: 0 Placebo-oxytocin: 2/91 (2%) Monitoring: 1/175 (< 1%) (no p-value reported) 5) ≥ 1 adverse neonatal outcome: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 3/91 (3%) Monitoring: 1/175 (< 1%) (no p-value reported) 6) Apgar score < 4 at 5 minutes: PGE_2-oxytocin: 0 Placebo-oxytocin: 0 Monitoring: 1/175 (< 1%) (no p-value reported) 7) Birthweight (mean ± SD): PGE_2-oxytocin: 3607 ± 382 g Placebo-oxytocin: 3532 ± 464 g Monitoring: 3606 ± 440 g (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on perinatal morbidity/mortality and maternal mortality.

(continued on next page)

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	cm; or abnormal NST followed by positive CST. If NST nonreactive, but CST negative, then testing repeated in 24 hours. Patients undelivered by 308 days (44 completed weeks) were released from the protocol and managed as "appropriate for the clinical situation." Dates: Dec 1987 - July 1989 Location: Multiple sites in US Setting: University hospitals Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	by Doppler; 2) LMP + compatible uterine size estimation at ≤ 24 weeks; 3) LMP + positive pregnancy test obtained early enough to assure that gestation exceeded 41 weeks; 4) if LMP uncertain, then fetal heartbeat documented for ≥ 32 weeks by Doppler; 5) U/S before 26 weeks Parity (% nulliparous): PGE_2-oxytocin: 60% Placebo-oxytocin: 59% Monitoring: 54% Bishop score (mean ± SD): PGE_2-oxytocin: 4.0 ± 1.4 Placebo-oxytocin: 3.8 ± 1.4 Monitoring: 3.9 ± 1.5		8) Birthweight ≥ 4500 g: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 3/91 (3%) Monitoring: 6/175 (4%) (no p-value reported) 9) Time from randomization to delivery (median, with range): PGE_2-oxytocin: 36 hours (6-492) Placebo-oxytocin: 35 hours (7-487) Monitoring: 85 hours (5-538) p < 0.001 10) Gestational age at delivery: 287-293 days / 294-301 days / >302 days PGE_2-oxy: 64% / 34% / 1% Placebo-oxy: 66% / 32% / 2% Monitoring: 38% / 47% / 14% p < 0.001 11) Maternal infection: PGE_2-oxytocin: 33/174 (19%) Placebo-oxytocin: 13/91 (14%) Monitoring: 25/175 (14%) p = not significant 12) Maternal transfusion: PGE_2-oxytocin: 2/174 (1%) Placebo-oxytocin: 0 Monitoring: 3/175 (2%) p = not significant 13) Hyperstimulation: PGE_2-oxytocin: 2/174 (1%) Placebo-oxytocin: 1/91 (1%) Monitoring: 0 p = not significant 14) C-sections: PGE_2-oxytocin: 39/174 (22%) Placebo-oxytocin: 16/91 (18%) Monitoring: 32/175 (18%) p = not significant	

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Ohel, Rahav, Rothbart, et al., 1996	Design: RCT, assignment to group by even/odd registration number Interventions: 1) Induction (n = 70) Protocol: NST + U/S assessment of AFV performed before treatment. If NST normal, then 3-mg vaginal tablet of PGE_2 inserted into the posterior vaginal fornix. Patients sent home and instructed to return in 3-4 days for repeat testing and a further dose of PGE_2. 2) Monitoring (n = 104) Protocol: Patients "seen" twice weekly (monitoring protocol not described). Labor induced if patient passed 42 completed weeks of gestation. Dates: NR Location: Tiberias, Israel Setting: Unspecified hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 200 Dropouts: 26 Loss to follow-up: NA No. of subjects at end: 174 Inclusion criteria: Uncomplicated, singleton pregnancy; within 4 days after expected date of confinement Exclusion criteria: None specified Age (mean ± SD): Induction, 28.9 ± 4.0; monitoring, 28.2 ± 5.3 Race: NR Gestational age at entry: NR; *at delivery* (mean ± SD), Induction, 40.2 ± 0.5 weeks; monitoring, 40.9 ± 0.7 weeks Dating criteria: "Early" U/S Parity (mean ± SD): Induction, 2.2 ± 1.1; monitoring, 2.4 ± 1.5 Bishop score (mean ± SD): Induction, 4.1 ± 1.6; monitoring, 4.6 ± 1.6	1) Apgar scores at 5 minutes 2) Meconium staining 3) Birthweight > 4 kg 4) C-sections	1) Apgar scores at 5 minutes (mean ± SD): Induction: 9.5 ± 0.6 Monitoring: 9.4 ± 0.6 p = not significant 2) Meconium staining: Induction: 5/70 (7.1%) Monitoring: 20/104 (19.2%) p < 0.02 3) Birthweight > 4 kg: Induction: 6/70 (8.6%) Monitoring: 9/104 (8.7%) p = not significant 4) C-sections: Induction: 4/70 (5.7%) Monitoring: 6/104 (5.8%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: - Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + 26 patients randomized to the induction group refused treatment and were excluded from analysis. Results not stratified by parity.

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Witter and Weitz, 1987	Design: RCT, randomization by computer-generated table of random numbers Interventions: 1) Induction at 42 weeks by oxytocin infusion + amniotomy (n = 103) Protocol: All patients instructed to keep 3-times-daily fetal motion charts. If decreased fetal motion, then OCT administered. If OCT positive, then patient delivered. If OCT negative, then patient continued with protocol. At 42 weeks, undelivered patients scheduled for induction of labor. Oxytocin infusion started at 7:00 AM with 1 mU/min and increased by 1 mU/min every 10 min until a dose of 30 mU/min reached or a regular pattern of adequate uterine contractions established. Amniotomy performed as soon as possible, but always after oxytocin had established regular contractions. If patient had intact membranes and was not in active phase labor by evening, the induction was stopped and the patient was rested overnight. The induction was restarted in the morning. If the patient failed to enter the active phase of labor by 20 hours of induction, then C-section performed. 2) Monitoring (principally by 24-hour urinary estriol	No. of subjects at start: 200 Dropouts: 5 (but included in analysis) Loss to follow-up: NA No. of subjects at end: 195 (200 included in analysis) Inclusion criteria: 41 completed weeks' gestation; uncomplicated pregnancy Exclusion criteria: None stated Age (mean ± SD): Induction, 20.95 ± 4.01; monitoring, 20.98 ± 3.67 Race: Induction, 20% White; monitoring, 34% White ($p < 0.05$) Gestational age at entry: NR; *at delivery* (mean ± SD), induction, 42.15 ± 1.92 weeks; monitoring, 42.41 ± 1.45 weeks Dating criteria: 2 or more of the following: certain LMP; basal body temperature indicating ovulation temperature shift for the present pregnancy; positive urinary pregnancy test at 6 weeks from LMP; fetal heart tones heard with DeLee stethoscope at 18-20 weeks; fundal height at the umbilicus at 20 weeks; fundal height in cm equal to gestational age in weeks within 2 cm from 20-34 weeks; early registration with dates equal to exam prior to 13 weeks; U/S dating by crown-rump length between 6 and 14 weeks or	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Small for gestational age 5) Large for gestational age 6) Post-maturity syndrome 7) Meconium aspiration 8) Endometritis 9) C-sections 10) Hospital stay	1) Apgar score < 7 at 1 minute: Induction: 20/103 (19.4%) Monitoring: 20/97 (21.1%) p = not significant 2) Apgar score < 7 at 5 minutes: Induction: 0 Monitoring: 2/97 (2.08%) p = not significant 3) Birthweight (mean ± SD): Induction: 3556.5 ± 436.3 g Monitoring: 3614.7 ± 472.2 g p = not significant 4) Small for gestational age: Induction: 0 Monitoring: 4/97 (4.43%) $p < 0.05$ 5) Large for gestational age: Induction: 21/103 (20.03%) Monitoring: 29/97 (29.59%) p = not significant 6) Post-maturity syndrome: Induction: 1/103 (0.97%) Monitoring: 2/97 (2.06%) p = not significant 7) Meconium aspiration: Induction: 2/103 (1.94%) Monitoring: 1/97 (1.03%) p = not significant 8) Endometritis: Induction: 12/103 (11.65%) Monitoring: 12/97 (12.37%) p = not significant 9) C-sections: *Overall:* Induction: 30/103 (29.13%) Monitoring: 27/97 (27.83%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Results not stratified by parity. *(continued on next page)*

Evidence Table 2: Studies relevant to Key Question 2 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	creatinine ratio) (n = 97) Protocol: All patients instructed to keep 3-times-daily fetal motion charts. If decreased fetal motion, then OCT administered. If OCT positive, then patient delivered. If OCT negative, then patient continued with protocol. In addition, 24-hour urinary estriol creatinine ratio determined between 41 and 42 weeks. This increased to twice weekly at 42 completed weeks and three times weekly at 43 completed weeks. If 24-hour urinary estriol creatinine ratio ≤ 14 mg/g, then OCT performed. If OCT "reassuring," then patient kept as inpatient and given daily urinary estriol creatinine ratio tests and twice weekly OCTs until spontaneous labor occurred, or until delivery required (Bishop score ≥ 9 or signs of fetal compromise). If estriol creatinine ratio > 14 mg/g, the patient followed as outpatient until spontaneous labor occurred. Dates: NR Location: Baltimore, MD Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	by biparietal diameter prior to 26 weeks Parity: Induction, 51% nulliparous; monitoring, 41% nulliparous Bishop score: NR		p = not significant *For fetal distress:* Induction: 11/30 (36.67%) Monitoring: 13/27 (48.15%) p = not significant *For cephalopelvic disproportion/failure to progress:* Induction: 11/30 (36.67%) Monitoring: 13/27 (48.15%) p = not significant *For prolonged latent phase:* Induction: 7/30 (23/33%) Monitoring: 0 p < 0.01 *For breech presentation:* Induction: 1/30 (3.33%) Monitoring: 1/27 (3.70%) p = not significant 10) Hospital stay (mean ± SD): Induction: 4.74 ± 2.80 days Monitoring: 4.06 ± 1.90 days p < 0.05	

Evidence Table 3: Studies relevant to Key Question 3

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Allott and Palmer, 1993	Design: RCT, randomization by computer-generated list and sealed envelope Interventions: 1) Cervical exam to assess Bishop score + sweeping of the membranes (n = 99) Protocol: Examiner's index finger inserted as far as possible through internal cervical os and rotated twice through 360 degrees. Patients allowed to go home with a fetal movement chart. Instructed to telephone labor ward if they experienced decreased fetal movements, rupture of the membranes, or onset of labor. 2) Cervical exam to assess Bishop score alone (control) (n = 96) Protocol: Not described. Patients in both groups given deadline date for labor to be induced in the absence of spontaneous onset. Dates: NR (18-month period) Location: Reading, UK Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 195 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 195 Inclusion criteria: > 40 weeks gestation; no risk factors (e.g., IUGR or hypertension); able to introduce finger into cervix Exclusion criteria: None specified Age (mean ± SD): Sweeping, 27.7 ± 5.7; control, 27.5 ± 4.9 Race: NR Gestational age at entry (mean ± SD): Sweeping, 284.7 ± 3.3 days; control, 285.3 ± 3.5 days Dating criteria: Mid-trimester U/S Parity: Sweeping, 43% nulliparous; control, 46% nulliparous Bishop score: Both groups, 44% ≤ 6, 56% ≥ 7	1) Apgar score < 6 at 1 minute 2) Apgar score < 6 at 5 minutes 3) Serious neonatal infection 4) Antibiotics given 5) "Other serious neonatal outcome" 6) Induction of labor 7) C-sections 8) Epidural 9) Duration of labor 10) Precipitate labor (< 2 hours) 11) Time to delivery	1) Apgar score < 6 at 1 minute: Sweeping: 4/99 (4.0%) Control: 9/96 (9.4%) (no p-value reported) 2) Apgar score < 6 at 5 minutes: Sweeping: 0 Control: 0 (no p-value reported) 3) Serious neonatal infection: Sweeping: 0 Control: 1/96 (1%) (no p-value reported) 4) Antibiotics given: Sweeping: 0 Control: 1/96 (1%) (no p-value reported) 5) Other serious neonatal outcome: Sweeping: 0 Control: 0 (no p-value reported) 6) Induction of labor: Sweeping: 8/99 (8.1%) Control: 18/96 (18.8%) p = 0.035 7) C-sections: Sweeping: 4/99 (4.0%) Control: 5/96 (5.2%) p = not significant 8) Epidural: Sweeping: 19/99 (19.2%) Control: 20/96 (20.8%) p = not significant 9) Duration of labor (mean): Sweeping: 8.2 hours Control: 7.7 hours p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on induction rates. Significant differences seen when results stratified by parity and Bishop score, except among primigravida with high Bishop score. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				10) Precipitate labor (< 2 hours): Sweeping: 14/99 (14.1%) Control: 19/96 (19.8%) p = not significant 11) Time to delivery (mean ± SEM): Sweeping: 2.24 ± 0.22 days Control: 5.18 ± 0.47 days p = 0.0001	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Atad, Hallak, Auslender, et al., 1996	Design: RCT, randomization by computer-generated list of random numbers Interventions: 1) PGE_2 (n = 30) Protocol: 3-mg tablet placed intravaginally. If contractions had not started or patient did not need analgesic agents 6 hours later, then second dose administered. If Bishop score still ≤ 4 at 12 hours, then patient treated with ARD. 2) Oxytocin (n = 30) Protocol: Oxytocin infusion given in initial dose of 1.5 mIU/min, with an increase of 1.5 mIU/min every 20 minutes until 3 contractions/10 minutes achieved. If Bishop score still ≤ 4 at 12 hours, then patient treated with ARD. 3) Atad Ripener Device (ARD) = double-balloon device invented by lead author (n = 35). Protocol: Device inserted into the cervix, and both balloons inflated with 100 ml or normal saline. Balloons deflated and device removed after 12 hours. If Bishop score still ≤ 4 at that time, then patient given PGE_2. Dates: NR Location: Haifa, Israel Setting: Community hospital	No. of subjects at start: 95 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 95 Inclusion criteria: Indication for induction; Bishop score ≤ 4; not in labor; singleton pregnancy; vertex presentation; intact membranes Exclusion criteria: Placenta previa; abnormal fetal monitoring; previous C-section Age (mean ± SD): PGE_2, 28.5 ± 5.2; oxytocin, 27.8 ± 5.7; ARD, 27.3 ± 4.2 Race: NR Gestational age at entry (mean ± SD): PGE_2, 38.8 ± 2.0 weeks; oxytocin, 39.6 ± 1.7 weeks; ARD, 40.0 ± 1.6 weeks Dating criteria: NR Parity: PGE_2, 57% primipara; oxytocin, 57 primipara; ARD, 54% primipara Bishop score (median, with range): 2 (0-4) all three groups Other: Indications for induction: Pregnancy-induced hypertension: 45% Postterm: 18% Diabetes mellitus: 7% Fetal growth restriction: 7%	1) Neonatal outcomes 2) Cervical dilation ≥ 3 cm at 12 hours 3) Failure of primary method 4) Time from induction to delivery 5) Success rate for vaginal delivery 6) C-sections	1) Neonatal outcomes: No quantitative data reported. Simply stated that neonatal outcome was "the same" for all 3 methods with respect to mean weight, Apgar scores at 1 and 5 minutes, and perinatal morbidity. 2) Cervical dilation ≥ 3 cm at 12 hours: PGE_2: 15/30 (50%) Oxytocin: 7/30 (23%) ARD: 30/35 (86%) $p < 0.01$ for ARD vs. PGE_2 and ARD vs. oxytocin 3) Failure of primary method: PGE_2: 6/30 (20%) Oxytocin: 16/30 (53%) ARD: 2/35 (6%) $p < 0.01$ for PGE_2 vs. oxytocin and ARD vs. oxytocin 4) Time from induction to delivery (mean ± SD): PGE_2: 23.2 ± 12.5 hours Oxytocin: 28.2 ± 14.7 hours ARD: 21.3 ± 7.0 hours p = not significant 5) Success rate for vaginal delivery: PGE_2: 21/30 (70%) Oxytocin: 8/30 (27%) ARD: 27/35 (77%) $p < 0.01$ for PGE_2 vs. oxytocin and ARD vs. oxytocin 6) C-sections: *Among patients successful with primary induction method:* PGE_2: 3/24 (13%) Oxytocin: 6/14 (43%) ARD: 6/33 (18%) $p < 0.05$ for PGE_2 vs. oxytocin and ARD vs. oxytocin	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (18% of total study population). Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Not specified Length of follow-up: None	Elective induction: 6% Nonreassuring NST: 6% Fetal death: 3% Other: 6%		*Among patients not successful with primary induction method:* PGE_2: 1/6 (17%) Oxytocin: 8/16 (50%) ARD: 1/2 (50%) (no p-value reported)	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Bell, Permezel, MacLennan, et al., 1993	Design: RCT, randomization by list of random numbers Interventions: 1) Relaxin gel (recombinant human, 1.5 mg) (n = 18) Protocol: Relaxin gel inserted into the posterior vaginal fornix on evening before scheduled induction. Patient remained recumbent for 1 hour. Spontaneous uterine activity, FHR, and maternal observations monitored overnight. If no labor after 15 hours, then induction protocol begun. This included surgical rupture of the membranes and IV administration of oxytocin at different dose schedules, according to the accepted regimen at each hospital. 2) Placebo gel (n = 22) Protocol: Same as above, except placebo gel used instead of relaxin. Dates: NR Location: Melbourne, Adelaide, and Clayton, Australia Setting: 4 hospitals of unspecified type Type(s) of providers: NR Length of follow-up: 6 weeks (relaxin levels and infant weight measured)	No. of subjects at start: 40 Dropouts: 0 Loss to follow-up: NR No. of subjects at end: NR (for 6-week follow-up) Inclusion criteria: Good maternal health, uncomplicated singleton pregnancy; gestational age 40-43 weeks; scheduled for induction for postdates pregnancy; cephalic presentation; unscarred uterus; maternal height > 1.5 m; normal blood pressure; no current medication Exclusion criteria: Abnormal placental location; antepartum hemorrhage; ruptured membranes; Calder score > 6 (modified Bishop score); fetal malformation; abnormal FHR tracing; IUGR; macrosomia; reduced AFV Age (mean ± SD): Relaxin, 25.7 ± 4.5; placebo, 27.3 ± 4.4 Race: NR Gestational age at entry (mean ± SD): Relaxin, 41.2 ± 0.4 weeks; placebo, 41.4 ± 0.7 weeks Dating criteria: NR Parity: Relaxin, 56% primiparas; placebo, 59% primiparas Bishop score: NR	1) Stillbirths 2) Neonatal deaths 3) Abnormal FHR tracings warranting intervention 4) Apgar scores at 1, 5, and 10 minutes 5) Cord blood gases 6) Birthweight 7) Forceps-assisted deliveries 8) C-sections 9) Time to delivery 10) Duration of labor	1) Stillbirths: None in either group. 2) Neonatal deaths: None in either group. 3) Abnormal FHR tracings warranting intervention: Relaxin: 7/18 (39%) Placebo: 7/22 (32%) p = not significant 4) Apgar scores at 1, 5, and 10 minutes: No statistically significant differences between two groups (no quantitative data reported) 5) Cord blood gases: No statistically significant differences between two groups (no quantitative data reported) 6) Birthweight (mean ± SD): Relaxin: 3634 ± 403 g Placebo: 3673 ± 310 g p = 0.73 7) Forceps-assisted deliveries Relaxin: 6/18 (33.3%) Placebo: 6/22 (27.3%) p = not significant 8) C-sections: Relaxin: 2/18 (11.1%) Placebo: 4/22 (18.2%) p = not significant 9) Time to delivery (mean ± SD): Relaxin: 23.6 ± 4.8 hours Placebo: 24.8 ± 4.8 hours p = 0.33 10) Duration of labor (mean ± SD): Relaxin: 7.1 ± 3.4 hours Placebo: 7.5 ± 3.4 hours	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + First trial ever conducted of recombinant human relaxin in pregnant women. Low dose used deliberately. Primarily interested in establishing safety in pregnant women. Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Other: Calder score: Score Relaxin Placebo ≤ 4 33% 32% 5 50% 41% 6 17% 27%		p = 0.49	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Berghella, Rogers, and Lescale, 1996	Design: RCT, randomization by computer-generated random number table and sealed envelopes Interventions: 1) Stripping of the membranes (n = 73) Protocol: Stripping of the membranes performed weekly starting at 38 weeks by separating an approximately 2-3-cm section the lower membranes from its cervical attachment with at least two circumferential passes of the index finger. 2) Cervical exam (control) (n = 69) Protocol: "Gentle cervical examination" performed weekly starting at 38 weeks. Dates: Jul - Oct 1991 and Jul - Oct 1993 Location: New York, NY Setting: Outpatient clinic/ physician office Type(s) of providers: General OB/GYN Length of follow-up: NA	No. of subjects at start: 149 Dropouts: 7 (excluded at 38 weeks due to long, closed cervices not amenable to stripping) Loss to follow-up: NA No. of subjects at end: 142 Inclusion criteria: First presented to clinic at gestational age ≤ 20 weeks Exclusion criteria: Multiple pregnancy; placenta previa; low-lying placenta; nonvertex presentation; IUGR; any medical complication of pregnancy; long, closed cervix not amenable to stripping at time of intervention (38 weeks) Age (mean ± SD): Stripping, 27.19 ± 6.1; control, 27.12 ± 5.6 Race: 100% Asian Gestational age at entry: 38 weeks Dating criteria: Pelvic exam during first 12 menstrual weeks to confirm size appropriate for dates and/or U/S before 20th week Parity: Stripping, 48% nulliparas; control, 62% nulliparas (p = not significant) Bishop score (mean ± SD): Stripping, 3.49 ± 2.7; control, 2.46 ± 2.3	1) Delivery after 41 weeks 2) Vacuum-assisted delivery 3) Forceps-assisted delivery 4) C-sections 5) Days to delivery (overall and broken down by Bishop score and parity)	1) Delivery after 41 weeks: Stripping: 4/73 (5%) Control: 15/69 (22%) p < 0.01 2) Vacuum-assisted delivery: Stripping: 2/73 (3%) Control: 3/69 (4%) p = not significant 3) Forceps-assisted delivery: Stripping: 5/73 (7%) Control: 4/69 (6%) p = not significant 4) C-sections: Stripping: 0/73 Control: 3/69 (4%) p = not significant 5) Days to delivery (mean ± SD): *Overall:* Stripping: 8.2 ± 6.3 Control: 12.2 ± 7.1 p < 0.002 *Broken down by Bishop score:* Bishop score ≤ 3: Stripping (n =39): 8.6 ± 6.4 Control (n = 44): 12.5 ± 6.8 p ≤ 0.02 Bishop score > 3: Stripping (n = 34): 6.5 ± 5.4 Control (n = 25): 11.5 ± 8.2 p = 0.10 *Broken down by parity:* Nulliparas: Stripping (n = 35): 7.8 ± 6.0 Control (n = 43): 12.9 ± 6.6 p < 0.09 Multiparas: Stripping (n = 38): 7.2 ± 5.9	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on proportion of patients delivering at ≥ 41 weeks. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Control (n = 26): 11.0 ± 7.9 p = 0.10	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Boulvain, Fraser, Marcoux, et al., 1998	Design: RCT, randomization by computer-generated list of random numbers and sealed envelopes Interventions: 1) Sweeping of the membranes (n = 99) Protocol: Sweeping performed using circular movements of examining finger between the lower segment of the uterus and the fetal membranes. If membranes could not be reached, then examiner attempted to dilate cervix manually. If successful, then sweeping performed; if not, then cervical massage performed. 2) Control (n = 99) Protocol: Vaginal exam performed for Bishop scoring only In both groups, post-intervention management, including method of induction and intrapartum interventions, were left to the discretion of the treating obstetrician. Dates: Apr 1995 - Oct 1996 Location: 3 sites in the province of Quebec, Canada Setting: 3 university hospitals Type(s) of providers: General OB/GYN	No. of subjects at start: 200 Dropouts: 2 Loss to follow-up: NA No. of subjects at end: 198 Inclusion criteria: Medical indication for nonurgent induction; gestational age ≥ 266 days; single fetus; cephalic presentation Exclusion criteria: None specified Age (mean ± SD): Sweeping, 28.5 ± 5.5; control, 29.2 ± 4.6 Race: NR Gestational age at entry (mean ± SD): Sweeping, 281.9 ± 5.0 days; control, 281.5 ± 4.5 days Dating criteria: LMP plus 2nd trimester U/S Parity: Sweeping, 58% nulliparous; control, 49% nulliparous Bishop score (mean ± SD): Sweeping, 5.8 ± 2.2; control, 5.3 ± 2.3 Other: Indications for induction: Postterm (> 287 days): 85% Hypertension: 4% Diabetes: 2.5% IUGR: 1.5% Other: 7%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Admission to NICU 5) Neonatal infection 6) Cephalhematoma 7) Convulsions 8) Respiratory distress 9) Induction of labor 10) Fever during labor or postpartum 11) Forceps/vacuum delivery 12) C-sections 13) Time from randomization to onset of labor	1) Apgar score < 7 at 1 minute: Sweeping: 5/99 Control: 8/99 p = 0.40 2) Apgar score < 7 at 5 minutes: Sweeping: 3/99 Control: 0/99 p = 0.25 3) Birthweight (mean ± SD): Sweeping: 3501 ± 436 g Control: 3633 ± 438 g p = 0.04 4) Admission to NICU: Sweeping: 6/99 Control: 6/99 p = 1.00 5) Neonatal infection: Sweeping: 1/99 Control: 1/99 p = 1.00 6) Cephalhematoma: Sweeping: 5/99 Control: 2/99 p = 0.44 7) Convulsions: Sweeping: 1/99 Control: 0/99 p = 1.00 8) Respiratory distress: Sweeping: 0/99 Control: 1/99 p = 1.00 9) Induction of labor: Sweeping: 49/99 Control: 59/99 p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (85% of total study population). Positive effect in multiparas with Bishop score > 6 (RR, 0.55; 95% CI, 0.31-0.98), but not in other groups. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None			10) Fever during labor or postpartum: Sweeping: 8/99 Control: 8/99 p = not significant 11) Forceps/vacuum delivery: Sweeping: 36/99 Control: 27/99 (no p-value reported) 12) C-sections: Sweeping: 12/99 Control: 12/99 p = 0.37 13) Time from randomization to onset of labor (mean): Sweeping: 76 hours Control: 98 hours p = 0.01	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Brennand, Calder, Leitch, et al., 1997	Design: RCT, randomized by computer-generated list Interventions: 1) 4 mg recombinant human relaxin (n = 25) given between 37 and 42 weeks gestation. Gel introduced into posterior fornix; NST monitored for 4 hours post-treatment, then every 4 hours for 24 hours or until delivery. 2) 2 mg relaxin (n = 25), given in same manner 3) 1 mg relaxin (n = 23), given in same manner 4) Placebo gel (n = 23), given in same manner In all groups, induction started by placing 2 mg PGE_2 gel intravaginally 15 hours after relaxin, amniotomy ± additional PGE_2 Dates: NR Location: Edinburgh, Glasgow, Manchester, and Oxford, UK Setting: University hospitals Providers: Unspecified OB/GYN Length of follow-up: None	No of subjects at start: 96 Drop-outs: 0 Loss to follow-up: NA No of subjects at end: 96 Inclusion criteria: Gestational age ≥ 37 weeks, Bishop score ≤ 4 Exclusion criteria: Uterine scar; ruptured membranes; evidence of placental abruption or previa; systemic disease; recent ingestion of NSAIDs; fetal malformation; abnormalities in fetal growth, size, or amniotic fluid volume Age (mean): 4 mg: 25.8 2mg: 26.7 1 mg: 26.8 Placebo: 27.0 Race: NR Gestational age at entry: 4 mg: 40.1 weeks 2 mg: 39.9 1 mg: 39.6 Placebo: 40.0 Dating criteria: NR Parity (% nulliparous): 4 mg: 76% 2 mg: 88% 1 mg: 87% Placebo: 78% Bishop score (mean): 4 mg: 2.5 2 mg: 2.8	1) Change in Bishop score between baseline and 15 hours 2) Spontaneous labor 3) Treatment to delivery 4) Cesarean delivery 5) Perinatal morbidity/mortality	1) Change in Bishop score between baseline and 15 hours: 4 mg: 1.32 2 mg: 1.76 1 mg: 1.36 Placebo: 1.64 p = 0.85 2) Spontaneous labor: 4 mg: 2/25 2 mg: 5/25 1 mg: 1/23 Placebo: 2/23 p = 0.93 3) Treatment to delivery (mean): 4 mg: 36.7 hours 3 mg: 39.3 hours 1 mg: 29.9 hours Placebo: 28.0 hours p = 0.31 3) Cesarean delivery: 4 mg: 4/25 3 mg: 8/25 1 mg: 3/23 Placebo: 4/23 p = 0.45 4) Perinatal morbidity/mortality: No deaths in any group. No significant differences reported except higher baseline fetal heart rates in all relaxin groups compared to placebo.	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy. Study underpowered to detect differences in important outcomes. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		1 mg: 3.0 Placebo: 2.9 Other: Indications for induction: "Most" pregnancy-induced hypertension or prolonged pregnancy; numbers not given			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Buser, Mora, and Arias, 1997	Design: RCT, randomization by random numbers table and sealed envelopes Interventions: 1) Misoprostol (n = 76) Protocol: 50-µg tablet placed in posterior vaginal fornix using a speculum. Dose repeated every 4 hours until patient developed an adequate contraction pattern (≥ 3 contractions in 10 minutes), cervix reached ≥ 3 cm dilation and 100% effacement, or SROM occurred. Maximum of 3 doses. Oxytocin augmentation started 4 hours after last dose if adequate pattern of contraction still not obtained. 2) PGE_2 (n = 79) Protocol: PGE_2 gel (0.5 mg) administered intracervically using a speculum. Dose repeated every 6 hours until patient developed an adequate contraction pattern (≥ 3 contractions in 10 minutes), cervix reached ≥ 3 cm dilation and 100% effacement, or SROM occurred. Maximum of 3 doses. Oxytocin augmentation started 6 hours after last dose if adequate pattern of contraction still not obtained. Dates: July 1994 - Dec 1995 Location: St. Louis, MO	No. of subjects at start: 155 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 155 Inclusion criteria: Admitted for induction; singleton pregnancy at term; cephalic presentation; reassuring FHR tracing; Bishop score ≤ 5 Exclusion criteria: Ruptured membranes; low-lying placenta; partial or complete placenta previa; prior C-section; parity ≥ 6; strong clinical suspicion of fetopelvic disproportion; history of asthma, glaucoma, or cardiac disease Age (mean ± SD): Misoprostol, 27.7 ± 5.6; PGE_2, 27.1 ± 5.8 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.2 ± 1.9 weeks; PGE_2, 39.3 ± 1.8 weeks Dating criteria: NR Parity: Misoprostol, 84% nulliparas; PGE_2, 82% nulliparas Bishop score (mean ± SD): Misoprostol, 2.66 ± 1.3; PGE_2, 2.64 ± 1.4 Other: Indications for induction: Postterm: 35% Preeclampsia: 28%	1) Apgar score < 6 at 5 minutes 2) Birthweight 3) Admission to NICU 4) Number of days in NICU 5) Nonreassuring FHR tracing with hyper-stimulation 6) Change in Bishop score 7) Time from induction to delivery 8) C-sections 9) Spontaneous vaginal delivery	1) Apgar score < 6 at 5 minutes: Misoprostol: 2/76 (3%) PGE_2: 0/79 p = not significant 2) Birthweight (mean ± SD): Misoprostol: 3435 ± 564 g PGE_2: 3383 ± 618 g p = not significant 3) Admission to NICU: Misoprostol: 7/76 (9%) PGE_2: 0/79 p = not significant 4) Number of days in NICU (mean): Misoprostol: 14 days PGE_2: 13 days p = not significant 5) Nonreassuring FHR tracing with hyper-stimulation: Misoprostol: 14/76 (18%) PGE_2: 0/79 p < 0.001 6) Change in Bishop score (mean ± SD): Misoprostol: 3.53 ± 2.1 PGE_2: 2.7 ± 1.8 p = 0.01 7) Time from induction to delivery (mean ± SD): Misoprostol: 15.8 ± 7.0 hours PGE_2: 24.2 ± 11.0 hours p < 0.01 8) C-sections: *Overall:* Misoprostol: 27/76 (36%) PGE_2: 17/79 (22%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (35% of total study population, unevenly distributed: 41% of miso-prostol group, 29% of PGE_2 group [p = not significant, but study underpowered]). Sample size estimates based on change in Bishop score, active labor, and C-section rate. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Setting: Community hospital Type(s) of providers: Not specified Length of follow-up: None	Decreased amniotic fluid: 10% Large for gestational age: 10% Gestational diabetes: 3% Fetal growth restriction: 3% Other: 11%		*For nonreassuring FHR tracing:* Misoprostol: 19/76 (25%) PGE_2: 4/79 (5%) $p < 0.001$ 9) Spontaneous vaginal delivery: Misoprostol: 25/76 (33%) PGE_2: 37/79 (47%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Buttino and Garite, 1990	Design: RCT, randomization performed by dispensing pharmacy Interventions: 1) PGE_2 gel (0.5 mg) (n = 23) Protocol: Patient underwent CST/NST, which had to be negative/reactive before treatment administered. PGE_2 gel placed intra-cervically using a syringe. Patient observed on external fetal monitor for 1 hour and then allowed to go home. 2) Placebo (n = 20) Protocol: Same as above, except that placebo gel used in place of PGE_2. Dates: NR Location: Long Beach, CA Setting: Unspecified hospital Type(s) of providers: NR Length of follow-up: NA	No. of subjects at start: 43 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 43 Inclusion criteria: Gestational age ≥ 41-6/7 weeks (279 days); no contraindications to prostaglandins Exclusion criteria: None stated Age (mean): PGE_2, 24.9; placebo, 25.8 Race: NR Gestational age at entry (mean): PGE_2, 42.3 weeks; placebo, 42.5 weeks Dating criteria: Any two of the following: LMP; 1st trimester pelvic exam consistent with dates; U/S demonstrating either a crown-rump length at 6-11 weeks or biparietal diameter and femur measurements at 17-20 weeks consistent with dates Parity: PGE_2, 43% primigravidas; placebo, 30% primigravidas (p = not significant) Bishop score (mean ± SD): PGE_2, 2.8 ± 0.8; placebo, 2.2 ± 1.3	1) Apgar scores at 1 minute 2) Apgar scores at 5 minutes 3) Birthweight 4) Time to delivery 5) Duration of labor 6) Change in Bishop score 7) C-sections	1) Apgar scores at 1 minute (mean ± SD): PGE_2: 7.8 ± 1.1 Placebo: 8.2 ± 0.8 p = not significant 2) Apgar scores at 5 minutes (mean ± SD): PGE_2: 8.9 ± 0.3 Placebo: 9.0 ± 0.2 p = not significant 3) Birthweight (mean ± SD): PGE_2: 3644.6 ± 416.7 g Placebo: 3840.8 ± 574.4 p = not significant 4) Time to delivery (mean ± SD): PGE_2: 311.2 ± 244.8 hours Placebo: 379.6 ± 186.7 hours p = not significant 5) Duration of labor (mean ± SD): PGE_2: 10.6 ± 6.9 hours Placebo: 9.0 ± 4.2 hours p = not significant 6) Change in Bishop score (mean ± SD): PGE_2: 3.8 ± 2.3 Placebo: 3.0 ± 2.3 p = not significant 7) C-sections: PGE_2: 5/23 (21.7%) Placebo: 7/20 (35.0%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: + Dating criteria: + Bishop score: + Underpowered to detect differences either at baseline or at outcome time points. Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Cammu and Haitsma, 1998	Design: RCT, randomization by computer-generated list of random numbers and sealed envelopes Interventions: 1) Sweeping of the membranes (n = 140) Protocol: Sweeping of the membranes performed weekly beginning at 39 completed weeks. Digital separation of 2-3 cm of the membranes from the lower uterine segment performed, rotating the finger at least twice through 360 degrees. Closed cervix stretched digitally until membrane sweeping could be carried out. Closed cervix that would not admit a finger was vigorously massaged. 2) Control (n = 138) Protocol: Routine pelvic exam performed weekly beginning at 39 completed weeks. In both groups, induction planned from 41 completed weeks onward and performed according to standard protocol (amniotomy ± oxytocin, with cervical ripening beforehand, if necessary). Dates: NR (patients enrolled over a 25-month period) Location: Brussels, Belgium Setting: Antenatal clinic of university hospital	No. of subjects at start: 287 Dropouts: 9 Loss to follow-up: NA No. of subjects at end: 278 Inclusion criteria: Gestational age 39 weeks; nulliparous; singleton fetus; cephalic presentation; no risk factors Exclusion criteria: None specified Age (mean ± SD): Sweeping, 27.6 ± 3.8; control, 27.6 ± 4.0 Race: NR; clinic said to serve "mostly urban middle class Caucasian women" Gestational age at entry (mean ± SD): 273.3 ± 2.4 days; 273.2 ± 2.5 days Dating criteria: U/S (not specified whether 1st or 2nd trimester) Parity: 100% nulliparous Bishop score (mean ± SD): Sweeping, 3.35 ± 1.8; control, 3.39 ± 1.6	1) Apgar score < 7 at 5 minutes 2) Arterial cord blood pH < 7 3) Birthweight 4) Gestational age at delivery 5) Induction of labor 6) Instrumental delivery 7) C-sections 8) Time from randomization to delivery	1) Apgar score < 7 at 5 minutes: Sweeping: 3/140 (2%) Control: 5/138 (4%) p = 0.490 2) Arterial cord blood pH < 7: Sweeping: 7/140 (5%) Control: 8/138 (6%) p = 0.976 3) Birthweight (mean ± SD): Sweeping: 3400 ± 375 g Control: 3459 ± 411 g p = not significant 4) Gestational age at delivery: *Mean ± SD:* Sweeping: 282.8 ± 5 days Control: 283.8 ± 6 days p = not significant *Percentage > 287 days:* Sweeping: 27/140 (19%) Control: 45/138 (33%) OR = 0.49 (95% CI, 0.29-0.86) 5) Induction of labor: Sweeping: 15/140 (11%) Control: 36/138 (26%) OR = 0.34 (95% CI, 0.18-0.66) 6) Instrumental delivery: Sweeping: 23/140 (16%) Control: 18/138 (13%) OR = 1.31 (95% CI, 0.67-2.55) 7) C-sections: Sweeping: 5/140 (4%) Control: 8/138 (6%) OR = 0.60 (95% CI, 0.19-1.89) 8) Time from randomization to delivery (mean ± SD): Sweeping: 9.4 ± 5 days	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + 24/140 women in the membrane-sweeping group (17%) had cervixes inaccessible to an examining finger and received cervical massage only. These women were not excluded from the analysis. Sample size estimates based on proportion of patients reaching 41 weeks. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: NR Length of follow-up: None			Control: 10.6 ± 6 days (no p-value reported)	
Chang and Chang, 1997	Design: RCT, method of randomization not described Interventions: 1) PGE_2 (n = 30) Protocol: 3-mg tablet placed in posterior vaginal fornix. Dose repeated every 6 hours until satisfactory uterine activity achieved. Maximum dose permitted was 9 mg. 2) Misoprostol (n = 30) Protocol: 50-µg tablet placed in posterior vaginal fornix. Dose repeated every 4 hours until satisfactory uterine activity achieved. Maximum dose permitted was 600 µg. In both groups, oxytocin augmentation initiated if Bishop score ≥ 9, but uterine contractions inadequate (< 3 per 10 minutes). Dates: July 1994 - June 1995 Location: Tainan, Taiwan Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: Dropouts: Loss to follow-up: No. of subjects at end: Inclusion criteria: Scheduled for induction; term singleton pregnancy; Bishop score ≤ 5; no regular uterine contractions Exclusion criteria: Contra-indications to vaginal prostaglandins; any maternal illness for which induction of labor not appropriate Age (mean ± SD): PGE_2, 28.9 ± 5.3; misoprostol, 27.6 ± 6.7 Race: NR Gestational age at entry (mean ± SD): PGE_2, 39.3 ± 2.4 weeks; misoprostol, 38.9 ± 3.1 weeks Dating criteria: NR Parity: 100% nulliparous in both groups Bishop score (mean ± SD): PGE_2, 4.3 ± 1.1; misoprostol, 4.2 ± 0.5 Other: Indications for induction: Excess maternal weight gain (> 16 kg): 42% Postterm: 40% Hypertension: 18%	1) Apgar scores < 7 at 1 and 5 minutes 2) Birthweight 3) Cord arterial pH 4) Time from induction to delivery 5) Hyperstimulation 6) Vacuum extractions 7) C-sections	1) Apgar scores < 7 at 1 and 5 minutes: No quantitative data reported. Simply stated that proportion of neonates with Apgar ≤ 7 at 1 and 5 minutes was "the same" in both groups. 2) Birthweight (mean ± SD): PGE_2: 3376 ± 432 g Misoprostol: 3285 ± 580 g p = not significant 3) Cord arterial pH (mean ± SD): PGE_2: 7.32 ± 0.91 Misoprostol: 7.29 ± 0.73 p = not significant 4) Time from induction to delivery (mean ± SD): PGE_2: 25.7 ± 3.8 hours Misoprostol: 16.5 ± 2.7 hours p < 0.001 5) Hyperstimulation: PGE_2: 8.9% Misoprostol: 13.4% p < 0.05 6) Vacuum extractions: PGE_2: 6% Misoprostol: 10% p = not significant 7) C-sections: PGE_2: 6% Misoprostol: 10% p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (40% of total study population).

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chatterjee, Ramchandran, Ferlita, et al., 1991	Design: RCT, randomization by card shuffling Interventions: 1) 2 mg PGE_2 gel applied in posterior fornix (n = 15) 12 hours prior to induction with oxytocin 2) Placebo gel (n = 18) In both groups, second application possible if induction unsuccessful. Dates: Jul 1983 - Apr 1984 Location: Newark, NJ Setting: University hospital Providers: Unspecified OB/GYN Length of follow-up: None	No of subjects at start: 38 Dropouts: 0 Loss to follow-up: NA No of subjects at end: 38 Inclusion criteria: NR Exclusion criteria: NR Age (mean ± SD): PGE_2: 24.2 ± 1.1 Placebo: 25.1 ± 1.3 Race: NR Gestational age at entry: PGE_2: 39.1 ± 0.5 Placebo: 38.4 ± 0.9 Dating criteria: NR Parity: NR Bishop score: NR Other: 18% induced for prolonged pregnancy	1) Change in Bishop score 2) Cesarean section 3) Mean Apgar score at 1 minute 4) Mean Apgar score at 5 minutes	1) Change in Bishop score: Data presented graphically; statistically significant greater change with PGE_2 ($p < 0.01$). 2) Cesarean section: PGE_2: 7/15 Placebo: 5/18 3) Mean Apgar score at 1 minute: PGE_2: 6.8 Placebo: 6.8 4) Mean Apgar score at 5 minutes: PGE_2: 7.9 Placebo: 8.1	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: +

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chayen, Tejani, and Verma, 1986	Design: RCT, allocation to treatment group by even/odd hospital ID number Interventions: 1) Nipple stimulation using breast pump (n = 30) Protocol: Patients admitted to labor ward, placed on an external monitor, and assigned a Bishop score. Vaseline applied to nipple. Breast pump turned on to normal setting (250 mmHg of negative pressure). Pump alternated from right to left breast every 15 minutes. Once regular contractions occurred and cervix ≥ 2 cm dilated, then patient under-went amniotomy and had internal pressure catheter placed. If active phase not reached or active phase arrested, then patient switched to oxytocin protocol. 2) Induction using oxytocin (control) (n = 32) Protocol: Patients admitted to labor ward, placed on an external monitor, and assigned a Bishop score. Induction initiated with 2 µm/min of oxytocin, with gradual increments until "adequate uterine activity" (≥ 200 Montevideo units) achieved. Once regular contractions occurred and cervix ≥ 2 cm dilated, then patient underwent amniotomy and had internal pressure catheter placed. Patients who	No. of subjects at start: 62 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 62 Inclusion criteria: Admitted for induction of labor Exclusion criteria: None specified Age: NR Race: NR Gestational age at entry (mean ± SD): Breast pump, 39.31 ± 2.33 weeks, 9/30 (30%) "postdates"; oxytocin, 40.18 ± 1.90 weeks, 8/32 (25%) "postdates" Dating criteria: NR Parity: Breast pump, 43% nulliparous; oxytocin, 53% nulliparous Bishop score (mean ± SD): Breast pump, 5.48 ± 1.87; oxytocin, 6.62 ± 1.77 (p = 0.05) Other: Indications for induction: Preeclampsia: 44% Postterm: 29% Other: 27%	1) Failure to reach active phase 2) Time to regular contractions 3) Time to adequate labor 4) Time to active phase 5) C-sections	1) Failure to reach active phase: Breast pump: 3/30 (10%) Oxytocin: 4/32 (12.5%) p = not significant 2) Time to regular contractions (mean ± SD): Breast pump: 5.68 ± 6.13 minutes Oxytocin: 61.55 + 42.62 minutes p = 0.0005 3) Time to adequate labor (mean ± SD): Breast pump: 1.52 ± 1.075 hours Oxytocin: 3.41 ± 2.22 hours p = 0.0005 4) Time to active phase (mean ± SD): Breast pump: 4.84 ± 3.33 hours Oxytocin: 6.90 ± 4.21 hours p = 0.05 5) C-sections: Breast pump: 8/30 (26.7%) Oxytocin: 14/32 (43.7)% p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (29% of total study population). Significant difference in baseline Bishop scores – bias in favor of oxytocin. Results not stratified by parity. Study underpowered to detect difference at baseline or in outcomes. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	failed induction delivered by C-section. Dates: NR Location: Stony Brook, NY Setting: University hospital; community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None				

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Chuck and Huffaker, 1995	Design: RCT, randomization by computer and sealed envelopes Interventions: 1) Misoprostol (n = 49) Protocol: 50-µg tablet placed in posterior vaginal fornix. Additional doses given every 4 hours for a maximum of 5 doses. 2) PGE_2 (n = 50) Protocol: Gel (0.5 mg) placed intracervically. Additional doses given every 4 hours for a maximum of 5 doses. In both groups, dosing halted for hyperstimulation or if patient having ≥ 3 contractions/10 minutes. Oxytocin used if no labor after maximum dose or if labor progress arrested for > 2 hours. AROM performed when cervix > 3 cm. Dates: Sep 1993 - Jan 1994 Location: Los Angeles, CA Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 103 Dropouts: 4 (excluded from analysis due to protocol violations) Loss to follow-up: NA No. of subjects at end: 99 Inclusion criteria: Gestational age 35-42 weeks; admitted for induction of labor Exclusion criteria: Nonvertex presentation; uterine scar other than from prior low-transverse C-section; ominous FHR tracing; multiple gestation; complete cervical effacement Age (mean ± SD): Misoprostol, 29.3 ± 6.7; PGE_2, 28.7 ± 6.4 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 29.7 ± 1.7 weeks; PGE_2, 39.7 ± 1.3 weeks Dating criteria: NR Parity (mean ± SD): Misoprostol, 0.8 ± 0.9 (52% nulliparous); PGE_2, 0.8 ± 0.9 (48% nulliparous) Bishop score: Misoprostol, 53% ≤ 3; PGE_2, 52% ≤ 3 Other: Indications for induction: PROM: 28% Postterm: 18% Diabetes mellitus: 17% Oligohydramnios: 10% Hypertensive disorders: 10%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Admission to NICU 5) Meconium 6) Time to (vaginal) delivery 7) Vaginal deliveries within 24 hours 8) Cost of study medication 9) Time to vaginal delivery 10) Vaginal delivery within 24 hours	1) Apgar score < 7 at 1 minute: Misoprostol: 6/49 (12%) PGE_2: 4/50 (8%) p = 0.525 2) Apgar score < 7 at 5 minutes: Misoprostol: 0/49 PGE_2: 0/50 p = not significant 3) Birthweight (mean ± SD): Misoprostol: 3326.8 ± 529.7 g PGE_2: 3331.4 ± 509.7 g p = 0.965 4) Admission to NICU: Misoprostol: 0/49 PGE_2: 0/50 p = not significant 5) Meconium: Misoprostol: 4/49 (8%) PGE_2: 5/50 (10%) p = 0.950 6) Time to (vaginal) delivery (mean ± SD): Misoprostol (n = 39): 11.4 ± 5.9 hours PGE_2 (n = 40): 18.9 ± 12.7 hours p = 0.001 7) Vaginal deliveries within 24 hours: Misoprostol: 39/39 (100%) PGE_2: 27/40 (68%) p = 0.001 8) Cost of study medication: Misoprostol: $0.20 per dose PGE_2: $65 per kit (no p-value reported) 9) Time to vaginal delivery (mean ± SD):	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (18% of total study population). Sample size estimates based on time to delivery. Study underpowered to detect differences at baseline and for some outcomes – e.g.: 1) Nulliparous with Bishop score ≤ 3: 61% misopostol, 48% PGE_2; p = not significant, but study insufficiently powered. Bias against misoprostol. 2) Prior C-section: 10% misoprostol, 20% PGE_2; bias in favor of misoprostol. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Nonreassuring FHR: 8% IUGR: 5% Other: 4%		*Among nulliparas:* Misoprostol (n = 16): 14.4 ± 6.5 hours PGE_2 (n = 16): 26.7 ± 14.3 hours p = 0.004 *Among multiparas:* Misoprostol (n = 23): 9.4 ± 4.7 hours PGE_2 (n = 24): 13.8 ± 8.3 hours p = 0.032 10) Vaginal delivery within 24 hours Misoprostol: 39/39 (100%) PGE_2: 27/40 (68%) p = 0.001	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Crane, Bennett, Young, et al., 1997	Design: RCT, randomization by computer-generated random numbers and sealed envelopes; stratified by status of cervix at initial exam Interventions: 1) Sweeping of membranes (n = 76) Protocol: "As much membrane as possible" separated from lower segment by circumferential sweeping of examining finger two times. Performed between 38 and 40 weeks. "Vigorous" massage by rubbing external os in circular manner if cervix closed. 2) Control exam only (n = 74) Dates: NR Location: Newfoundland, Canada Setting: University hospital (antenatal clinic) Type(s) of providers: NR Length of follow-up: None	No. of subjects at start: 150 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 150 Inclusion criteria: Low-risk pregnancy; gestational age 38-40 weeks Exclusion criteria: Medical disease; pregnancy complications; fetal growth restriction; history of perinatal mortality or low birthweight infant; PROM; abnormal presentation; placenta previa; scheduled cesarean section; other contraindications to vaginal delivery Age (mean ± SD): Sweeping, 27.9 ± 4.8; control, 28.3 ± 4.4 Race: 95% white Gestational age at entry: Sweeping, 39.7 weeks; control, 39.5 weeks Dating criteria: "Firm" LMP or ultrasound prior to 18 weeks Parity: Sweeping: median, 0; 61% nulliparous; control: median, 1.0; 47% nulliparous (p = 0.10) Bishop score: Sweeping: Median, 5; 28% < 7 Control: Median, 5; 16% < 7	1) Spontaneous labor within 7 days 2) Spontaneous labor before 41 weeks 3) Spontaneous labor 4) C-section 5) Epidural 6) PROM 7) Maternal infection 8) Apgar score < 7 at 1 minute 9) Apgar score < 7 at 5 minutes	1) Spontaneous labor within 7 days: Sweeping: 33% Control: 38% p = 0.39 2) Spontaneous labor before 41 weeks: Sweeping: 45% Control: 51% p = 0.66 3) Spontaneous labor: Sweeping: 54% Control: 68% 4) C-section: Sweeping: 13% Control: 14% 5) Epidural: Sweeping: 66% Control: 43% p = 0.006 6) PROM: Sweeping: 6.6% Control: 22% p = 0.008 7) Maternal infection: Sweeping: 6.6% Control: 8.1% 8) Apgar score < 7 at 1 minute: Sweeping: 12% Control: 5.4% p = 1.0 9) Apgar score < 7 at 5 minutes: Sweeping: 0 Control: 0	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: +/- Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + No differences observed when results stratified by open cervix or by parity. More nulliparous women, with less favorable cervix, in sweeping group. Secondary multivariate analyses: Logistic regression: Bishop score < 7, gestational age at entry both predictors of spontaneous labor within 7 days. Log-rank test done for number of days to delivery: median 6.5 for sweeping, 8 for control (p = 0.88). Not clear whether study powered to detect this difference.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Doany and McCarty, 1997	Design: RCT, randomization by table of random numbers Interventions: 1) No membrane stripping + placebo gel (n = 28) Protocol: Placebo gel (4 ml) placed, via syringe, in posterior vaginal fornix. Continuous external fetal and uterine monitoring for 1 hour; if no sign of fetal distress, then patient allowed to go home (instructed to do daily kick counts). Repeat testing at 294 days and every 3-4 days after that. Treatment re-administered at each visit after obtaining reactive NST, normal AFI, and Bishop score. Patients referred to labor and delivery suite if painful contractions every 5 minutes, spontaneous amniorrhexis, decreased fetal movement, nonreactive NST, oligo-hydramnios (AFI < 5), fetal distress, hyperstimulation, or attainment of 307 days of gestation. Labor and delivery managed by appropriate staff (not part of controlled trial). 2) No membrane stripping + PGE_2 gel (n = 37) Protocol: Same as 1), above, except that PGE_2 gel (2 mg) substituted for placebo 3) Membrane stripping + placebo gel (n = 50) Protocol: For membrane stripping, examining finger introduced into the cervical	No. of subjects at start: 150 Dropouts: 7 Loss to follow-up: NA No. of subjects at end: 143 Inclusion criteria: Singleton pregnancy; cephalic presentation; referred for fetal surveillance at ≥ 287 days; reactive NST; AFI 5-25 cm; fetal weight 2500-4500 g; contractions less frequent than every 5 minutes Exclusion criteria: No prenatal care; previous uterine surgery; acute or chronic medical or psychiatric illness; drug use Age (median, with range): No stripping + placebo: 23 (19-26) No stripping + PGE_2: 23 (21-30) Stripping + placebo: 22 (19-26) Stripping + PGE_2: 25 (22-27) Race: No stripping + placebo: 100% Hispanic No stripping + PGE_2: 100% Hispanic Stripping + placebo: 94% Hispanic Stripping + PGE_2: 96% Hispanic Gestational age at entry (median, with 25-75th percentile): No stripping + placebo: 288 days (287-290) No stripping + PGE_2: 288 days (287-291) Stripping + placebo: 288 days (287-290)	1) Apgar score < 7 at 5 minutes 2) Birthweight 3) Admission to NICU 4) Probable neonatal sepsis 5) Amnionitis 6) Preeclampsia 7) Maternal hemorrhage 8) Gestational age at delivery 9) Inductions 10) Oxytocin augmenta-tion 11) Meconium 12) C-sections 13) Operative vaginal deliveries 14) Time from enrollment to delivery	1) Apgar score < 7 at 5 minutes: No stripping + placebo: 0 No stripping + PGE_2: 3% Stripping + placebo: 4% Stripping + PGE_2: 4% p = 0.99 2) Birthweight (mean [in grams] ± SD): No stripping + placebo: 3613 ± 273 No stripping + PGE_2: 3527 ± 333 Stripping + placebo: 3605 ± 365 Stripping + PGE_2: 3614 ± 479 p = 0.70 3) Admission to NICU: No stripping + placebo: 0 No stripping + PGE_2: 5% Stripping + placebo: 2% Stripping + PGE_2: 4% p = 0.70 4) Probable neonatal sepsis: No stripping + placebo: 7% No stripping + PGE_2: 11% Stripping + placebo: 6% Stripping + PGE_2: 7% p = 0.86 5) Amnionitis: No stripping + placebo: 0 No stripping + PGE_2: 11% Stripping + placebo: 10% Stripping + PGE_2: 11% p = 0.32 6) Preeclampsia: No stripping + placebo: 0 No stripping + PGE_2: 14% Stripping + placebo: 0 Stripping + PGE_2: 7% p = 0.01	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	canal and a total of 3 circumferential sweeps made between the lower uterine segment and the chorionic membranes. When cervical canal not accessible, then cervix pulled anteriorly and massaged. Rest of protocol as in 1), above. 4) Membrane stripping + PGE_2 gel (n = 28) Protocol: Membrane stripping as in 3), above. Rest of protocol as in 2), above. Dates: NR Location: Sylmar, CA Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	Stripping + PGE_2: 288 days (287-289) Dating criteria: LMP confirmed by uterine size, fetal heart tones, and U/S (no date given) Parity (% nulliparous): No stripping + placebo: 54% No stripping + PGE_2: 38% Stripping + placebo: 50% Stripping + PGE_2: 43% Bishop score (% ≤ 6): No stripping + placebo: 50% No stripping + PGE_2: 69% Stripping + placebo: 63% Stripping + PGE_2: 63%		7) Maternal hemorrhage: No stripping + placebo: 7% No stripping + PGE_2: 0 Stripping + placebo: 0 Stripping + PGE_2: 4% p = 0.05 8) Gestational age at delivery (median [in days], with 25-75th percentile): No stripping + placebo: 297 (292-302) No stripping + PGE_2: 294 (290-298) Stripping + placebo: 294 (291-298 Stripping + PGE_2: 290 (289-293) p = 0.005 9) Inductions: No stripping + placebo: 33% No stripping + PGE_2: 28% Stripping + placebo: 27% Stripping + PGE_2: 14% p = 0.42 10) Oxytocin augmentation: No stripping + placebo: 48% No stripping + PGE_2: 47% Stripping + placebo: 37% Stripping + PGE_2: 36% p = 0.65 11) Meconium: No stripping + placebo: 30% No stripping + PGE_2: 19% Stripping + placebo: 26% Stripping + PGE_2: 21% p = 0.67 12) C-sections: No stripping + placebo: 4% No stripping + PGE_2: 8% Stripping + placebo: 8% Stripping + PGE_2: 11% p = 0.08 *(continued on next page)*	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				13) Operative vaginal deliveries: No stripping + placebo: 4% No stripping + PGE_2: 3% Stripping + placebo: 18% Stripping + PGE_2: 7% (no p-value reported) 14) Time from enrollment to delivery (median [in days], with 25-75th percentile): No stripping + placebo: 7 (3.5-11.5) No stripping + PGE_2: 2 (0-7) Stripping + placebo: 4 (2-8) Stripping + PGE_2: 1 (0-4) p = 0.001	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Elliott, Brennand, and Calder, 1998	Design: RCT, randomization method not detailed but implied by computer-generated random numbers Interventions: 1) Mifepristone 50 mg (n = 25) Protocol: 50 mg given orally in women with indication for induction between 37 weeks and 41 weeks, 4 days. 2) Mifepristone 200 mg (n = 25) Protocol: Same as above, except dose 200 mg. 3) Placebo (n = 30) In all groups, patients had NST and cervical exam at 24 and 48 hours after initial dose. Induction scheduled for 72 hours after medication if no labor. Induction performed using 1 mg PGE_2 gel as initial dose, with oxytocin as clinically indicated. Dates: NR Location: Edinburgh, UK Setting: University hospital Type(s) of providers: NR Length of follow-up: None	No. of subjects at start: 80 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 80 Inclusion criteria: Single gestation; vertex presentation; Bishop score ≤ 4 Exclusion criteria: Signs and symptoms of labor; placental insufficiency; contraindications to mifepristone Age (mean ± SD): Placebo: 26.2 ± 5.9 50 mg: 25.8 ± 4.5 200 mg:25.6 ± 3.3 Race: NR Gestational age at entry (mean ± SD): Placebo: 40 weeks, 6 days (± 3.6 days) 50 mg: 40 weeks, 5 days (± 5.5 days) 200 mg: 40 weeks, 6 days (± 5.1 days Dating criteria: 1st trimester U/S Parity: 100% nulliparous Bishop score (median, with range): Placebo: 3 (1-4) 50 mg: 4 (2-4) 200 mg: 3 (1-4)	1) Proportion in spontaneous labor within 72 hours 2) Proportion with Bishop score ≥ 6 at induction 3) Time to onset of labor 4) Time to delivery 5) Fetal distress in labor requiring intervention 6) Cesarean delivery 7) Neonatal outcomes	1) Proportion in spontaneous labor within 72 hours: Placebo: 23.3% 50 mg: 32% 200 mg: 36% 2) Proportion with Bishop score ≥ 6 at induction: Placebo: 6.7% 50 mg: 16% 200 mg: 28% 3) Time to onset of labor (median): Placebo: 81 hours 15 minutes 50 mg: 80 hours 20 minutes 200 mg: 75 hours 50 minutes 4) Time to delivery (median): Placebo: 88 hours 14 minutes 50 mg: 85 hours 15 minutes 200 mg: 84 hours 6 minutes 5) Fetal distress in labor requiring intervention: Placebo: 13.3% 50 mg: 24% 200 mg: 48% 6) Cesarean delivery: Placebo: 25% 50 mg: 5% 200 mg: 38% p=0.033, Placebo vs. 50 mg p=0.075, Placebo vs. 200 mg 200 mg group: 8/9 for fetal distress, 1 for dystocia Placebo: 3/8 for fetal distress, 5 for dystocia 7) Neonatal outcomes: Jaundice: Placebo: 6.7% 50 mg: 8%	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Study underpowered to detect differences in cesarean rates, neonatal outcomes. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				200 mg: 28% Trends toward lower ACTH, higher cortisol in infants in 200 mg group	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Elliott and Flaherty, 1984	Design: RCT, randomization by table of random numbers Interventions: 1) Breast stimulation (n = 100) Protocol: Patients instructed to manually stimulate the nipple, areola, and distal breast with the balls of the fingertips, one breast at a time, for 15 minutes at a time, for 1 hour. Encouraged to do this 3 x per day (total of 3 hours per day). Re-evaluation at 42 weeks. If Bishop score ≥ 8, then labor induced. If Bishop score < 8, then CST administered. If CST reactive (negative), then further week of treatment. If CST abnormal, then labor induced. 2) Pelvic exam (control) (n = 100) Protocol: Pelvic exam given. Patients instructed to abstain from sexual intercourse and to avoid breast stimulation. Re-evaluation at 42 weeks. If Bishop score ≥ 8, then labor induced. If Bishop score < 8, then CST administered. If CST abnormal, then labor induced. If CST reactive (negative), then patient randomly assigned a second time to breast stimulation or control for further treatment. Dates: NR Location: San Francisco, CA	No. of subjects at start: 200 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 200 Inclusion criteria: Uncomplicated prenatal course; ≥ 39 weeks gestation Exclusion criteria: None specified Age (mean ± SD): Breast stimulation, 25.0 ± 4.75; control, 24.4 ± 4.88 Race: NR Gestational age at entry: NR; all subjects "approximately" 39 weeks Dating criteria: Reliable menstrual history, early pregnancy test, early vaginal estimation of uterine size, fetal heart auscultation at 20 weeks, and/or obstetric sonograms Parity (mean ± SD): Breast stimulation, 0.79 ± 1.04; control, 0.84 ± 1.10 Bishop score (mean ± SD): Breast stimulation, 4.67 ± 2.27; control, 4.15 ± 2.34	1) Apgar scores < 7 at 1 minute 2) Apgar scores < 7 at 5 minutes 3) Birthweight 4) Meconium aspiration 5) Meconium in labor 6) Inductions 7) C-sections 8) Dysmature infant 9) Death 10) Proportion of patients reaching 43 weeks with Bishop score < 8	1) Apgar scores < 7 at 1 minute: *Among women delivering at ≤ 42 weeks:* Breast stimulation: 6/95 (6%) Control: 1/83 (1%) p = not significant *Among women delivering at > 42 weeks:* Breast stimulation: 1/5 (20%) Control: 2/17 (12%) p = not significant 2) Apgar scores < 7 at 5 minutes: *Among women delivering at ≤ 42 weeks:* Breast stimulation: 1/95 (1%) Control: 0 p = not significant *Among women delivering at > 42 weeks:* Breast stimulation: 0 Control: 0 3) Birthweight (mean ± SD): Breast stimulation: 3594 ± 441 g Control: 3649 ± 394 g p = not significant 4) Meconium aspiration: Breast stimulation: 0 Control: 0 5) Meconium in labor: *Among women delivering at ≤ 42 weeks:* Breast stimulation: 25/95 (26%) Control: 22/83 (26%) p = not significant *Among women delivering at > 42 weeks:* Breast stimulation: 0 Control: 11/17 (65%) p < 0.01	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: +

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Setting: Military hospital Type(s) of providers: General OB/GYN Length of follow-up: None			6) Inductions: *Among women delivering at ≤ 42 weeks:* Breast stimulation: 6/95 (6%) Control: 6/83 (7%) p = not significant *Among women delivering at > 42 weeks:* Breast stimulation: 0 Control: 2/17 (12%) p = not significant 7) C-sections: *Among women delivering at ≤ 42 weeks:* Breast stimulation: 9/95 (9%) Control: 5/83 (6%) p = not significant *Among women delivering at > 42 weeks:* Breast stimulation: 0 Control: 5/17 (29%) p = not significant 8) Dysmature infant: Breast stimulation: 3/100 (3%) Control: 5/100 (5%) p = not significant 9) Death: Breast stimulation: 0/100 Control: 0/100 p = not significant 10) Proportion of patients reaching 43 weeks with Bishop score < 8: Breast stimulation: 5/100 Control: 17/100 p < 0.01	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Escudero and Contreras, 1997	Design: RCT, randomization by table of random numbers Interventions: 1) Misoprostol (n = 53) Protocol: Misoprostol 50 µg placed in posterior vaginal fornix. Dose repeated every 4 hours until adequate labor achieved (≥ 3 contractions of 40-50 seconds each in 10 min). Maximum total dose 350 µg. AROM performed as soon as possible. Patients with arrest of dilatation managed with oxytocin infusion, as below. 2) Oxytocin (n = 67) Protocol: Oxytocin infusion started at 4 mIU/min for 45 minutes, then increased by 2 mIU/min at 15-minute intervals up to 20 mIU/min. Dates: Sep 1994 - Mar 1995 Location: Lima, Peru Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 123 Dropouts: 3 (excluded from analysis due to protocol violations) Loss to follow-up: NA No. of subjects at end: 120 Inclusion criteria: Obstetric or medical indication for induction; no labor or fetal distress; no previous uterine; singleton pregnancy with vertex presentation; no contraindication to vaginal delivery Exclusion criteria: None specified Age (mean ± SD): Misoprostol, 27.1 ± 6.1; oxytocin, 25.5 ± 6.0 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.0 ± 2.2 weeks; oxytocin, 39.3 ± 2.1 weeks Dating criteria: NR Parity (mean ± SD): Misoprostol, 0.8 ± 1.2; oxytocin, 0.5 ± 1.0 Bishop score (mean ± SD): Misoprostol, 2.6 ± 1.5; oxytocin, 2.9 ± 1.5 Other: Indications for induction: Preeclampsia: 43% Postterm: 25% PROM: 25% Fetal demise: 4% Other: 3%	1) Apgar scores at 1 minute 2) Apgar scores at 5 minutes 3) Birthweight 4) Interval from induction to delivery 5) C-sections 6) Vaginal deliveries within 24 hours 7) Hyperstimulation 8) Any labor complication	1) Apgar scores at 1 minute (mean ± SD): Misoprostol (n = 51): 8.0 ± 1.4 Oxytocin (n = 41): 8.0 ± 1.5 p = 1.0000 2) Apgar scores at 5 minute (mean ± SD): Misoprostol (n = 51): 9.1 ± 0.9 Oxytocin (n = 41): 9.0 ± 1.3 p = 0.6646 3) Birthweight (mean ± SD): Misoprostol (n = 55): 3090.5 ± 556.9 g Oxytocin (n = 41): 3254.4 ± 493.2 g p = 0.1378 4) Interval from induction to delivery (mean ± SD): Misoprostol: 11.3 ± 6.9 hours Oxytocin: 8.4 ± 4.1 hours p = 0.0050 5) C-sections: Misoprostol : 10/57 (17.6%) Oxytocin: 4/63 (6.4%) p = 0.0560 6) Vaginal deliveries within 24 hours: Misoprostol: 45/57 (78.9%) Oxytocin: 37/63 (58.7%) p = 0.0017 7) Hyperstimulation: Misoprostol: 5/57 (8.8%) Oxytocin: 0/63 p = 0.0160 8) Any labor complication: Misoprostol: 12/57 (21.1%) Oxytocin: 5/63 (7.9%) p = 0.0400	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: - Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (25% of total study population). Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Evans, Dougan, Moawad, et al., 1983	Design: RCT, method of randomization not described Interventions: 1) Relaxin 4 mg (n = 10) Protocol: 4-mg pellet inserted into, or placed closely against, the cervix, as permitted by cervical dilatation. Cervical diaphragm placed behind the pellet to maintain its position until it dissolved (approximately 30 minutes). Patient then allowed to go home. Standard management protocol of estriols 3 times per week and NSTs 1-2 times per week was followed. If patient reached 42 weeks' gestation, then she was admitted for induction. 2) Relaxin 2 mg (n = 13) Protocol: Same as above, except that 2-mg pellet used. 3) Placebo (n = 14) Protocol: Same as above, except that placebo pellet used. Dates: NR Location: Chicago, IL Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 37 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 37 Inclusion criteria: ≥ 41 weeks gestation; scheduled to undergo oxytocin induction of labor Exclusion criteria: None specified Age (mean ± SD): Relaxin 4 mg: 26.0 ± 5.7 Relaxin 2 mg: 23.3 ± 5.4 Placebo: 21.3 ± 4.4 Race: NR Gestational age at entry (mean ± SD): Relaxin 4 mg: 41.0 ± 0.2 weeks Relaxin 2 mg: 41.2 ± 0.3 weeks Placebo: 41.1 ± 0.2 weeks Dating criteria: NR Parity (mean ± SD): Relaxin 4 mg: 1.0 ± 1.2 Relaxin 2 mg: 1.2 ± 1.1 Placebo: 1.1 ± 0.9 Bishop score: NR Other: Initial cervical coefficient (dilatation x % effacement): Relaxin 4 mg: 38.0 ± 44.5 Relaxin 2 mg: 49.6 ± 44.4 Placebo: 70.0 ± 62.6	1) Apgar scores at 5 minutes 2) Birthweight 3) Days to admission 4) Number admitted in labor 5) Time to delivery	1) Apgar scores at 5 minutes (mean ± SD): Relaxin 4 mg: 8.6 ± 1.2 Relaxin 2 mg: 9.0 ± 0.4 Placebo: 9.0 ± 0.4 p = not significant 2) Birthweight (mean ± SD): Relaxin 4 mg: 3113 ± 447 g Relaxin 2 mg: 3256 ± 613 g Placebo: 3245 ± 479 g p = not significant 3) Days to admission (mean ± SD): Relaxin 4 mg: 4.6 ± 1.6 Relaxin 2 mg: 5.3 ± 2.2 Placebo: 5.3 ± 2.1 p = not significant 4) Number admitted in labor: Relaxin 4 mg: 3/10 (30%) Relaxin 2 mg: 7/13 (54%) Placebo: 6/14 (43%) p = not significant 5) Time to delivery (mean ± SD): Relaxin 4 mg: 11.3 ± 7.2 hours Relaxin 2 mg: 7.7 ± 5.0 hours Placebo: 14.8 ± 12.2 hours p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Article describes two trials; only the trial conducted on "postdate" women abstracted here. Investigators used the "cervical coefficient" (dilatation x % effacement) instead of the Bishop score as a measure of cervical ripeness. See Hendricks CH, Brenner WE, Kraus G. Normal cervical dilatation pattern in late pregnancy and labor. Am J Obstet Gynecol 1970;106: 1065-82. Improvement in time to delivery in both nullipara and multipara.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Farah, Sanchez-Ramos, Rosa, et al., 1997	Design: RCT, randomization by computer-generated table of random numbers Interventions: 1) Misoprostol 25 µg (n = 192) Protocol: Tablet placed in posterior vaginal fornix. Dose repeated every 3 hours until adequate labor achieved (≥ 3 contractions/10 minutes). Maximum total dose 200 µg, or 8 applications. 2) Misoprostol 50 µg (n = 207) Protocol: Same as above, except maximum total dose 400 µg. In both groups, amniotomy performed as soon as cervical dilation permitted. Patients in active phase of labor with arrest of dilation and those who failed to achieve active labor after the maximum dose of misoprostol were given oxytocin. Dates: July 1994 - Sep 1995 Location: Jacksonville and Gainesville, FL Setting: 2 university hospitals Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 430 Dropouts: 31 Loss to follow-up: NA No. of subjects at end: 399 Inclusion criteria: Obstetric or medical indication for induction; Bishop score < 5; no active labor or fetal distress; no history of uterine surgery; singleton 3rd-trimester pregnancy; vertex presentation; no contraindication to vaginal delivery; no contra-indication to prostaglandins Exclusion criteria: None specified Age (mean ± SD): 25 µg, 23.8 ± 6.2; 50 µg, 23.7 ± 6.4 Race: 25 µg, 52% non-White; 50 µg, 59% non-White Gestational age at entry (mean ± SD): 25 µg, 28.9 ± 2.3 weeks; 50 µg, 38.4 ± 2.8 weeks Dating criteria: NR Parity: 25 µg, 59% nulliparous; 50 µg, 60% nulliparous Bishop score: 25 µg, 86% < 6; 50 µg, 88% < 6 Other: Indications for induction: PROM: 27% Pregnancy-induced hypertension: 22% Postterm: 14% IUGR: 8%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Cord pH < 7.6 4) Mean cord pH 5) Admission to NICU 6) Interval from induction to delivery 7) C-sections 8) Tachysystole 9) Hyperstimulation 10) Delivery within 24 hours	1) Apgar score < 7 at 1 minute: 25-µg dose: 33/192 (17.2%) 50-µg dose: 39/207 (18.8%) p = not significant 2) Apgar score < 7 at 5 minutes: 25-µg dose: 1/192 (0.5%) 50-µg dose: 7/207 (3.4%) p = 0.07 3) Cord pH < 7.6: 25-µg dose: 13/192 (6.8%) 50-µg dose: 27/207 (13.0%) p = 0.04 4) Mean cord pH (± SD): 25-µg dose: 7.26 ± 0.07 50-µg dose: 7.25 ± 0.09 p = not significant 5) Admission to NICU: 25-µg dose: 11/192 (5.7%) 50-µg dose: 23/207 (11.1%) p = 0.07 6) Interval from induction to delivery (mean ± SD): 25-µg dose: 970 ± 684 minutes 50-µg dose: 826 ± 554 minutes p = 0.02 7) C-sections: 25-µg dose: 23/192 (12%) 50-µg dose: 33/207 (15.9%) p = not significant 8) Tachysystole: 25-µg dose: 30/192 (15.6%) 50-µg dose: 68/207 (32.8%) p = 0.0001 9) Hyperstimulation: 25-µg dose: 10/192 (5.2%) 50-µg dose: 12/207 (5.8%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (14% of total study population). Sample size estimates based on incidence of tachysystole. Differences in indications for C-sections (e.g., fetal distress 30% 25 µg vs. 48.5% 50 µg; difference not significant, but study underpowered). *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Abnormal FHR: 5% Diabetes mellitus: 3% Other: 21%		p = not significant 10) Delivery within 24 hours: 25-µg dose: 79/192 (41.1%) 50-µg dose: 101/207 (48.8%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Fletcher, Mitchell, Frederick, et al., 1994	Design: RCT, randomization by drawing odd/even numbers in sealed envelopes Interventions: 1) Misoprostol (n = 32) Protocol: 100-µg tablet placed in posterior vaginal fornix. 2) PGE_2 (n = 31) Protocol: 3-mg tablet placed in posterior vaginal fornix. In both groups, patients not in labor at 12 hours were sent to the labor ward for oxytocin infusion. Dates: Sep-Oct 1992 Location: Kingston, Jamaica Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 64 Dropouts: 1 (excluded from analysis due to protocol violation) Loss to follow-up: NA No. of subjects at end: 63 Inclusion criteria: Scheduled for induction Exclusion criteria: Known contraindications to vaginal prostaglandins, including a previous scar on the uterus; antepartum hemorrhage; fetal distress; PROM; abnormal lie; cephalopelvic disproportion; any maternal illness for which induction contraindicated Age (mean ± SD): Misoprostol, 27.1 ± 6.0; PGE_2, 28.0 ± 5.1 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 38.8 ± 2.8 weeks; PGE_2, 39.7 ± 1.5 weeks Dating criteria: NR Parity (mean ± SD): Misoprostol, 0.6 ± 0.8; PGE_2, 1.1 ± 1.2 Bishop score (mean ± SD): Misoprostol, 4.1 ± 2.3; 4.4 ± 2.5 Other: Indications for induction: Hypertension: 38% Postterm: 33% Diabetes: 11%	1) Apgar scores at 1 and 5 minutes 2) Perinatal deaths 3) Time from induction to delivery 4) Forceps deliveries 5) Vacuum deliveries 6) C-sections	1) Apgar scores at 1 and 5 minutes (mean): *At 1 minute:* Misoprostol: 7.6 PGE_2: 8.3 p = 0.12 *At 5 minutes:* Misoprostol: 8.8 PGE_2: 9.1 p = 0.45 2) Perinatal deaths: None in either group 3) Time from induction to delivery (mean ± SD): Misoprostol: 21.8 ± 29.3 hours PGE_2: 32.3 ± 36.6 hours p = 0.21 4) Forceps deliveries: Misoprostol: 1/32 (3%) PGE_2: 0/31 (no p-value reported) 5) Vacuum deliveries: Misoprostol: 3/32 (9%) PGE_2: 0/32 (no p-value reported) 6) C-sections: Misoprostol: 1/32 (3%) PGE_2: 3/31 (10%) p = 0.17	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: ?? Mode of delivery: ?? Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (33% of total study population). Results not stratified by parity. Study underpowered to detect differences. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Excess weight gain: 3% Cardiac: 3% IUGR, previous stillbirth, poor weight gain, eclampsia, low biological profile score, weight loss at term, and unstable lie: 1.6% each			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Fletcher, Mitchell, Simeon, et al., 1993	Design: RCT, method of randomization not described Interventions: 1) Misoprostol (n = 24) Protocol: Misoprostol 100 µg powder mixed with sterile gel and placed in posterior vaginal fornix using a syringe. At 12 hours, patients not in labor were sent to the labor ward for oxytocin infusion. 2) Placebo (n = 21) Protocol: Same as above, except placebo powder (0.05 mg ethinyl oestradiol) used instead of misoprostol. Dates: NR Location: Kingston, Jamaica Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 48 Dropouts: 3 Loss to follow-up: NA No. of subjects at end: 45 Inclusion criteria: Indication for induction; 3rd trimester pregnancy; unripe cervix; no contraindication to prostaglandins Exclusion criteria: None specified Age (mean ± SD): Misoprostol, 25.8 ± 6.3; placebo, 26.0 ± 4.9 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.5 ± 2.2 weeks; placebo, 39.8 ± 1.7 weeks Dating criteria: NR Parity: Misoprostol, 54% nulliparous; placebo, 43% nulliparous Bishop score (mean ± SD): Misoprostol, 3.1 ± 1.5; placebo, 3.1 ± 2.0 Other: Indications for induction: Postterm: 51% Preeclampsia: 27% Preeclampsia with IUD: 4% Diabetes mellitus: 7% IUGR: 2% UTI: 2% Rheumatic heart: 2% Previous stillbirth: 2% Oligohydramnios: 2%	1) Apgar scores at 1 and 5 minutes (for women receiving oxytocin augmentation) 2) Meconium staining 3) Fetal tachycardia 4) Time from induction to delivery 5) Forceps deliveries 6) C-sections	1) Apgar scores at 1 and 5 minutes (mean ± SD) (for women receiving oxytocin augmentation): *At 1 minute:* Misoprostol (n = 7): 8.1 ± 2.3 Placebo (n = 13): 7.7 ± 2.2 p = 0.34 *At 5 minutes:* Misoprostol (n = 7): 8.9 ± 2.2 Placebo (n = 13): 8.9 ± 2.2 p = 0.73 2) Meconium staining: Misoprostol: 2/24 (8%) Placebo: 0/21 p = not significant 3) Fetal tachycardia: Misoprostol: 0/24 Placebo: 2/21 (9.5%) p = not significant 4) Time from induction to delivery (mean ± SD): Misoprostol: 15.6 ± 12.5 hours Placebo: 43.2 ± 20.5 hours p < 0.001 5) Forceps deliveries: Misoprostol: 1/24 (4%) Placebo: 1/21 (5%) p = not significant 6) C-sections: Misoprostol: 2/24 (8%) Placebo: 3/21 (14%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (51% of total study population). Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Frydman, Lelaidier, Baton-Saint-Mleux, et al., 1992	Design: RCT, randomized by computer-generated tables Interventions: 1) Mifepristone (n = 60), in women from 37.5-41.4 weeks, given as two 200-mg oral doses 24 hours apart 2) Placebo (n = 60) In both groups, NST performed each day until day 4, when induction done with vaginal PGE_2 if no labor. Dates: Apr 1990 - Jan 1991 Location: Clamart, France Setting: Unspecified hospital Type(s) of providers: NR Length of follow-up: None	No of subjects at start: 120 Drop-outs: 8 Loss to follow-up: NA No of subjects at end: 112 Inclusion criteria: Indication for induction (48% "postdates"); Bishop score < 4 Exclusion criteria: Medical condition; nonvertex presentation; more than one prior cesarean; multiple gestation; premature rupture of membranes Age (mean ± SD) Mifepristone: 31 ± 4.1 Placebo: 29 ± 3.6 Gestational age at entry (mean ± SD): Mifepristone: 39.9 ± 1.2 Placebo: 39.7 ± 1.2 Parity (% nulliparous) Mifepristone: 65% Placebo: 60% Bishop score: NR (100% < 4)	1) Proportion in spontaneous labor 2) Bishop score < 4 on day 4 3) Interval from randomization to start of labor 4) Cesarean delivery 5) Epidural anesthesia 6) Apgar < 7 at 1 minute 7) Apgar <7 at 5 minutes	1) Proportion in spontaneous labor: Mifepristone: 54% Placebo: 18% p < 0.001 2) Bishop score < 4 on day 4: Mifepristone: 23% Placebo: 58% p < 0.001 3) Interval from randomization to start of labor: Mifepristone: mean 51 h 45 min Placebo: mean 74 h 30 min P < 0.001 4) Cesarean delivery: Mifepristone: 30% Placebo: 30% No detectable differences by indication 5) Epidural anesthesia: Mifepristone: 73% Placebo: 82% p = not significant 6) Apgar < 7 at 1 minute: Mifepristone: 5/57 Placebo: 4/55 p = not significant NS 7) Apgar <7 at 5 minutes: Mifepristone: 0/57 Placebo: 0/55	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: - Mode of delivery: + Sample size: - Statistical tests: - Gestational age: + Dating criteria: + Bishop score: - Study underpowered to detect differences in categorical outcomes.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Garry, Figueroa, Guillaume, et al., 2000	Design: RCT, patients alternately assigned to one of two study groups Interventions: 1) Castor oil (n = 52) Protocol: Single 60-ml oral dose given, diluted in apple or orange juice. 2) No treatment (n = 48) Dates: July 1992 - Feb 1993 Location: Brooklyn, NY Setting: Community hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 103 Dropouts: 3 Loss to follow-up: NA No. of subjects at end: 100 Inclusion criteria: Gestational age 40-42 weeks; Bishop score ≤ 4; no regular uterine contractions Exclusion criteria: Ruptured membranes; multiple gestations; oligohydramnios; IUGR; abnormal FHR tracings; biophysical profile score ≤ 8; noncephalic presentation; maternal medical complications Age (mean ± SD): Castor oil, 24.8 ± 6.7; no treatment, 24.4 ± 4.9 Race: NR Gestational age at entry (mean ± SD): Castor oil, 284.4 ± 4.2 days; no treatment, 284.7 ± 3.6 days Dating criteria: LMP or early U/S (obtained in 1st or 2nd trimester) Parity: Castor oil, 42.3% nulliparous; no treatment, 43.8% nulliparous Bishop score: NR; score ≤ 4 required for entry into study Other: Indications for induction not reported	1) Birthweight 2) Meconium staining 3) Labor within 24 hours 4) C-sections	1) Birthweight (mean ± SD): Castor oil: 3486 ± 434 g No treatment: 3437 ± 420 g p = 0.56 2) Meconium staining: Castor oil: 10.4% No treatment: 11.5% p = 3) Labor within 24 hours: Castor oil: 30/52 (57.7%) No treatment: 2/48 (4.2%) p < 0.001 4) C-sections: Castor oil: 10/52 (19.2%) No treatment: 4/48 (8.3%) p = 0.20	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not stratified by parity or by indication for induction. Study underpowered to detect differences in C-section rate.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Giacalone, Targosz, Laffargue, et al., 1998	Design: RCT, randomization by permutation blocks and sealed envelope Interventions: 1) Mifepristone for cervical ripening (n = 41) Protocol: Mifepristone 400 mg given as a single oral dose. Patients re-examined 24 and 48 hours later. If Bishop score ≥ 6, then patient induced with oxytocin and amniotomy. If Bishop score < 6, then cervical ripening/induction considered to have failed, and patient managed in accordance with physician's "usual induction techniques." FHR tracing done at each exam visit and during labor. 2) Placebo (n = 42) Protocol: Same as above, but with identical placebo used in place of mifepristone. Dates: Jan 1991 - Feb 1992 Location: Montpellier and Nantes, Frances Setting: 2 university hospitals Type(s) of providers: Unspecified OB/GYN Length of follow-up: Follow-up visit scheduled for neonates 1-2 months after birth	No. of subjects at start: 84 Dropouts: 1 Loss to follow-up: 7 (not available for 1-2 month follow-up) No. of subjects at end: 76 Inclusion criteria: Gestational age ≥ 41 weeks and 3 days; Bishop score < 6; labor induction post-ponable for 48 hours Exclusion criteria: Contra-indication to vaginal delivery; multiple gestation; > 4 previous deliveries; uterine scar; premature rupture of the membranes; FHR abnormality; impaired renal, adrenal, or hepatic function; corticosteroid therapy during pregnancy; abnormal hemostasis; anticoagulant therapy Age (mean ± SD): Mifepristone, 28.5 ± 4.3; placebo, 28.3 ± 5.0 Race: NR Gestational age at entry: NR*; at delivery* mifepristone, 41.5 ± 0.2 weeks; placebo, 41.6 ± 0.2 weeks Dating criteria: NR Parity: Mifepristone, 20/41 (49%) nulliparous; placebo, 20/42 (48%) nulliparous Bishop score (median, with range): Mifepristone, 3 (1 to 5); placebo, 3 (1 to 5)	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Umbilical artery pH < 7.2 5) Glycemia ≤ 40 mg/dL 6) Cortisol levels 7) Post-natal abnormalities 8) C-sections 9) Cervical ripening in patients with Bishop score < 6 10) Instrumental delivery 11) Time to onset of labor 12) Time to delivery (excluding C-sections)	1) Apgar score < 7 at 1 minute: Mifepristone: 3/41 (7.3%) Placebo: 2/42 (4.8%) p = not significant 2) Apgar score < 7 at 5 minutes: Mifepristone: 0 Placebo: 0 3) Birthweight (mean ± SD): Mifepristone: 3418 ± 380 g Placebo: 3502 ± 364 g p = not significant 4) Umbilical artery pH < 7.2: Mifepristone: 3/41 (7.3%) Placebo: 2/42 (4.8%) p = not significant 5) Glycemia ≤ 40 mg/dL: *Day 1:* Mifepristone: 1/41 (2.4%) Placebo: 6/42 (14.3%) p = not significant *Day 2:* Mifepristone: 1/41 (2.4%) Placebo: 1/42 (2.4%) p = not significant 6) Cortisol levels (median, with range): Mifepristone: 153.5 nmol/L (42 to 537) Placebo: 94.5 nmol/L (28 to 223) (no p-value reported) 7) Post-natal abnormalities (at 1-2 month follow-up): Mifepristone: 5/38 (13%) Placebo: 2/38 (5.3%) p = 0.42	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not stratified by parity.

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				8) C-sections: Mifepristone: 7/41 (17%) Placebo: 6/42 (14.3%) p = not significant 9) Cervical ripening in patients with Bishop score < 6: Mifepristone: 7/41 (17.1%) Placebo: 17/42 (40.4%) p = not significant 10) Instrumental delivery: Mifepristone: 9/41 (22%) Placebo: 6/42 (14.3%) (no p-value reported) 11) Time to onset of labor (median, with range): Mifepristone: 31.7 hours (9.5 to 117.8) Placebo: 53.9 hours (2.5 to 192.0) p = 0.02 12) Time to delivery (excluding C-sections) (median, with range): Mifepristone: 31.3 hours (13.2 to 123.3) Placebo: 58.5 hours (5.8 to 193.7) p = 0.02	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Gottschall, Borgida, Mihalek, et al., 1997	Design: RCT, randomization by random-numbers table and sealed envelopes Interventions: 1) Misoprostol (n = 38) Protocol: 100 µg placed in posterior vaginal fornix. 2) PGE_2 gel (n = 37) Protocol: PGE_2 gel (5 mg) placed in posterior vaginal fornix by syringe. In both groups, patients were re-examined at 6 hours after placement of study medication. If patient in labor (≥ 3 contractions/10 minutes, with changes in cervical dilatation), then amniotomy performed. If patient not in labor, then oxytocin augmentation initiated. Dates: Nov 1995- Aug 1996 Location: New Britain, CT Setting: Community hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 75 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 75 Inclusion criteria: Indication for cervical ripening and induction; live, singleton fetus; cephalic presentation; intact membranes; reactive FHR tracing; no contra-indications to a vaginal delivery Exclusion criteria: Previous uterine scar; allergy to prostaglandin agents Age (mean ± SD): Misoprostol, 28.4 ± 5.7; PGE_2, 26.9 ± 6.4 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.8 ± 1.7 weeks; PGE_2, 39.8 ± 2.2 weeks Dating criteria: NR Parity: Misoprostol, 61% nulli-parous; PGE_2, 68% nulliparous Bishop score (median): 4, both groups Other: Indications for induction: Postterm: 40% Preeclampsia: 27% Oligohydramnios: 16% IUGR: 7% Chronic hypertension: 3% Diabetes: 1% Other: 7%	1) Apgar scores at 1 and 5 minutes 2) Birthweight 3) Time from induction to delivery 4) Delivery by 24 hours 5) Hyperstimulation 6) Tachysystole 7) C-sections	1) Apgar scores at 1 and 5 minutes (median): *At 1 minute:* Misoprostol: 8 PGE_2: 8 p = not significant *At 5 minutes:* Misoprostol: 9 PGE_2: 9 p = not significant 2) Birthweight (mean ± SD): Misoprostol: 3438 ± 536 g PGE_2: 3435 ± 591 g p = not significant 3) Time from induction to delivery (mean ± SD): Misoprostol: 14.7 ± 6.4 hours PGE_2: 20.4 ± 10.2 hours p = 0.005 4) Delivery by 24 hours: Misoprostol: 95% PGE_2: 70% p = 0.005 5) Hyperstimulation: Misoprostol: 2.8% PGE_2: 0 p = not significant 6) Tachysystole: Misoprostol: 15.8% PGE_2: 2.7% p = not significant 7) C-sections: Misoprostol: 18% PGE_2: 27% p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (40% of total study population). Sample size estimates based on time to delivery. Underpowered to detect differences in some outcomes. Findings similar when nulliparas analyzed separately.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Graves, Baskett, Gray, et al., 1985	Design: RCT, randomization method not specified Interventions: 1) 3 mg PGE_2 gel (n = 20) inserted into posterior vaginal fornix via catheter 2) 2 mg PGE_2 gel (n = 20) 3) 1 mg PGE_2 gel (n = 20) 4) Placebo gel (n = 20) In all groups, patients monitored for 1 hour after insertion. If no labor after 12-16 hours, induction with oxytocin ± amniotomy. Dates: NR Location: Halifax, Canada Setting: University hospital Providers: Unspecified OB/GYN Length of follow-up: None	No of subjects at start: 80 Drop-outs: 0 Loss to follow-up: NA No of subjects at end: 80 Inclusion criteria: Gestational age ≥ 36 weeks; Bishop score ≤ 4 Exclusion criteria: regular uterine contractions; contraindication to vaginal delivery; asthma or hypersensitivity to prostaglandins; prior attempts at ripening or induction in this pregnancy; malpresentation; multiple gestation; intrauterine death; polyhydramnios; antepartum hemorrhage; premature rupture of membranes; uterine scar Age (mean): 3 mg: 27.3 2 mg: 24.7 1 mg: 27.2 Placebo: 26.8 Race: NR Gestational age at entry (mean): 3 mg: 38.9 2 mg: 39.0 1 mg: 39.0 Placebo: 40.0 Dating criteria: NR Parity (% nulliparous): 3 mg: 40% 2 mg: 65% 1 mg: 65% Placebo: 55%	1) Change in Bishop score 2) Labor after gel alone 3) Cesarean section 4) Uterine hyper-contractility	1) Change in Bishop score: 3 mg: 3.8 2 mg: 2.6 1 mg: 2.7 Placebo: 1.4 $p < 0.01$ 2) Labor after gel alone: 3 mg: 50% 2 mg: 25% 1 mg: 5 % Placebo: 0% 3) Cesarean section: 3 mg: 20% 2 mg: 25% 1 mg: 35% Placebo: 15% 4) Uterine hypercontractility: 3 mg: 20% 2 mg: 10% 1 mg: 5% Placebo: 0%	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: - Gestational age: + Dating criteria: - Bishop score: + Underpowered to detect many important differences or trends. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Bishop score (mean): 3 mg: 2.6 2 mg: 3.0 1 mg: 2.7 Placebo: 2.4 Other: 18% of subjects induced for prolonged pregnancy			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Grünberger and Spona, 1986	Design: RCT, method of randomization not described Interventions: 1) PGE_2 (1.5 mg) in saline (n = 15) Protocol: PGE_2 injected through syringe, using cervical cap. If labor within 6 hours, then cap removed; if no labor, then administration repeated. If no labor by 24 hours, then patient crossed over to other treatment group. Amniotomy performed when labor established and cervix sufficiently dilated (≥ 4 cm). 2) Placebo (n = 15) Protocol: Same as above, but with saline alone Dates: NR Location: Vienna, Austria Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 30 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 30 Inclusion criteria: 41-42 weeks gestation; unfavorable cervix Exclusion criteria: Maternal or fetal risk factors; twin pregnancy; breech presentation; previous C-section; previous surgery on cervix Age: NR Race: NR Gestational age at entry: NR (gestational age of 41-42 weeks required for entry into study) Dating criteria: NR Parity: Two groups "equal" (no further information provided) Bishop score (mean): PGE_2, 4.7; placebo, 4.6	1) Treatment failure (neither cervical ripening nor delivery)	1) Treatment failure: PGE_2: 1/15 (6.6%) Placebo: 10/15 (66.6%) $p < 0.001$	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: - Dating criteria: - Bishop score: + Results summarized for period before crossover.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Gupta, Vasishta, Sawhney, et al., 1998	Design: RCT, randomization by computer-generated list and sealed envelope Interventions: 1) Stripping of membranes (n = 50) Protocol: Stripping of membranes performed at 38 weeks by digital separation of 2-3 cm of chorionic membranes from lower uterine segment using two circumferential passes of the examining fingers. Performed "under aseptic precautions." Patients then followed weekly (no details provided) until delivery or scheduled induction. 2) Gentle cervical exam (control) (n = 50) Protocol: Exam not described. Performed at 38 weeks "under aseptic precautions." Patients then followed weekly (no details provided) until delivery or scheduled induction. Dates: NR Location: Chandigarh, India Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 100 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 100 Inclusion criteria: Confirmed gestational age; early confirmation of pregnancy, cephalic presentation; no contraindication to vaginal delivery Exclusion criteria: Closed cervix at 38 weeks gestation; known medical disease or medical complications of pregnancy; multiple pregnancy; hydramnios; premature rupture of membranes; vaginal or cervical infection; low-lying placenta; intrauterine fetal death; malpresentation; labor; cephalopelvic disproportion Age (mean ± SD): Stripping, 24.46 ± 3.07; control, 23.52 ± 2.55 Race: NR Gestational age at entry (mean ± SD): Stripping, 38.00 ± 0.44 weeks; control, 38.02 ± 0.10 Dating criteria: NR Parity: 100% primigravidae Bishop score: Stripping, 86% < 6; control, 82% < 6	1) Apgar scores at 1 minute 2) Apgar scores at 5 minutes 3) Birthweight 4) Admission to NICU 5) Stillbirths 6) Gestational age at onset of labor 7) Days from intervention to delivery 8) Pregnancy continuing beyond 40 weeks 9) Induction of labor 10) C-sections 11) Assisted vaginal delivery 12) Microbiological flora	1) Apgar scores at 1 minute (mean ± SD): Stripping: 7.80 ± 0.17 Control: 7.74 ± 0.16 p > 0.05 2) Apgar scores at 5 minutes (mean ± SD): Stripping: 8.96 ± 0.19 Control: 9.12 ± 0.12 p > 0.05 3) Birthweight (mean ± SD): Stripping: 2882 ± 340 g Control: 2894 ± 420 g (no p-value reported) 4) Admission to NICU: Stripping: 0 Control: 2/50 (4%) (no p-value reported) 5) Stillbirths: Stripping: 1/50 (2%) Control: 0 p > 0.05 6) Gestational age at onset of labor (mean ± SD): Stripping: 38.70 ± 0.63 weeks Control: 39.83 ± 0.56 weeks p < 0.001 7) Days from intervention to delivery (mean ± SD): Stripping: 4.62 ± 4.15 Control: 11.95 ± 8.27 p < 0.005 8) Pregnancy continuing beyond 40 weeks: Stripping: 2/50 (4%) Control: 17/50 (34%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				(no p-value reported) 9) Induction of labor: Stripping: 1/50 (2%) Control: 16/50 (32%) $p < 0.05$ 10) C-sections: *Overall:* Stripping: 6/50 (12%) Control: 8/50 (16%) $p > 0.05$ *For fetal distress:* Stripping: 3/50 (6%) Control: 5/50 (10%) (no p-value reported) *For nonprogress of labor:* Stripping: 3/50 (6%) Control: 3/50 (6%) (no p-value reported) 11) Assisted vaginal delivery: Stripping: 13/50 (26%) Control: 9/50 (18%) (no p-value reported) 12) Microbiological flora: No significant difference in the microbiological flora of cervical swabs (taken at time of intervention and at onset of labor) or the placental membrane in the two groups.	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Herabutya, Prasert-sawat, and Pokpirom, 1997	Design: RCT, blocked randomization scheme Interventions: 1) Misoprostol (n = 60) Protocol: 100-µg tablet placed in posterior vaginal fornix. 2) PGE_2 gel (n = 50) Protocol: PGE_2 gel (1.5 mg) placed via catheter into the endocervix In both groups, patients re-examined at 12 hours. Amniotomy carried out if cervix 80% effaced and 3 cm dilated. Patients who did not enter active labor or who had SROM without adequate uterine contractions were given oxytocin augmentation. At 24 hours, those still not in labor were sent to the labor ward for induction by amniotomy and oxytocin. Dates: May 1995 - Apr 1996 Location: Bangkok, Thailand Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 110 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 110 Inclusion criteria: Medical or obstetric indication for induction; singleton pregnancy; cephalic presentation; intact membranes; Bishop score ≤ 4 Exclusion criteria: None specified Age (mean ± SD): Misoprostol, 29.12 ± 4.69; PGE_2, 28.18 ± 4.72 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.33 ± 1.41 weeks; PGE_2, 39.74 ± 1.43 weeks Dating criteria: NR Parity: Misoprostol, 73% nulliparous; PGE_2, 82% nulliparous Bishop score (mean ± SD): Misoprostol, 2.22 ± 1.06; PGE_2, 2.50 ± 1.15 Other: Indications for induction: Preeclampsia: 44% Postterm: 34% Decreased fetal movement: 9% Diabetes mellitus: 4% IUGR: 3% Previous dead fetus: 4% Nonreactive NST: 4%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Time from induction to delivery 4) Hyperstimulation 5) C-sections	1) Apgar score < 7 at 1 minute: Misoprostol: 4/60 (6%) PGE_2: 4/50 (8%) p = 1.00 2) Apgar score < 7 at 5 minutes: Misoprostol: 0/60 PGE_2: 1/50 (2%) p = 0.45 3) Time from induction to delivery (mean ± SD): Misoprostol: 19.14 ± 10.64 hours PGE_2: 21.37 ± 13.09 hours p = 0.33 4) Hyperstimulation: Misoprostol: 1/60 PGE_2: 0/50 (no p-value reported) 5) C-sections: Misoprostol: 19/60 (31.7%) PGE_2: 16/50 (32.0%) p = 0.87	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (34% of total study population). Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Howarth, Funk, Steytler, et al., 1996	Design: RCT, randomization by computer-generated list of random numbers and sealed envelopes Interventions: 1) Misoprostol (n = 36) Protocol: 100 µg misoprostol placed in posterior vaginal fornix. 2) PGE_2 gel (n = 36) Protocol: 1 mg PGE_2 gel placed in posterior vaginal fornix. In both groups, second dose administered after 6 hours if cervix remained unfavorable. Patients not in labor by 12 hours were managed according to their physician's preference. C-section was performed for suspected fetal distress. Dates: Apr - June 1995 Location: Pretoria, South Africa Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None	No. of subjects at start: 72 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 72 Inclusion criteria: Singleton pregnancy, longitudinal lie; cephalic presentation; fetal well-being; anticipated fetal mass > 2000 g; intact membranes; unfavorable cervix Exclusion criteria: Contra-indication to vaginal delivery; previous C-section; parity > 4; contraindication to prostaglandins Age (median, with range): Misoprostol, 27 (18-41); PGE_2, 27 (18-24) Race: NR Gestational age at entry (median, with range): Misoprostol, 40 weeks (35-43); PGE_2, 40 weeks (34-42) Dating criteria: NR Parity (median, with range): Misoprostol, 1 (0-4); PGE_2, 1 (0-4) Bishop score (median, with range): Misoprostol, 4 (2-7); PGE_2, 5 (2-7) Other: Indications for induction: Hypertension: 47% Postterm: 33% Other: 19%	1) Apgar score at 5 minutes 2) Birthweight 3) C-sections 4) Delivery within 12 hours 5) Tachysystole	1) Apgar score at 5 minutes (median, with range): Misoprostol: 10 (7-10) PGE_2: 10 (8-10) p = not significant 2) Birthweight (median, with range): Misoprostol: 3220 g (2260-4200) PGE_2: 2880 g (2100-4020) p = not significant 3) C-sections *Overall:* Misoprostol: 6/36 (17%) PGE_2: 15/36 (42%) $p < 0.05$ *For failed induction:* Misoprostol: 1/36 (3%) PGE_2: 6/36 (17%) p = not significant *For prolonged 1st stage of labor:* Misoprostol: 0/36 PGE_2: 7/36 (19%) $p < 0.01$ *For suspected fetal distress:* Misoprostol: 5/36 (14%) PGE_2: 2/36 (5.5%) p = not significant 4) Delivery within 12 hours: Misoprostol: 30/36 (83%) PGE_2: 13/36 (36%) $p < 0.05$ 5) Tachysystole: Misoprostol: 14/36 (39%) PGE_2: 3/36 (8%) $p < 0.01$	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (33% of total study population). Results not stratified by parity. 42% of patients in the misoprostol group were postdates vs. 25% in the PGE_2 group. Difference not significant, but study underpowered to detect differences at baseline or for outcomes.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Idrisa, Obisesan, and Adeleye, 1993	Design: RCT, patients assigned alternately to one of two treatment groups Interventions: 1) Membrane sweeping (n = 100) Protocol: Membrane sweeping performed at 41 weeks using the examiner's index finger. If no labor within 6 days, then patient induced with oxytocin. 2) Control (n = 100) Management of control group not specified Dates: Jan 1988 - Dec 1990 Location: Ibadan, Nigeria Setting: University hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 200 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 200 Inclusion criteria: Gestational age 41 weeks; no spontaneous labor Exclusion criteria: Contra-indications to vaginal delivery Age (mean ± SD): Membrane sweeping, 26 ± 3.1; control, 26 ± 3.3 Race: NR Gestational age at time of induction (mean ± SD): Both groups, 292 ± 2 days Dating criteria: 2nd trimester U/S Parity: NR Bishop score: NR	1) Birthweight 2) Perinatal death 3) Complications 4) Vacuum extraction/ forceps-assisted delivery 5) C-sections 6) Spontaneous labor	1) Birthweight (mean ± SD): Sweeping: 3.05 ± 0.25 kg Control: 3.05 ± 0.25 kg p = not significant 2) Perinatal death: Sweeping: 0/100 Control: 0/100 p = not significant 3) Complications: "No severe maternal or neonatal complication attributable to membrane sweeping was observed." 4) Vacuum extraction/ forceps-assisted delivery: Sweeping: 3/100 Control: 6/100 p = not significant 5) C-sections: Sweeping: 2/100 Control: 3/100 p = not significant 6) Spontaneous labor: Sweeping: 92/100 Control: 33/100 p < 0.001	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Kadanali, Küçüköz-kan, Zor, et al., 1996	Design: RCT, randomization by sealed envelope Interventions: 1) Misoprostol (n = 112) Protocol: Misoprostol 100 µg tablet inserted intravaginally in the posterior fornix. Same dose repeated orally every 2 hours until adequate labor established (at least 3 contractions in 10 minutes). If labor not achieved by 24 hours, then patient infused with 10 IU oxytocin in 1000 ml 5% glucose solution. Infusion started at rate of 4 mIU/min and doubled every 30 minutes (to maximum of 32 mIU/min) until contractions began. If no active labor after 12 hours of oxytocin administration, then C-section performed. 2) PGE_2 gel + oxytocin (n = 112) Protocol: PGE_2 gel instilled into cervix. If no labor after 6 hours, then oxytocin infusion initiated "according to a uniform protocol." Dates: Mar-Aug 1995 Location: Erzurum, Turkey Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 224 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 224 Inclusion criteria: Medical or obstetrical indication for induction; no labor or fetal distress; gestational age 37-42 weeks; singleton vertex presentation Exclusion criteria: Previous uterine surgery, including C-section; Bishop score ≥ 6 Age (mean ± SD): Misoprostol, 22.3 ± 5.7; PGE_2/oxytocin, 22.5± 5.3 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 38.2 ± 3.4 weeks; PGE_2/oxytocin, 38.8 ± 2.8 weeks Dating criteria: NR Parity: Misoprostol, 70% nulliparous; PGE_2/oxytocin, 73% nulliparous Bishop score (mean ± SD): Misoprostol, 4.0 ± 1.4; PGE_2/oxytocin, 3.8 ± 1.4 Other: Indications for induction were as follows: Postdates: 41% Preeclampsia: 22% PROM: 11%	1) Apgar score < 5 at 5 minutes 2) Birthweight 3) Cord pH < 7.16 4) Vacuum extraction 5) C-sections for obstetric indication 6) C-sections for failed induction 7) Cost per patient 8) Time to delivery	1) Apgar score < 5 at 5 minutes: Misoprostol: 2/112 (1.8%) PGE_2/oxytocin: 2/112 (1.8%) p = not significant 2) Birthweight (mean ± SD): Misoprostol: 3382 ± 702.3 g PGE_2/oxytocin: 3302 ± 771.9 g p = not significant 3) Cord pH < 7.16: Misoprostol: 8/112 (7.1%) PGE_2/oxytocin: 10/112 (8.9%) p = not significant 4) Vacuum extraction: Misoprostol: 4/112 (3.6%) PGE_2/oxytocin: 5/112 (4.5%) p = not significant 5) C-sections for obstetric indication: Misoprostol: 5/112 (4.5%) PGE_2/oxytocin: 6/112 (5.4%) p = not significant 6) C-sections for failed induction: Misoprostol: 7/112 (6.3%) PGE_2/oxytocin: 15/112 (13.4%) p = 0.001 7) Cost per patient: Misoprostol: $1.50 PGE_2/oxytocin: $28.00 (no p-value reported) 8) Time to delivery (mean ± SD): Misoprostol: 9.2 ± 2.4 hours PGE_2/oxytocin: 15.2 ± 3.2 hours p = 0.001	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Mean gestational age 38 weeks, but 41% induced for "postdates." Results not reported separately for subgroup of patients induced for postterm pregnancy (41% of total study population). Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Diabetes: 5% IUGR: 6% Other: 15%			
Kadar, Tapp, and Wong, 1990	Design: RCT, randomization by hospital number Interventions: 1) Nipple stimulation (NS) (n =62) Protocol: Women given written instructions for NS and instructed to perform unilateral NS manually each day "for as long as was practically feasible." Told to stop NS if contractions occurred more frequently than 5 in 10 minutes if a contractions lasted more than 90 seconds; NS could be resumed once the contractions had abated. 2) Control (no nipple stimulation) (n = 76) Dates: NR Location: London, England Setting: Outpatient clinic/physician office; university hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 155 Dropouts: 17 (11%) Loss to follow-up: NA No. of subjects at end: 138 Inclusion criteria: Low-risk pregnancy; ≥ 39 weeks gestation Exclusion criteria: None specified; patients withdrawn if pregnancy complications developed during the study Age (median): NS, 26.5; control, 25.0 Race: NS, 81% White; control, 75% White Gestational age at entry: Median, 281 days in both groups Dating criteria: LMP *or* U/S before 20 weeks Parity: NS, 52% nulliparous; control, 50% nulliparous Bishop score (median): Both groups, 5.0	1) Birthweight 2) Spontaneous delivery 3) Spontaneous labor 4) Postterm deliveries (> 294 days) 5) Pregnancy duration	1) Birthweight (median): NS: 3500 g Control: 3500 g (no p-value reported) 2) Spontaneous delivery: NS: 48/62 (77%) Control: 64/76 (84%) (no p-value reported) 3) Spontaneous labor: NS: 60/62 (97%) Control: 70/76 (92%) (no p-value reported) 4) Postterm deliveries: NS: 9/62 (14.5%) Control: 8/76 (10.5%) (no p-value reported) 5) Pregnancy duration (median): NS: 281 days Control: 281 days (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Compliance with nipple stimulation was poor. 70% of the women assigned to the NS group either failed to perform NS altogether or did so for < 2 hours in total. Survival analysis showed that duration of pregnancy was influenced only by the gestational age at enrollment and the Bishop score at enrollment. Nipple stimulation did not significantly affect the duration of pregnancy or the frequency of postterm deliveries. Women assigned to the nipple stimulation group who refused to participate were included with controls in the analysis.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Kemp, Winkler, and Rath, 2000	Design: RCT, randomization by stratified block Interventions: 1) PGE_2 vaginal gel (2 mg) (n = 229) Protocol: Gel administered in the posterior fornix. Repeated every 6-8 hours up to 3 times until Bishop score > 7. When Bishop score > 7, oxytocin administered 8 hours after last PGE_2 administration. If no labor and no improvement in Bishop score after 3 applications of gel, then 24-hour rest, followed by either induction with prostaglandins or C-section, as clinically indicated. FHR monitored for 2 hours following PGE_2 application and intermittently thereafter. 2) PGE_2 intracervical gel (0.5 mg) (n = 241) Protocol: Same as above, except that 0.5-mg gel administered "high into the cervical canal." Dates: Apr 1995 - July 1997 Location: Aachen, Germany Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 470 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 470 Inclusion criteria: Singleton pregnancy; vertex presentation; medical indication for induction (> 10 days postterm, premature rupture of the membranes, IUGR, hypertension; gestational or pre-existing diabetes); Bishop score 3-4 Exclusion criteria: Known contraindications for prostaglandins; previous uterine surgery; previous vertical C-section; uterine abnormality; FHR abnormality Age: NR Race: NR Gestational age at entry: NR; vaginal gel, 32.9% > 10 days postterm; intracervical gel, 29.2% > 10 days postterm Dating criteria: NR Parity: NR Bishop score: NR (required to be 3 or 4 for entry into study)	1) Apgar score ≤ 7 at 5 minutes 2) Umbilical artery pH < 7.20 3) C-sections 4) Change in Bishop score (before/after 1st administration) 5) Vaginal delivery within 24 hours 6) Time from induction to delivery 7) "Maternal side effects" 8) Hyperstimulation	1) Apgar score ≤ 7 at 5 minutes: Vaginal gel: 1.3% Intracervical gel: 2.1% p = not significant 2) Umbilical artery pH < 7.20: Vaginal gel: 12.3% Intracervical gel: 8.7% p = not significant 3) C-sections: Vaginal gel: 22.3% Intracervical gel: 26.7% p = not significant 4) Change in Bishop score (before/after 1st administration) (mean): Vaginal gel: 1.9 Intracervical gel: 1.35 p = 0.001 5) Vaginal delivery within 24 hours: Vaginal gel: 81.6% Intracervical gel: 67.8% p = 0.001 6) Time from induction to delivery (median): Vaginal gel: 15.7 hours Intracervical gel: 19.1 hours p = 0.01 7) "Maternal side effects": Vaginal gel: 5.7% Intracervical gel: 6.7% p = not significant 8) Hyperstimulation: Vaginal gel: 14.5% Intracervical gel: 13.0% p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: - Dating criteria: - Bishop score: +

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Kramer, Gilson, Morrison, et al., 1997	Design: RCT, randomization schedule computer-generated by hospital pharmacy Interventions: 1) Misoprostol (100 µg) (n = 60) Protocol: Misoprostol 100 µg placed in posterior vaginal fornix every 4 hours until adequate uterine contractions achieved (defined as > 200 Montevideo units). No lubricating gel used to place tablets. Repeat dosing (up to max of 5 doses) permitted if uterine activity inadequate and fetus tolerating labor. Oxytocin started if labor had not progressed by 4 hours after last dose of misoprostol. 2) Oxytocin infusion (n = 66) Protocol: Intravenous oxytocin started at an infusion rate of 1 mU/min. Dose increased every 30 min until adequate uterine activity achieved (> 200 Montevideo units). Maximal infusion rate permitted was 36 mU/min. Women in both groups were monitored by external tocodynamometry. Fetal scalp monitoring, cord blood gas sampling, and administration of terbutaline left to discretion of managing physician. Amniotomy generally performed at 3-4 cm dilation. Dates: June 1995 - Apr 1996	No. of subjects at start: 130 Dropouts: 4 women excluded from analysis after randomization Loss to follow-up: NA No. of subjects at end: 126 Inclusion criteria: None stated Exclusion criteria: Multiple gestation; nonvertex presentation; abnormal FHR tracing; previous uterine surgery; allergy to misoprostol; history of asthma; digital exam with lubricant immediately before induction; spontaneous uterine contractions more frequently than every 5 minutes; contraindications to vaginal delivery (e.g., active genital herpes, placenta previa) Age (mean ± SD): Misoprostol, 26.2 ± 5.9; oxytocin, 25.4 ± 5.7 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.6 ± 2.6 weeks; oxytocin, 38.3 ± 3.2 weeks Dating criteria: NR Parity: Misopr / Oxytocin Nulliparous 60% / 49% Primiparous 20% / 29% Multiparous 20% / 22% Bishop score (% with score ≤ 3): Misoprostol, 58%; oxytocin, 38% Other: Indications for induction	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Arterial cord blood pH 4) Birthweight 5) Vacuum delivery 6) Forceps delivery 7) C-section for nonreassuring FHR tracing 8) C-section for dystocia 9) C-section for worsening maternal status 10) Duration of labor 11) Tachystole 12) Estimated hospital charges	1) Apgar score < 7 at 1 minute: Misoprostol: 8/60 (13%) Oxytocin: 12/66 (18%) p = not significant 2) Apgar score < 7 at 5 minutes: Misoprostol: 0/60 Oxytocin: 3/66 (5%) p = not significant 3) Arterial cord blood pH (mean ± SD): Misoprostol (n = 16): 7.21 ± 0.08 Oxytocin (n = 9): 7.19 ± 0.16 p = not significant 4) Birthweight (mean ± SD): Misoprostol: 3262 ± 679 g Oxytocin: 3092 ± 786 p = not significant 5) Vacuum delivery: Misoprostol: 2/60 (3%) Oxytocin: 3/66 (5%) p = not significant 6) Forceps delivery: Misoprostol: 6/60 (10%) Oxytocin: 6/66 (9%) p = not significant 7) C-section for nonreassuring FHR tracing: Misoprostol: 7/60 (12%) Oxytocin: 4/66 (6%) p = not significant 8) C-section for dystocia: Misoprostol: 6/60 (10%) Oxytocin: 14/66 (21%) p < 0.05 9) C-section for worsening maternal status: Misoprostol: 0/60	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (29% of total study population). Results not stratified by parity. 4/60 women in the misoprostol group received oxytocin, but were analyzed in intention-to-treat fashion as part of the misoprostol group. Difference in baseline characteristics suggests problem with randomization. Underpowered to detect some differences in baseline and other variables. Sample size estimates based on time to delivery. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Location: Albuquerque, New Mexico Setting: University hospital Type(s) of providers: General OB/GYN resident physicians under direct supervision of faculty member Length of follow-up: None	were as follows: Preeclampsia: 41% Postterm: 29% Oligohydramnios: 11% Diabetes mellitus: 2% Fetal growth restriction: 1% Other: 16%		Oxytocin: 1/66 (2%) p = not significant 10) Duration of labor (median, with range): Misoprostol: 585 minutes (120-1890) Oxytocin: 885 minutes (120-1890) p < 0.001 11) Tachystole: Misoprostol: 42/60 (70%) Oxytocin: 7/66 (11%) p < 0.001 12) Estimated hospital charges (total charges per patient [mean ± SD]): Misoprostol: $2081 ± $984 Oxytocin: $2616 ± $1035 p < 0.005	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Lee, 1997	Design: RCT, randomization by sealed envelope Interventions: 1) Misoprostol (n = 25) Protocol: 200 µg given intravaginally at 6-hour interval up to a maximum of 2 doses. Patient examined every 6 hours and transferred to labor room when "ready for labor." If no established labor, then oxytocin given. If cervix still unripe after 24 hours, then C-section performed. 2) PGE_2 (n = 25) Protocol: Same as above, except that PGE_2 3 mg used instead of misoprostol. Dates: Beginning Jan 1996 (no end date specified) Location: Pahang, Malaysia Setting: Community hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 50 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 50 Inclusion criteria: At least term + 10 days' gestation; para ≤ 3; singleton pregnancy; cephalic presentation; no prior C-section; no contraindication to prostaglandins; uncomplicated gestation; Bishop score ≤ 6 Exclusion criteria: None specified Age (mean ± SD): Misoprostol, 26.3 ± 4.8; PGE_2, 26.5 ± 4.4 Race: Misoprostol, 84% Malay; PGE_2, 72% Malay Gestational age at entry (mean number of days postdate [± SD]): Misoprostol, 12.5 ± 2.1 days; PGE_2, 12.6 ± 2.6 days Dating criteria: NR Parity (mean ± SD): Misoprostol, 1.3 ± 1.2; PGE_2, 1.1 ± 1.0 Bishop score (mean ± SD): Misoprostol, 4.1 ± 1.1; PGE_2, 4.1 ± 1.2	1) Apgar score at 1 minute 2) Apgar score at 5 minutes 3) Neonatal complication 4) Neonatal hospital stay 5) Moderate meconium aspiration 6) Established labor rate 7) Time to delivery 8) C-sections 9) Polysystole	1) Apgar score at 1 minute (mean ± SD): Misoprostol: 7.7 ± 0.7 PGE_2: 7.6 ± 1.3 p = 0.69 2) Apgar score at 5 minutes (mean ± SD) Misoprostol: 8.9 ± 0.4 PGE_2: 8.7 ± 1.1 p = 0.39 3) Neonatal complication: Misoprostol: 4/25 (16%) PGE_2: 1/25 (4%) p = 0.17 4) Neonatal hospital stay (mean ± SD): Misoprostol: 2.9 ± 2.3 days PGE_2: 2.7 ± 1.0 days p = 0.69 5) Moderate meconium aspiration: Misoprostol: 2/25 (8%) PGE_2: 1/25 (4%) (no p-value reported) 6) Established labor rate: Misoprostol: 23/25 (92%) PGE_2: 16/25 (64%) p = 0.04 7) Time to delivery: *Mean ± SD:* Misoprostol: 676.1 ± 411 minutes PGE_2: 874.9 ± 406 minutes p = 0.09 *Delivered by 6 hours:* Misoprostol: 5/25 (20%) PGE_2: 3/25 (12%) p = 0.35	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Delivered by 12 hours: Misoprostol: 18/25 (72%) PGE_2: 7/25 (28%) p = 0.047 8) C-sections: Misoprostol: 2/25 (8%) PGE_2: 4/25 (16%) p = 0.33 9) Polysystole: Misoprostol: 7/25 (28%) PGE_2: 3/25 (12%) p = 0.28	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Lien, Morgan, Garite, et al., 1998	Design: RCT, randomization by computer-generated table of random numbers Interventions: 1) PGE_2 gel (n = 43) Protocol: PGE_2 gel administered into the endocervical canal. Patient monitored continuously for ≥ 40 minutes. If FHR monitoring "reassuring," then patient instructed to return in 3-4 days for another NST, AFI determination, and gel insertion (up to maximum of 4 doses). Patient induced at 42 weeks, or before then if Bishop score > 9 or "an obstetric factor other than postdate pregnancy developed." Obstetric management during labor determined by patient's obstetrician. 2) Placebo gel (n = 47) Protocol: Same as above, except that identical placebo gel used instead of PGE_2 gel. Dates: NR Location: Anaheim, CA, and Portland, OR Setting: 1 university hospital and 3 community hospitals Type(s) of providers: General OB/GYN; nurse midwives Length of follow-up: None	No. of subjects at start: 92 Dropouts: 2 Loss to follow-up: NA No. of subjects at end: 90 Inclusion criteria: Gestational age ≥ 40 weeks, 3 days; Bishop score ≤ 6; AFI > 5 cm; reactive NST Exclusion criteria: Evidence of hyperstimulation; suspicious FHR patterns; ≥ 5 previous deliveries; nonvertex presentation; multiple gestation; previous C-section; major uterine surgery; placenta previa; other contraindications to vaginal delivery Age (mean ± SD): PGE_2, 25.9 ± 7.0; placebo, 26.4 ± 5.8 Race: PGE_2: 84% White, 12% Hispanic, 5% Asian/Black/other; placebo: 85% White, 6% Hispanic, 9% Asian/Black/other Gestational age at entry (mean ± SD): PGE_2, 40.9 ± 0.3 weeks; placebo, 40.7 ± 0.3 weeks (p = 0.01) Dating criteria: LMP confirmed by either 1st trimester pelvic exam or U/S before 24 weeks Parity: PGE_2, 67% nulliparous; placebo, 55% nulliparous Bishop score (median, with range): PGE_2, 3 (1-6); placebo, 3 (0-5)	1) Apgar score ≤ 7 at 5 minutes 2) Birthweight (mean) 3) Birthweight > 4000 g 4) Shoulder dystocia 5) Gestational age at delivery 6) Time from enrollment to delivery 7) C-sections 8) Vacuum- or forceps-assisted delivery 9) Chorioamnionitis 10) Endometritis	1) Apgar score ≤ 7 at 5 minutes: PGE_2: 0 Placebo: 1/47 (2.1%) p = not significant 2) Birthweight (mean ± SD): PGE_2: 3765 ± 446 Placebo: 3684 ± 411 p = not significant 3) Birthweight > 4000 g: PGE_2: 14/43 (32.6%) Placebo: 7/47 (14.9%) p < 0.05 4) Shoulder dystocia: PGE_2: 3/43 (7.0%) Placebo: 1/47 (2.1%) p = not significant 5) Gestational age at delivery (mean ± SD): PGE_2: 41.7 ± 0.5 weeks Placebo: 41.6 ± 0.4 weeks p = not significant 6) Time from enrollment to delivery (mean ± SD): PGE_2: 5.5 ± 3.5 days Placebo: 6.0 ± 2.8 days p = not significant 7) C-sections: Overall: PGE_2: 6/43 (14.0%) Placebo: 8/47 (17.0%) p = not significant For fetal distress: PGE_2: 0 Placebo: 1/47 (2.1%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				8) Vacuum- or forceps-assisted delivery: PGE_2: 6/43 (14.0%) Placebo: 3/47 (6.4%) p = not significant 9) Chorioamnionitis: PGE_2: 5/43 (11.6%) Placebo: 2/47 (4.3%) p = not significant 10) Endometritis: PGE_2: 1/43 (2.3%) Placebo: 1/47 (2.1%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
MacKenzie and Burns, 1997	Design: RCT, randomization by computer-generated random numbers and sealed envelopes Interventions: 1) PGE_2 gel, 1 dose (n = 483) Protocol: PGE_2 gel (2 mg) applied vaginally. If labor had not started 14-20 hours after initial treatment, then amniotomy performed and IV oxytocin infusion started 1-2 hours later. If amniotomy not technically possible, it was deferred until 4 hours after oxytocin started. 2) PGE_2 gel, 2 doses (n = 472) Protocol: Same as above, but second dose of PGE_2 gel applied 6 hours after the first if labor not established or cervical score < 9. Dates: NR Location: Oxford, England Setting: Unspecified hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 1000 Dropouts: 45 (excluded due to protocol violations) Loss to follow-up: NA No. of subjects at end: 955 Inclusion criteria: Modified Bishop score ≤ 8; singleton viable pregnancy; cephalic presentation; no previous C-section Exclusion criteria: None specified Age: 1 dose / 2 doses < 20: 5% / 5% 20-29: 57% / 60% 30-39: 36% / 34% ≥ 40: 2% / 1% Race: NR Gestational age at entry (weeks): 1 dose / 2 doses < 40: 21% / 22% 40-42: 74% / 72% > 42: 5% / 6% Dating criteria: NR Parity: 1 dose / 2 doses 0: 49% / 55% 1-2: 46% / 39% ≥ 3: 5% / 6% Bishop score: 1 dose / 2 doses < 4: 25% / 29% 4-5: 44% / 39% ≥ 6: 31% / 31%	1) Apgar score < 8 at 1 minute 2) Apgar score < 5 at 1 minute 3) Apgar score < 8 at 5 minutes 4) Apgar score < 9 at 10 minutes 5) Birthweight 6) Admission to NICU 7) C-sections 8) Time to delivery	1) Apgar score < 8 at 1 minute: 1-dose nulliparae: 38/237 (16%) 2-dose nulliparae: 63/262 (24%) 1-dose multiparae: 43/246(17%) 2-dose multiparae: 37/210 (18%) (no p-value reported) 2) Apgar score < 5 at 1 minute: 1-dose nulliparae: 9/237 (4%) 2-dose nulliparae: 15/262 (6%) 1-dose multiparae: 15/246 (6%) 2-dose multiparae: 7/210 (3%) (no p-value reported) 3) Apgar score < 8 at 5 minutes: 1-dose nulliparae: 1/237 (< 1%) 2-dose nulliparae: 7/262 (3%) 1-dose multiparae: 5/246 (2%) 2-dose multiparae: 3/210 (1%) (no p-value reported) 4) Apgar score < 9 at 10 minutes: 1-dose nulliparae: 0/237 2-dose nulliparae: 3/262 (1.2%) 1-dose multiparae: 2/246 (0.8%) 2-dose multiparae: 0/210 (no p-value reported) 5) Birthweight (mean ± SD): 1-dose nulliparae: 3499 ± 546 g 2-dose nulliparae: 3512 ± 508 g p = 0.783 1-dose multiparae: 3646 ± 483 g 2-dose multiparae: 3642 ± 542 g p = 0.934 6) Admission to NICU: 1-dose nulliparae: 4/237 (2%) 2-dose nulliparae: 13/262 (5%) 1-dose multiparae: 6/246 (2%) 2-dose multiparae: 6/210 (3%) (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (68% of total study population).

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
		Other: Indications for induction: Postterm: 68% Hypertension: 15% Fetal concerns: 6% Maternal health concerns: 1% Maternal request: 8% Past obstetric history: 2%		7) C-sections: 1-dose nulliparae: 35/237 (15%) 2-dose nulliparae: 30/262 (11%) RR = 1.0 (95% CI, 0.90-1.03) 1-dose multiparae: 4/246 (2%) 2-dose multiparae: 5/210 (2%) RR = 0.7 (95% CI, 0.19-2.51) 8) Time to delivery (mean ± SD): 1-dose nulliparae: 1240 ± 540 minutes 2-dose nulliparae: 1197 ± 503 minutes p = 0.358 1-dose multiparae: 927 ± 519 minutes 2-dose multiparae: 785 ± 394 minutes p = 0.001	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Magann, Chauhan, Nevils, et al., 1998	Design: RCT, randomization by table of random numbers and sealed envelopes Interventions: 1) PGE_2 gel (n = 35) Protocol: PGE_2 gel (0.5 mg) placed into cervix on a daily basis. Modified biophysical profile performed, and patient sent home only when monitoring revealed that any contractions caused had begun to dissipate. Labor induced when Bishop score = 8 or when patient reached 42nd week of pregnancy. 2) Membrane stripping (n = 35) Protocol: Membrane stripping performed daily + modified biophysical profile every 3 days. Membranes separated from the lower uterine segment by two circumferential sweeps of examining finger. If cervix unfavorable for stripping, it was stretched by examining finger daily until membrane stripping could be accomplished. Labor induced when Bishop score = 8 or when patient reached 42nd week of pregnancy. 3) Cervical exam (control) (n = 35) Protocol: Gentle cervical exam performed daily + modified biophysical profile every 3 days. Labor induced when Bishop score = 8 or when patient reached 42nd	No. of subjects at start: 105 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 105 Inclusion criteria: ≥ 41 weeks gestation; uncomplicated pregnancy; no contraindications to vaginal delivery; Bishop score ≤ 4 Exclusion criteria: None specified Age (mean ± SD): PGE_2: 24.5 ± 5.2 Stripping: 25.1 ± 5.1 Control: 25.5 ± 5 Race: PGE_2: 71% White, 11% Black, 17% Hispanic Stripping: 74% White, 11% Black, 11% Hispanic, 3% Asian Control: 63% White, 20% Black, 11% Hispanic, 6% Asian Gestational age at entry (mean ± SD): All 3 groups, 41.1 ± 0.1 weeks Dating criteria: LMP, early pelvic exam, auscultation of the fetal heart by U/S stethoscope, and ("in nearly all cases") U/S before 20th week Parity: PGE_2: 0, 74%; 1, 14%; ≥ 2, 11% Stripping: 0, 51%; 1, 31%; ≥ 2, 17% Control: 0, 60%; 1, 26%; ≥ 2, 14%	1) Apgar score < 7 at 5 minutes 2) Birthweight 3) Uterine artery pH (mean) 4) Uterine artery pH < 7.2 5) Admitted to well-baby nursery 6) Inductions at 42 weeks 7) C-sections 8) Forceps-assisted deliveries 9) Total antepartum costs (per group) 10) Total intrapartum costs (per group)	1) Apgar score < 7 at 5 minutes: PGE_2: 1/35 (3%) Stripping: 0 Control: 1/35 (3%) p = 0.6 2) Birthweight (mean ± SD): PGE_2: 3694 ± 419 g Stripping: 3835 ± 489 g Control: 3770 ± 430 g p = 0.19 3) Uterine artery pH (mean ± SD): PGE_2: 7.22 ± 0.05 Stripping: 7.22 ± 0.05 Control: 7.21 ± 0.05 p = 0.77 4) Uterine artery pH < 7.2: PGE_2: 7/35 (20%) Stripping: 8/35 (23%) Control: 7/35 (20%) p = 0.94 5) Admitted to well-baby nursery: PGE_2: 32/35 (91%) Stripping: 33/35 (94%) Control: 35/35 (100%) p = 0.23 6) Inductions at 42 weeks: PGE_2: 7/35 (20%) Stripping: 6/35 (17%) Control: 22 (63%) p < 0.0001 7) C-sections: PGE_2: 8/35 (23%) Stripping: 5/35 (14%) Control: 5/35 (14%) (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on post hoc analysis of proportion of patients induced at 42 weeks. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	week of pregnancy. Dates: Mar-Sep 1996 Location: San Diego, CA Setting: Military hospital Type(s) of providers: General OB/GYN Length of follow-up: None	Bishop score (mean ± SD): PGE_2: 2.6 ± 1 Stripping: 2.8 ± 0.7 Control: 2.6 ± 0.7		8) Forceps-assisted deliveries: PGE_2: 3/35 (9%) Stripping: 4/35 (11%) Control: 5/35 (14%) (no p-value reported) 9) Total antepartum charges (per group): PGE_2: $30,800 Stripping: $7420 Control: $9520 (no p-value reported) 10) Total intrapartum charges (per group): PGE_2: $11,445 Stripping: $9240 Control: $14,735 (no p-value reported)	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Magann, McNamara, Whitworth, et al., 1998	Design: RCT, randomization by table of random numbers and sealed envelopes Interventions: 1) Membrane stripping (n = 33) Protocol: Membrane stripping performed every 3 days by placing a finger through the cervix and performing 2 circumferential sweeps. If the cervix would not admit a finger, then examining finger placed into the cervix every 3 days until the sweeping could be performed. 3) Vaginal exam (control) (n = 32) Protocol: Gentle vaginal exam performed every 3 days. In both groups, treatment continued until spontaneous labor, rupture of the membranes, or completion of 41 weeks' gestation, at which time patient admitted for induction of labor. Dates: NR Location: San Diego, CA, and Jackson, MS Setting: 1 university hospital and 1 military hospital Type(s) of providers: Not specified Length of follow-up: None	No. of subjects at start: 65 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 65 Inclusion criteria: 39 weeks' gestation; negative fetal fibronectin test result; Bishop score ≤ 4; vertex presentation Exclusion criteria: Placenta previa; other contraindications to vaginal delivery Age (mean ± SD): Stripping, 24.5 ± 5; control, 24.3 ± 5.3 Race: Stripping: 64% White, 27% Black, 9% Hispanic Control: 66% White, 22% Black, 6% Hispanic, 6% other Gestational age at entry (mean ± SD): Both groups, 39.00 ± 0.00 Dating criteria: LMP, initial exam, first auscultation of fetal heart tones with an U/S stethoscope, or U/S before 20 weeks Parity: Stripping, 55% nulliparous; control, 56% nulliparous Bishop score (mean ± SD): Stripping, 2.5 ± 0.6; control, 2.6 ± 0.9	1) Birthweight 2) Umbilical artery pH 3) Admission to NICU 4) Gestational age at delivery 5) Bishop score ≥ 8 at delivery 6) Inductions at 42 weeks 7) C-sections 8) Time from admission to delivery	1) Birthweight (mean ± SD): Stripping: 3449 ± 442 g Control: 3531 ± 490 g p = 0.48 2) Umbilical artery pH (mean ± SD): Stripping: 7.24 ± 0.04 Control: 7.23 ± 0.06 p = 0.43 3) Admission to NICU: Stripping: 2/33 (6%) Control: 2/32 (6%) p =1.00 4) Gestational age at delivery (mean ± SD): Stripping: 39.9 ± 0.3 weeks Control: 41.5 ± 0.6 weeks p < 0.0001 5) Bishop score ≥ 8 at delivery: Stripping: 19/33 (58%) Control: 6/32 (19%) p = 0.0002 6) Inductions at 42 weeks: Stripping: 0 Control: 18/32 (56%) p < 0.0001 7) C-sections: Stripping: 4/33 (12%) Control: 5/33 (15%) p = not significant 8) Time from admission to delivery (mean ± SD): Stripping: 10.4 ± 5.5 hours Control: 13.0 ± 7.1 hours p = 0.10	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on reduction in 42-week inductions.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Mahmood, 1989	Design: RCT, randomization by sealed envelope Interventions: 1) PGE_2 gel 2 mg (n = 40), inserted into posterior fornix at 5 PM day before induction; patients monitored for 1 hour after insertion. Second dose if Bishop score < 5 next morning at 9 AM. If no labor or cervical change by 9 AM next day, third insertion. 2) PGE_2 3 mg tablet (n = 40), inserted into posterior fornix. Protocol same as above. Dates: NR Location: Abderdeen, UK Setting: Community hospital Type(s) of providers: NR Length of follow-up: None	No of subjects at start: 80 Drop-outs: 0 Loss to follow-up: NA No of subjects at end: 80 Inclusion criteria: Gestational age 37-43 weeks; singleton pregnancy; vertex presentation; Bishop score < 5 Exclusion criteria: None specified Age (mean ± SD): Gel: 25 ± 4.4 Tablet: 25 ± 5.3 Race: NR Gestational age at entry: NR Dating criteria: NR Parity: NR Bishop score (mean and range): Gel: 2.30 (0-4) Tablet: 2.55 (0-4) Other: 61% induced for prolonged pregnancy	1) Number of insertions required for spontaneous labor 2) Time from insertion to spontaneous labor 3) Posttreatment Bishop score 4) Need for oxytocin 5) Emergent cesarean section 6) Apgar score < 7 at 1 minute 7) Apgar score < 7 at 5 minutes	1) Number of insertions required for spontaneous labor: Gel: 1 insertion: 50% 2 insertions: 50% Tablet: 1 insertion: 20% 2 insertions: 50% 3 insertions: 30% $p < 0.05$ 2) Time from insertion to onset of labor: Gel: 15.1, if spontaneous; 20.6, if induction needed Table: 25.6, if spontaneous; 30.5, if induction needed $p < 0.02$ 3) Posttreatment Bishop score: Gel: 9.5 Tablet: 7.0 $p < 0.05$ 4) Need for oxytocin: Gel: 12.5% Tablet: 50% $p < 0.001$ 5) Emergent cesarean section: Gel: 15% Tablet: 30% p = not significant 6) Apgar score < 7 at 1 minute: Gel: 22% Tablet: 37% 7) Apgar score < 7 at 5 minutes: Gel: 0 Tablet: 2.5%	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: - Dating criteria: - Bishop score: +

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
McColgin, Hampton, McCaul, et al., 1990	Design: RCT, randomization by computer Interventions: 1) Membrane stripping (n = 90) Protocol: Performed weekly by digital separation of 2-3 cm of the membranes from the lower uterine segment using 2 circumferential passes of the examining finger. If cervix long and closed, then stretched digitally until membrane stripping could be accomplished. Treatment continued until patient admitted to labor and delivery or advanced beyond 42 completed weeks' gestation. 2) Cervicovaginal exam (control) (n = 90) Protocol: Weekly atraumatic assessment of the cervix for Bishop scoring. Treatment continued until patient admitted to labor and delivery or advanced beyond 42 completed weeks' gestation. Dates: Mar 1988 - June 1989 Location: Jackson, MS Setting: University hospital Type(s) of providers: General OB/GYN Length of follow-up: None	No. of subjects at start: 209 Dropouts: 29 (excluded post-randomization Loss to follow-up: NA No. of subjects at end: 180 Inclusion criteria: 38 weeks' gestation Exclusion criteria: Placenta previa; low-lying placenta; abnormal fetal presentation; known medical complication; vaginal or cervical infection Age (mean ± SEM): Stripping, 23.06 ± 0.55; control, 23.31 ± 0.58 Race: NR Gestational age at entry: 38 weeks Dating criteria: LMP, early assessment of uterine size, and U/S before 20 weeks Parity: Stripping, 40% nulliparous; control, 50% nulliparous Bishop score (mean ± SEM): Stripping, 3.51 ± 0.24; control, 3.82 ± 0.19 Other: Long/closed cervix: Stripping, 12/90 (13%); control, 10/90 (11%)	1) Fetal deaths 2) Delivery ≥ 42 weeks 3) Days to delivery 4) Delivery within 1 week	1) Fetal deaths: Stripping: 0 Control: 1/90 (1%) (no p-value reported) 2) Delivery ≥ 42 weeks: Stripping: 3/90 (3.3%) Control: 14/90 (15.6%) p < 0.004 3) Days to delivery (mean ± SEM): Stripping: 8.60 ± 0.74 Control: 15.14 ± 0.83 p < 0.001 4) Delivery within 1 week: Stripping: 49/90 (54.5%) Control: 14/90 (15.6%) p < 0.001	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: NA Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Investigators stated that "nulliparous patients and individuals with unfavorable Bishop scores benefited the most from membrane stripping in reduction of postterm pregnancies." No quantitative data provided.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
McColgin, Patrissi, and Morrison, 1990	Design: RCT, method of randomization not specified Interventions: 1) Sweeping membranes (n = 51) performed weekly from 38-42 weeks by digital separation of membranes from lower uterine segment; if cervix closed, "digitally stretched" to allow sweeping. 2) Control (n = 48): Bishop scoring only performed weekly from 38-42 weeks Both groups followed until 42 weeks, when induction scheduled Dates: NR Location: Jackson, MS Setting: Military hospital and university hospital antenatal clinics Type(s) of providers: NR Length of follow-up: None	No of subjects at start: 103 Drop-outs: 4 Loss to follow-up: NA No of subjects at end: 99 Inclusion criteria: Low-risk pregnancy Exclusion criteria: Uncertain dates; abnormal presentation; low-lying placenta; scheduled repeat cesarean; candidates for vaginal birth after cesarean section allowed to participate Age: NR ("comparable") Gestational age at entry: NR ("comparable") Dating criteria: LMP and ultrasound prior to 20 weeks Parity: NR ("comparable") Bishop score: NR ("comparable")	1) Days to delivery 2) Proportion delivering within 1 week 3) Number delivering after 42 weeks 4) Cesarean delivery	1) Days to delivery: Sweeping: 6.7 days Control: 13.3 days p = 0.003 2) Proportion delivering within 1 week: Sweeping: 59% Control: 21% p = 0.003 3) Number delivering after 42 weeks: Sweeping: 2 Control: 6 p = 0.12 4) Cesarean delivery: Sweeping: 7/51 Control: 5/48 p = NS Results similar when analysis restricted to those entering study at 38 weeks. No significant differences seen in group with Bishop score > 5.	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: - Dating criteria: + Bishop score: -

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Misra and Vavre, 1994	Design: RCT, method of randomization not described Interventions: 1) Intracervical PGE_2 gel (0.5 mg) (n = 136) Protocol: Gel administered into cervical canal at 7:30 PM the night before induction. If no labor (3-4 "good intensity" contractions, lasting 40-50 seconds each, every 10 minutes) after 12 hours, then patient induced with oxytocin. Amniotomy performed after cervical dilatation of ≥ 2.5 cm and effacement of ≥ 80%. If no labor after 12 hours and after receiving as much as 64 mU/min of oxytocin, then C-section performed. 2) Oxytocin infusion (n = 127) Protocol: Infusion started at 8:00 AM on day of planned induction, beginning with 2 mU/min and increasing the dose by 1-2 mU every 30 minutes. Amniotomy performed after cervical dilatation of ≥ 2.5 cm and effacement of ≥ 80%. If no labor after 12 hours and after receiving as much as 64 mU/min of oxytocin, then C-section performed. Dates: Aug 1992 - Jan 1994 Location: Bhilai, India Setting: Unspecified hospital Type(s) of providers:	No. of subjects at start: 263 Dropouts: 0 Loss to follow-up: 0 No. of subjects at end: 263 Inclusion criteria: Bishop score < 4; induction of labor required Exclusion criteria: Premature rupture of membranes; "major degrees of cephalopelvic disproportion"; malpresentations; intrauterine deaths; congenital anomalies not compatible with life; persistently nonreactive NST Age (mean ± SD): PGE_2 primigravidas (n =80): 23.7 ± 3.7 PGE_2 multigravidas (n = 56): 25.6 ± 4.0 Oxytocin primigravidas (n = 72): 23.3 ± 2.4 Oxytocin multigravidas (n = 55): 26.3 ± 3.3 Race: NR Gestational age at entry: PGE_2 primigravidas: 39.6 ± 2.7 weeks PGE_2 multigravidas: 39.4 ± 2.1 weeks Oxytocin primigravidas: 39.8 ± 2.0 weeks Oxytocin multigravidas: 39.5 ± 2.3 weeks Dating criteria: NR Parity: PGE_2, 59% primigravidas;	1) Apgar score < 7 at 5 minutes 2) Birthweight 3) Forceps/ventouse deliveries 4) C-sections	1) Apgar score < 7 at 5 minutes: PGE_2 primigravidas: 3/80 (3.8%) Oxytocin primigravidas: 2/72 (2.8%) p = not significant PGE_2 multigravidas: 1/56 (1.8%) Oxytocin multigravidas: 0 p = not significant 2) Birthweight (mean ± SD): PGE_2 primigravidas (n = 80): 2640 ± 580 g Oxytocin primigravidas (n = 72): 2660 ± 550 g p = 0.84 PGE_2 multigravidas (n = 56): 2670 ± 580 g Oxytocin multigravidas (n = 55): 2770 ± 620 g p = 0.38 3) Forceps/ventouse deliveries: PGE_2 primigravidas: 3/80 (3.8%) Oxytocin primigravidas: 4/72 (5.6%) (no p-value reported) PGE_2 multigravidas: 0 Oxytocin multigravidas: 2/55 (3.6%) (no p-value reported) 4) C-sections: PGE_2 primigravidas: 21/80 (26.3%) Oxytocin primigravidas: 34/72 (47.2%) p < 0.01 PGE_2 multigravidas: 7/56 (12.5%) Oxytocin multigravidas: 8/55 (14.6%) p = 0.75	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: - Gestational age: + Dating criteria: - Bishop score: +

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Unspecified OB/GYN Length of follow-up: None	oxytocin, 57% primigravidas Bishop score (mean ± SD): PGE_2 primigravidas: 2.2 ± 0.6 PGE_2 multigravidas: 2.5 ± 0.6 Oxytocin primigravidas: 2.3 ± 0.6 Oxytocin multigravidas: 2.6 ± 0.7			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Mundle and Young, 1996	Design: RCT, randomization by random-number tables and sealed envelopes Interventions: 1) Misoprostol (n = 111) Protocol: 50-µg tablet placed in upper vagina every 4 hours until patient experienced progressive labor, contractions 3 times/minute, ruptured membranes, non-reassuring FHR tracing, or delivery. No more than 16 applications permitted; no change in dosage permitted. 2) PGE_2 gel (n = 111) Protocol: Patient given PGE_2 gel in dose of either 0.5 mg intracervically (for ripening) or 1-2 mg intravaginally (for induction), as determined by treating physician. In both groups, amniotomy was performed at the discretion of the attending physician. Oxytocin administration was begun at 2 mU/min, then increased by 2-mU/min increments at 30-60-min intervals. Oxytocin not permitted within 4 hours of last dose of misoprostol or 6 hours of last dose of PGE_2 gel. Dates: Mar-Sep 1994 Location: St. John's, Newfoundland, Canada Setting: Unspecified hospital	No. of subjects at start: 222 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 222 Inclusion criteria: Indication for induction; single live fetus; gestational age > 37 weeks; cephalic presentation; intact membranes Exclusion criteria: Nonreassuring FHR tracing; prior uterine surgery; know hypersensitivity to misoprostol or other prostaglandins; contraindication to vaginal birth Age (mean ± SD): Misoprostol, 27.6 ± 5.1; PGE_2, 27.4 ± 5.5 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 286.4 ± 7.8 days; PGE_2, 285.5 ± 8.8 days Dating criteria: NR Parity (mean ± SD): Misoprostol, 0.5 ± 0.8; PGE_2, 0.6 ± 0.9 Bishop score (median, with 25% and 75% quartiles): Misoprostol, 4 (2, 5); PGE_2, 4 (2, 6) Other: Indications for induction were as follows: Postterm: 78% Hypertension: 8% Oligohydramnios: 7%	1) Median Apgar scores 2) Apgar score < 7 at 1 minute 3) Apgar score < 7 at 5 minutes 4) Cord pH 5) Base deficit 6) Birthweight 7) Episiotomy 8) Laceration 9) 3rd- or 4th-degree laceration 10) Intact perineum 11) Time from induction to delivery 12) Vacuum-assisted deliveries 13) C-sections	1) Median Apgar scores (with 25% and 75 % quartiles): At 1 minute: Misoprostol: 9 (7, 9) PGE_2: 9 (8, 9) p = 0.67 At 5 minutes: Misoprostol: 9 (8, 9) PGE_2: 9 (9, 10) p = 0.72 2) Apgar score < 7 at 1 minute: Misoprostol: 17/111 (15%) PGE_2: 13/111 (12%) p = 0.43 3) Apgar score < 7 at 5 minutes: Misoprostol: 2/111 (2%) PGE_2: 1/111 (1%) p = 1.00 4) Cord pH (mean ± SD): Misoprostol: 7.28 ± 0.09 PGE_2: 7.28 ± 0.10 p = 0.90 5) Base deficit (mean ± SD): Misoprostol: 5.1 ± 4.0 PGE_2: 5.6 ± 4.5 p = 0.38 6) Birthweight (mean ± SD): Misoprostol: 3728 ± 509 g PGE_2: 3631 ± 493 g (no p-value reported) 7) Episiotomy: Misoprostol: 33/111 (30%) PGE_2: 47/111 (42%) (no p-value reported) RR = 0.72 (95% CI, 0.51-1.02)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (78% of total study population). Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Unspecified hospital Length of follow-up: None	Other: 7%		8) Laceration: Misoprostol: 55/111 (50%) PGE_2: 49/111 (44%) (no p-value reported) RR = 1.16 (95% CI, 0.89-1.51) 9) 3rd- or 4th-degree laceration: Misoprostol: 6/111 (5%) PGE_2: 4/111 (4%) (no p-value reported) RR = 1.55 (95% CI, 0.45-5.31) 10) Intact perineum: Misoprostol: 17/111 (15%) PGE_2: 18/111 (16%) (no p-value reported) RR = 0.97 (95% CI, 0.53-1.78) 11) Time from induction to delivery (mean $\pm$ SD): Misoprostol: 753 $\pm$ 588 minutes PGE_2: 941 $\pm$ 506 minutes p = 0.018 12) Vacuum-assisted deliveries: Misoprostol: 3/111 (3%) PGE_2: 15/111 (14%) (no p-value reported) RR = 0.20 (95% CI, 0.06-0.67) 13) C-sections: Misoprostol: 15/111 (14%) PGE_2: 12/111 (11%) (no p-value reported) RR = 1.25 (95% CI, 0.61-2.55)	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
National Institute of Child Health and Human Develop-ment Network of Maternal-Fetal Medicine Units, 1994	Design: RCT, randomization by computer-generated random numbers Interventions: 1) PGE_2 gel + induction by oxytocin (n = 174) Protocol: PGE_2 gel (0.5 mg) inserted into intracervical canal within 24 hours of randomization. No repeat applications. FHR and uterine contractions monitored continuously for ≥ 4 hours. If no labor after 12 hours, then patient induced using amniotomy (where clinically feasible), followed by oxytocin infusion ("according to a uniform protocol"). If no active labor 24 hours after oxytocin infusion, then C-section performed or induction of labor continued. (Decision to perform C-section not dictated by study protocol.) 2) Placebo gel + induction by oxytocin (n = 91) Protocol: Same as in 1), above, except that placebo gel used instead of PGE_2 gel. 3) Monitoring (n = 175) Protocol: Weekly cervical exam + twice-weekly NST and U/S assessment of AFV. Spontaneous labor awaited, but labor could be induced if: Bishop score > 6; estimated fetal weight > 4500 g; medical or obstetric indication for delivery developed; largest pocket of amniotic fluid < 2	No. of subjects at start: 440 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 440 Inclusion criteria: Gestational age ≥ 287 days and < 301 days Exclusion criteria: Medical or obstetric complications requiring induction, C-section, or frequent monitoring; estimated fetal weight > 4500 g; Bishop score ≥ 7; non-reactive NST; amniotic fluid pocket < 2 cm Age (mean ± SD): PGE_2-oxytocin: 25.4 ± 5.7 Placebo-oxytocin: 25.4 ± 5.3 Monitoring: 26.1 ± 5.8 Race: PGE_2-oxytocin: 67% White, 32% Black, 1% not available Placebo-oxytocin: 63% White, 37% Black Monitoring: 60% White, 38% Black, 2% not available Gestational age at entry: PGE_2-oxytocin: 8I% 287-293 days; 19% 295-301 days Placebo-oxytocin: 79% 287-293 days; 21% 295-301 days Monitoring: 79% 287-293 days; 21% 295-301 days Dating criteria: Any one of following: 1) LMP + audible fetal heartbeat documented for ≥ 21 weeks by fetoscope or ≥ 30	1) Mechanical ventilation 2) Meconium aspiration 3) Nerve injury 4) Seizures 5) ≥ 1 adverse neonatal outcome 6) Apgar score < 4 at 5 minutes 7) Birthweight (mean) 8) Birthweight ≥ 4500 g 9) Time from randomization to delivery 10) Gestational age at delivery 11) Maternal infection 12) Maternal transfusion 13) Hyperstimulation 14) C-sections	1) Mechanical ventilation: PGE_2-oxytocin: 0 Placebo-oxytocin: 1/91 (1%) Monitoring: 1/175 (< 1%) (no p-value reported) 2) Meconium aspiration: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 1/91 (1%) Monitoring: 2/175 (1%) (no p-value reported) 3) Nerve injury: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 0 Monitoring: 0 (no p-value reported) 4) Seizures: PGE_2-oxytocin: 0 Placebo-oxytocin: 2/91 (2%) Monitoring: 1/175 (< 1%) (no p-value reported) 5) ≥ 1 adverse neonatal outcome: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 3/91 (3%) Monitoring: 1/175 (< 1%) (no p-value reported) 6) Apgar score < 4 at 5 minutes: PGE_2-oxytocin: 0 Placebo-oxytocin: 0 Monitoring: 1/175 (< 1%) (no p-value reported) 7) Birthweight (mean ± SD): PGE_2-oxytocin: 3607 ± 382 g Placebo-oxytocin: 3532 ± 464 g Monitoring: 3606 ± 440 g (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size estimates based on perinatal morbidity/mortality and maternal mortality.

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	cm; or abnormal NST followed by positive CST. If NST nonreactive, but CST negative, then testing repeated in 24 hours. Patients undelivered by 308 days (44 completed weeks) were released from the protocol and managed as "appropriate for the clinical situation." Dates: Dec 1987 - July 1989 Location: Multiple sites in US Setting: University hospitals Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	weeks by Doppler; 2) LMP + compatible uterine size estimation at ≤ 24 weeks; 3) LMP + positive pregnancy test obtained early enough to assure that gestation exceeded 41 weeks; 4) if LMP uncertain, then fetal heartbeat documented for ≥ 32 weeks by Doppler; 5) U/S before 26 weeks Parity (% nulliparous): PGE_2-oxytocin: 60% Placebo-oxytocin: 59% Monitoring: 54% Bishop score (mean ± SD): PGE_2-oxytocin: 4.0 ± 1.4 Placebo-oxytocin: 3.8 ± 1.4 Monitoring: 3.9 ± 1.5		8) Birthweight ≥ 4500 g: PGE_2-oxytocin: 1/174 (< 1%) Placebo-oxytocin: 3/91 (3%) Monitoring: 6/175 (4%) (no p-value reported) 9) Time from randomization to delivery (median, with range): PGE_2-oxytocin: 36 hours (6-492) Placebo-oxytocin: 35 hours (7-487) Monitoring: 85 hours (5-538) p < 0.001 10) Gestational age at delivery: 287-293 days / 294-301 days / >302 days PGE_2-oxy: 64% / 34% / 1% Placebo-oxy: 66% / 32% / 2% Monitoring: 38% / 47% / 14% p < 0.001 11) Maternal infection: PGE_2-oxytocin: 33/174 (19%) Placebo-oxytocin: 13/91 (14%) Monitoring: 25/175 (14%) p = not significant 12) Maternal transfusion: PGE_2-oxytocin: 2/174 (1%) Placebo-oxytocin: 0 Monitoring: 3/175 (2%) p = not significant 13) Hyperstimulation: PGE_2-oxytocin: 2/174 (1%) Placebo-oxytocin: 1/91 (1%) Monitoring: 0 p = not significant 14) C-sections: PGE_2-oxytocin: 39/174 (22%) Placebo-oxytocin: 16/91 (18%) Monitoring: 32/175 (18%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
O'Brien, Mercer, Cleary, et al., 1995	Design: RCT, randomization by table of random numbers Interventions: 1) PGE_2 gel (n = 50) Protocol: PGE_2 (2 mg) gel given intravaginally every day for 5 consecutive days. Patients monitored for minimum of 30 minutes after each dose. At 41 weeks, patients re-evaluated. If cervix favorable, NST non-reactive with a BPS ≤ 6, oligohydramnios, FHR decelerations, or evidence of growth restriction, then patient induced. Otherwise, patients evaluated with twice –weekly NSTs and weekly AFV assessments. 2) Placebo gel (n = 50) Protocol: Same as above, except that placebo gel used instead of PGE_2. Dates: June 1993 - June 1994 Location: Memphis, TN Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 100 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 100 Inclusion criteria: 38-40 weeks gestation; Bishop score ≤ 6; no medical indication for delivery; ≤ 1 previous low-transverse C-section Exclusion criteria: Nonreactive NST; oligohydramnios (AFI < 5.0 cm); macrosomia (estimated fetal weight > 4000 g); fetal growth restriction (estimated fetal weight < 10th percentile) Age: NR Race: NR Gestational age at entry (mean ± SD): PGE_2, 38.9 ± 0.54 weeks; placebo, 39.0 ± 0.66 weeks Dating criteria: NR Parity: PGE_2, 40% nulliparous; placebo, 56% nulliparous Bishop score (median, with range): PGE_2, 4 (1-6); placebo, 4 (1-6)	1) Birthweight 2) Macrosomia 3) Apgar score < 7 at 5 minutes 4) Admission to NICU 5) Meconium staining 6) Postdate pregnancies (delivery > estimated date) 7) Postterm pregnancies (delivery ≥ 294 days) 8) Inpatient inductions 9) Gestational age at delivery 10) Chorioamnionitis 11) C-sections	1) Birthweight (mean ± SD): PGE_2: 3320 ± 400 g Placebo: 3450 ± 400 g p = 0.11 2) Macrosomia: PGE_2: 1/50 (2%) Placebo: 4/50 (8%) p = 0.36 3) Apgar score < 7 at 5 minutes: PGE_2: 0 Placebo: 2/50 (4%) p = 0.50 4) Admission to NICU: PGE_2: 1/50 (2%) Placebo: 5/50 (10%) p = 0.20 5) Meconium staining: PGE_2: 8/50 (16%) Placebo: 15/50 (30%) p = 0.15 6) Postdate pregnancies (delivery > estimated date): PGE_2: 20/50 (40%) Placebo: 33/50 (66%) p =0.016 7) Postterm pregnancies (delivery ≥ 294 days): PGE_2: 2/50 (4%) Placebo: 3/50 (6%) p = 1.0 8) Inpatient inductions: PGE_2: 6/50 (12%) Placebo: 14/50 (28%) p = 0.08	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not stratified by parity.

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				9) Gestational age at delivery (mean ± SD): PGE_2: 39.9 ± 1.0 weeks Placebo: 40.5 ± 0.99 weeks p = 0.003 10) Chorioamnionitis: PGE_2: 4/50 (8%) Placebo: 7/50 (14%) p = 0.52 11) C-sections: PGE_2: 7/50 (14%) Placebo: 10/50 (20%) p = 0.59	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Papa-georgiou, Tsionou, Minaretzis, et al., 1992	Design: RCT, allocation to treatment group by even/odd admission number Interventions: 1) PGE_2 gel (n = 83) Protocol: PGE_2 gel (0.5 mg) instilled deeply into cervical canal by syringe. Patient monitored for 45 min before and after treatment. Pelvic exam done 6 hours after placement of gel. If Bishop score < 5, then second dose given. Pelvic exam repeated 6 hours after second dose. If Bishop score still < 5, then patient considered to have failed PGE_2 ripening and given oxytocin infusion. If Bishop score > 5, but regular contractions or progressive dilatation not observed, then oxytocin used for labor augmentation. 2) Oxytocin (n = 82) Protocol: Up to 3 trials of oxytocin infusion, each lasting 4 hours, with 4-hour rest period between trials. Infusion started at 5 mU/min and increased by 5 mU/min every half hour up to 30 mU/min. If no labor established after 3 trials, then patient delivered by C-section. Dates: NR Location: Athens, Greece Setting: University hospital	No. of subjects at start: 165 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 165 Inclusion criteria: Singleton pregnancy; vertex presentation; unripe cervix; no other obstetric complications; 41 completed weeks' gestation; nonreactive NST; normal AFI by U/S Exclusion criteria: None specified Age (mean ± SEM): PGE_2, 24.9 ± 0.5; oxytocin, 25.0 ± 0.5 Race: NR Gestational age at entry: NR (required to have completed 41 weeks' gestation for entry into study) Dating criteria: LMP + U/S at 20 weeks Parity (mean ± SEM): PGE_2, 1.6 ± 0.1; oxytocin, 1.5 ± 0.1 Bishop score (mean ± SEM): PGE_2, 2.9 ± 0.1; oxytocin, 3.1 ± 0.1	1) Apgar score < 7 at 5 minutes 2) Birthweight 3) C-sections for disproportion 4) C-sections for fetal distress 5) Vacuum delivery 6) Vaginal delivery 7) Hyperstimulation	1) Apgar score < 7 at 5 minutes: PGE_2: 2/83 (2.4%) Oxytocin: 8/82 (9.7%) $p < 0.05$ 2) Birthweight (mean ± SEM): PGE_2: 3601 ± 55 g Oxytocin: 3562 ± 43 g p = not significant 3) C-sections for disproportion: PGE_2: 4/83 (4.8%) Oxytocin: 4/82 (4.8%) p = not significant 4) C-sections for fetal distress: PGE_2: 2/83 (2.4%) Oxytocin: 3/82 (3.6%) p = not significant 5) Vacuum delivery: PGE_2: 7/83 (8.4%) Oxytocin: 9/82 (10.9%) p = not significant 6) Vaginal delivery: PGE_2: 74/83 (89%) Oxytocin: 58/82 (70.7%) $p < 0.01$ 7) Hyperstimulation: PGE_2: 2/83 (2.4%) Oxytocin: 4/82 (4.8%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: + Gestational age: - Dating criteria: + Bishop score: + Results not stratified by parity.

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: Unspecified OB/GYN Length of follow-up: None				

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Rayburn, Gosen, Ramadei, et al., 1988	Design: RCT, randomization by drawing a card Interventions: 1) Prostaglandin E_2 (PGE_2) gel (2.5 mg) (n = 55) 2) Placebo gel (n = 63) Treatment protocol: After assignment of Bishop score and a reactive NST, gel instilled into cervix using a 16-gauge angiocatheter tube. Patient remained in semi-Trendelenburg position while uterine contractions and FHR monitored for 2 hours. Induction of labor with oxytocin scheduled approximately 12 hours after instillation of study drug. Induction followed ACOG guidelines. Dates: Dec 1985 - Feb 1987 Location: Omaha, NE Setting: University hospital and military hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 118 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 118 Inclusion criteria: Singleton pregnancy; scheduled for induction at 42 weeks; unfavorable cervix (Bishop score ≤ 5) Exclusion criteria: None specified Age (mean ± SD, with range): PGE_2: 23 ± 1.2 (21.8 to 24.2) Placebo: 24 ± 1.6 (22.4 to 25.6) Race: NR Gestational age at entry: 42 weeks Dating criteria: LMP *plus* "compatible clinical milestones" or U/S results from first half of gestation Parity: PGE_2: 51% nulliparous Placebo: 63% nulliparous Bishop score (mean ± SD): PGE_2: 3.2 ± 1.0 Placebo: 3.4 ± 0.8	1) Vaginal delivery, spontaneous 2) Vaginal delivery, forceps-assisted 3) C-sections 4) Time to delivery	1) Vaginal delivery, spontaneous: PGE_2: 42/55 (76%) Placebo: 35/63 (56%) $p < 0.05$ 2) Vaginal delivery, forceps-assisted: PGE_2: 3/55 (5.5%) Placebo: 7/63 (11%) $p < 0.05$ 3) C-sections: *Overall:* PGE_2: 10/55 (18%) Placebo: 21/63 (33%) $p < 0.05$ *For fetal distress:* PGE_2: 1/55 (2%) Placebo: 6/63 (9.5%) (no p-value reported) *For failure to progress:* PGE_2: 9/55 (16%) Placebo: 13/63 (21%) (no p-value reported) *For other reasons:* PGE_2: 0 Placebo: 2/63 (3%) (no p-value reported) 4) Time to delivery (mean ± SD): PGE_2: 5.5 ± 1.6 hours Placebo: 9.5 ± 2.3 hours $p < 0.01$	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: - Statistical tests: - Gestational age: + Dating criteria: + Bishop score: +

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sala-malekis, Vitoratos, Kassanos, et al., 2000	Design: RCT, method of randomization not described Interventions: 1) Membrane stripping (n = 34) Protocol: Examiner's finger inserted as far as possible through the internal cervical os, separating the membranes from the lower uterine segment and rotating 360º. Patients followed up for 4 days. 2) Oxytocin (n = 35) Protocol: Oxytocin infusion given over 6 hours. Initial infusion 0.5 mU/min, then doubled hourly, reaching a maximum of 4 mU/min. Continuous cardiotocographic monitoring throughout 6-hour infusion period. Patients followed up for 4 days. 3) Vaginal exam (control) (n = 35) Protocol: "Gentle vaginal examination" given. Patients followed up for 4 days. Dates: NR Location: Athens, Greece Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 104 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 104 Inclusion criteria: Primigravida; gestational age 40-41 weeks; Bishop score ≤ 5; no maternal or fetal complications; singleton pregnancy; cephalic presentation Exclusion criteria: None specified Age (mean ± SD): Stripping: 26 ± 2.4 Oxytocin: 27.1 ± 4.5 Control: 26.3 ± 3.8 Race: 100% Greek Gestational age at entry (mean ± SD): Stripping: 283.3 ± 2.4 days Oxytocin: 284.1 ± 2.1 days Control: 282.9 ± 3.2 days Dating criteria: Clinical exam and U/S during 1st trimester Parity: 100% primigravida Bishop score: NR (required to be ≤ 5 for entry into study)	1) C-sections 2) Chorioamnionitis 3) Inductions 4) Spontaneous labor 5) Time to onset of labor	1) C-sections: Stripping: 2/34 (5.9%) Oxytocin: 3/35 (8.6%) Control: 1/35 (2.9%) p = not signifcant 2) Chorioamnionitis: Stripping: 0 Oxytocin: 0 Control: 0 p = not significant 3) Inductions: Stripping: 1/34 (2.9%) Oxytocin: 2/35 (5.7%) Control: 7/35 (20%) p = 0.05 4) Spontaneous labor: Stripping: 23/34 (67.6%) Oxytocin: 18/35 (51.4%) Control: 12/35 (34.2%) p = 0.05 5) Time to onset of labor (mean ± SD): Stripping: 1.9 ± 1.2 days Oxytocin: 2.1 ± 0.8 days Control: 2.5 ± 0.9 days p = not significant	QUALITY SCORE: Randomized: + Method of randomization: - Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Definition of "labor" used not reported.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sanchez-Ramos, Kaunitz, Del Valle, et al., 1993	Design: RCT, randomization by table of random numbers (generated by consecutive coin toss) and sealed envelopes Interventions: 1) Misoprostol (n = 64) Protocol: 50-µg misoprostol tablet placed in posterior vaginal fornix. Dose repeated every 4 hours until adequate labor achieved (3 contractions in 10 minutes). Maximum dose = 600 µg. Artificial rupture of the membranes performed as soon as cervical dilatation permitted. Patients in active labor with arrest of dilatation (no change in dilatation for 2+ hours at 5 cm or more) received oxytocin augmentation. 2) Oxytocin (n = 65) Protocol: Oxytocin infusion started at 1-2 mU/minute and gradually increased in dose increments of 1-2 mU/minute at 30-min intervals, as needed. If Bishop score < 5 before start of oxytocin infusion, then cervical ripening was performed with single or multiple doses of PGE_2 gel. Dates: Jan-Aug 1992 Location: Jacksonville, FL Setting: University hospital Type(s) of providers: Unspecified OB/GYN	No. of subjects at start: 130 Dropouts: 1 (excluded after randomization for breech presentation) Loss to follow-up: NA No. of subjects at end: 129 Inclusion criteria: Obstetric indication for labor; medical complications (including diabetes and renal disease); absence of labor or fetal distress; no previous C-section or other uterine surgery; singleton pregnancy with vertex presentation; no contraindications to vaginal delivery Exclusion criteria: None specified Age (mean + SD): Misoprostol, 23.7 ± 5.5; oxytocin, 23.1 ± 5.6 Race: Misoprostol, 50% non-White; oxytocin, 51% non-White Gestational age at entry (mean ± SD): Misoprostol, 38.8 ± 2.6 weeks; oxytocin, 38.8 ± 4.0 weeks Dating criteria: NR Parity (mean ± SD): Misoprostol, 0.8 ± 1.2; oxytocin, 0.7 ± 1.1 Bishop score (mean ± SD): Misoprostol, 4.0 ± 2.2; oxytocin, 4.2 ± 2.2 Other: Indications for induction were as follows: Preeclampsia: 34%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Cord pH < 7.16 5) Admission to NICU 6) Bleeding > 500 ml 7) Forceps delivery 8) Vacuum delivery 9) C-sections 10) Induction-agent costs 11) Time to delivery	1) Apgar score < 7 at 1 minute: Misoprostol: 11/64 (17.2%) Oxytocin: 9/65 (13.8%) p = not significant 2) Apgar score < 7 at 5 minutes: Misoprostol: 1/64 (1.6%) Oxytocin: 1/65 (1.5%) p = not significant 3) Birthweight (mean ± SD): Misoprostol: 3181.5 ± 731.8 g Oxytocin: 3231.4 ± 662.8 g p = not significant 4) Cord pH < 7.16: Misoprostol: 9/64 (14.1%) Oxytocin: 7/65 (10.8%) p = not significant 5) Admission to NICU: Misoprostol: 3/64 (4.7%) Oxytocin: 6/65 (9.2%) p = not significant 6) Bleeding > 500 ml: Misoprostol: 1/64 (1.6%) Oxytocin: 0/65 p = not significant 7) Forceps delivery: Misoprostol: 9/64 (14.1%) Oxytocin: 9/65 (13.8%) p = not significant 8) Vacuum delivery: Misoprostol: 4/64 (6.3%) Oxytocin: 7/65 (10.8%) p = not significant 9) C-sections: Misoprostol: 14/64 (21.9%) Oxytocin: 14/65 (21.5%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (19% of total study population). Results not stratified by parity. Study underpowered to detect differences in some outcomes (e.g., hyperstimulation 11% in misoprostol group, 4.6% in oxytocin group, but not significant). Sample size estimates based on time to delivery. Total dose and maximum rate of oxytocin significantly lower in misoprostol group. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None	Postterm: 19% (22% misoprostol, 15% oxytocin) PROM: 13% Abnormal fetal testing: 9% Diabetes: 7% Other: 18%		10) Induction-agent costs (per patient): Misoprostol (± oxytocin): $49 Oxytocin alone: $205 Oxytocin + PGE_2: $315 (no p-value reported) 11) Time to delivery (mean ± SD): Misoprostol: 661.9 ± 435.9 minutes Oxytocin: 1104.9 ± 968.1 minutes p = 0.004	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Satin, Hankins, and Yeomans, 1991	Design: RCT, randomization by sealed envelope Interventions: 1) Oxytocin, slow dose escalation (n = 32) Protocol: Initial dose 2 mU/min. Incremental increases of 1 mU/min given at 30-minute intervals to maximum dose of 40 mU/min. 2) Oxytocin, fast dose escalation (n = 48) Protocol: : Initial dose 2 mU/min. Incremental increases of 2 mU/min given at 15-minute intervals to maximum dose of 40 mU/min. In both groups, oxytocin doses were increased until an adequate labor pattern was achieved (defined as labor resulting in cervical change). Amniotomy performed in active labor. Internal FHR and pressure monitored. Pressure catheter used to titrate. Induction considered to have failed if no cervical dilatation or spontaneous rupture of membranes by 8-10 hours and no evidence of fetal distress or maternal illness. Dates: NR Location: San Antonio, TX Setting: Military hospital Type(s) of providers: Unspecified OB/GYN	No. of subjects at start: 80 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 80 Inclusion criteria: Cervical dilatation ≤ 2 cm; Bishop score ≤ 6; no regular uterine activity; intact membranes Exclusion criteria: Malpresentation; placenta previa; active herpes infection; hypertension; deviation from dosing protocol Age (mean + SD): Slow, 24.3 ± 3.6; fast, 24.7 ± 3.2 Race: NR Gestational age at entry: Mean NR. Slow, 31/32 (97%) ≥ 42 weeks; fast, 44/48 (92%) ≥ 42 weeks Dating criteria: NR Parity: Slow, 47% nulliparous; fast, 46% nulliparous Bishop score: NR	1) Apgar score ≤ 3 at 1 minute 2) Apgar score ≤ 6 at 5 minutes 3) Birthweight 4) Use of epidural 5) Induction failure 6) Hyperstimulation/FHR abnormalities requiring oxytocin to be stopped 7) C-sections (by parity) 8) Mid-forceps delivery (by parity) 9) Time to delivery (by parity)	1) Apgar score ≤ 3 at 1 minute: Slow: 0/32 Fast: 1/48 (2%) p = not significant 2) Apgar score ≤ 6 at 5 minutes: Slow: 1/32 (3%) Fast: 1/48 (2%) p =not significant 3) Birthweight (mean ± SD): Slow: 3623 ± 459 g Fast: 3670 ± 516 p = not significant 4) Use of epidural: Slow: 25% Fast: 27% p = not significant 5) Induction failure: Slow: 10/32 (31%) Fast: 4/48 (8%) p < 0.05 6) Hyperstimulation/FHR abnormalities requiring oxytocin to be stopped: Slow: 66%, 0 episodes; 25%, 1 episode; 3%, 2 episodes; 6%, ≥ 3 episodes Fast: 46%, 0 episodes; 29%, 1 episode; 8%, 2 episodes; 17%, ≥ 3 episodes p = not significant 7) C-sections (by parity): Slow, nulliparous: 1/32 (3%) Slow, multiparous: 0 Fast, nulliparous: 3/48 (6%) Fast, multiparous: 2/48 (4%) p = not significant 8) Mid-forceps delivery (by parity): Slow, nulliparous: 1/32 (3%) Slow, multiparous: 1/32 (3%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: - Hyperstimulation more common in fast protocol, but study underpowered to detect difference *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Length of follow-up: None			Fast, nulliparous: 2/48 (4%) Fast, multiparous: 0 p = not significant 9) Time to delivery (mean, by parity): Slow, nulliparous: 15 hours, 18 minutes Fast, nulliparous: 9 hours, 16 minutes $p < 0.05$ Slow, multiparous: 10 hours, 54 minutes Fast, multiparous: 8 hours, 2 minutes $p < 0.05$	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sawai, O'Brien, Mastro-giannis, et al., 1994	Design: RCT, computer-generated randomization Interventions: 1) Self-administered PGE_2 suppositories (2 mg) (n = 38) Protocol: Patients given explicit instructions on how to avoid intracervical placement of suppository. Enough suppositories given for daily use until next clinic visit. Telephone contact with investigator available on 24-hour basis. Patients returned for weekly sonogram for AFI and twice-weekly NST and Bishop scoring. Suppositories dispense at each clinic visit until spontaneous labor occurred or until patient admitted for induction of labor for Bishop score ≥ 9, oligohydramnios (AFI < 5 cm), "nonreassuring" FHR tracing, gestational age of 44 weeks, or the development of preeclampsia or other exclusion criteria. 2) Placebo suppositories (n = 42) Protocol: Same as above, except that placebo suppositories used instead of PGE_2. Dates: May 1990 - Sep 1991 Location: Tampa, FL Setting: University hospital Type(s) of providers:	No. of subjects at start: 91 Dropouts: 11 Loss to follow-up: NA No. of subjects at end: 80 Inclusion criteria: Gestational age ≥ 41 weeks; uncomplicated pregnancy; Bishop score < 9; reactive NST; normal U/S Exclusion criteria: Maternal medical problems; previous uterine surgery; previous stillbirth; abnormal FHR; vaginal bleeding; spontaneous rupture of membranes; regular uterine contractions; abnormal U/S findings; estimated fetal weight ≥ 4500 g Age: NR Race: NR Gestational age at entry (mean ± SD): PGE_2, 297.0 ± 5.4 days; placebo, 295.0 ± 4.5 days (p = 0.021) Dating criteria: NR ("reliable dating criteria") Parity: NR Bishop score: Baseline scores not reported	1) Apgar score < 7 at 5 minutes 2) Birthweight 3) Umbilical artery pH 4) Admission to NICU 5) C-sections 6) Chorioamnionitis 7) Time from admission to delivery 8) Antepartum testing charges (per patient)	1) Apgar score < 7 at 5 minutes: PGE_2: 1/38 (2.6%) Placebo: 1/42 (2.4%) p = not significant 2) Birthweight (mean ± SD): PGE_2: 3.50 ± 0.40 kg Placebo: 3.68 ± 0.39 kg p = 0.051 3) Umbilical artery pH (mean ± SD): PGE_2: 7.27 ± 0.07 Placebo: 7.27 ± 0.07 p = not significant 4) Admission to NICU: PGE_2: 2/38 (5.3%) Placebo: 4/42 (9.5%) p = not significant 5) C-sections: PGE_2: 1/38 (2.6%) Placebo: 6/42 (14.3%) p = not significant 6) Chorioamnionitis: PGE_2: 2/38 (5.3%) Placebo: 10/42 (24%) p = 0.04 7) Time from admission to delivery (mean ± SD): *Nulliparas:* PGE_2 (n = NR): 10.7 ± 5.1 hours Placebo (n = NR): 15.3 ± 7.6 hours p = 0.035 *Multiparas:* PGE_2 (n = NR): 11.2 ± 1.3 hours Placebo (n = NR): 7.1 ± 4.4 hours p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: - Baseline characteristics not reported. Underpowered to detect differences in categorical variables. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Unspecified OB/GYN Length of follow-up: None			8) Antepartum testing charges (per patient; mean ± SD): *All patients:* PGE_2: $476.97 ± $170.36 Placebo: $647.29 ± $257.36 p = 0.001 *Nulliparas:* PGE_2 (n = NR): $456.44 ± $141.55 Placebo (n = NR): $659.67 ± $271.38 p = 0.006 *Multiparas:* PGE_2 (n = NR): $495.45 ± $194.50 Placebo (n = NR): $630 ± $244.12 p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sawai, Williams, O'Brien, et al., 1991	Design: RCT, randomization by sealed envelope Interventions: 1) PGE_2 gel (n = 24; 14 nulliparas and 10 multiparas) Protocol: PGE_2 gel (2 mg) placed in the posterior vaginal fornix. Uterine activity and FHR tracings monitored for 1-2 hours after gel insertion. If no regular uterine contractions and NST reactive, then patient discharged and asked to return for weekly sonograms for AFI assessment and twice-weekly NSTs, cervical scoring, and application of gel. Labor induced if spontaneous labor did not occur and Bishop score > 9, if oligohydramnios present (AFI < 5 cm), if FHR tracing "not reassuring," or if a gestational age of 44 weeks was reached. 2) Placebo gel (n = 26; 16 nulliparas and 10 multiparas) Protocol: Same as above, except that placebo gel used instead of PGE_2. Dates: Aug 1988 - Aug 1989 Location: Tampa, FL Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 50 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 50 Inclusion criteria: Gestational age ≥ 287 days; unfavorable cervix (Bishop score < 9) Exclusion criteria: Diabetes; hypertension; previous uterine surgery; abnormal FHR tracings; vaginal bleeding; spontaneous rupture of membranes; regular uterine contractions; nonvertex presentation; macrosomia (estimated fetal weight > 4500 g); fetal anomalies; fetal growth retardation; oligohydramnios; multiple gestation Age: NR Race: NR Gestational age at entry: NR (gestational age ≥ 287 days required for entry into study) Dating criteria: LMP confirmed by early clinical exam and/or early U/S Parity: PGE_2, 58% nulliparous; placebo, 62% nulliparous Bishop score: NR; score < 9 required for entry into study	1) Apgar scores at 1 minute 2) Apgar scores at 5 minutes 3) Birthweight 4) Umbilical arterial blood pH 5) Admission to NICU 6) C-sections 7) Length of labor and delivery	1) Apgar scores at 1 minute (median): PGE_2 nulliparas (n = 14): 9.0 Placebo nulliparas (n = 16): 8.5 p = not significant PGE_2 multiparas (n = 10): 9.0 Placebo multiparas (n = 10): 9.0 p = not significant 2) Apgar scores at 5 minutes (median): PGE_2 nulliparas: 9.0 Placebo nulliparas: 9.0 p = not significant PGE_2 multiparas: 9.0 Placebo multiparas: 9.0 p = not significant 3) Birthweight (mean ± SEM): PGE_2 nulliparas: 3753.6 ± 126 Placebo nulliparas: 3910.7 ± 113 p = not significant PGE_2 multiparas: 3564.5 ± 119 Placebo multiparas: 3589.0 ± 74 p = not significant 4) Umbilical arterial blood pH (mean ± SEM): PGE_2 nulliparas: 7.28 ± 0.02 Placebo nulliparas: 7.28 ± 0.02 p = not significant PGE_2 multiparas: 7.32 ± 0.01 Placebo multiparas: 7.19 ± 0.06 p = not significant 5) Admission to NICU: PGE_2 nulliparas: 0 Placebo nulliparas: 0 p = not significant PGE_2 multiparas: 0	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: - Study underpowered to detect differences in categorical variables (e.g., C-sections).

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				Placebo multiparas: 2 p = not significant 6) C-sections (all for failure to progress or arrest of descent): PGE_2 nulliparas: 6/14 (43%) Placebo nulliparas: 3/16 (19%) p = not significant PGE_2 multiparas: 0 Placebo multiparas: 1/10 (10%) p = not significant 7) Length of labor and delivery (mean ± SEM): PGE_2 nulliparas: 17.6 ± 2.7 hours Placebo nulliparas: 13.9 ± 1.9 hours p = not significant PGE_2 multiparas: 5.4 ± 2.0 hours Placebo multiparas: 8.2 ± 1.2 hours p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Sciscione, Nguyen, Manley, et al., 2001	Design: RCT, randomization by computer-generated list of random numbers and sealed envelopes Interventions: 1) Transcervical Foley catheter (n = 58) Protocol: 16F Foley catheter with 30-ml balloon inserted into endocervical canal under direct visualization via a sterile speculum exam. Effort was made not to touch the catheter to vagina or ectocervix. Once balloon in place, 30 ml water injected. Traction applied by taping end of catheter to patient's leg. Catheter checked for extrusion every 6 hours by cervical exam. If not extruded, then catheter adjusted to maintain traction. FHR monitoring started after placement, and patient allowed to ambulate. Oxytocin given after catheter extrusion, beginning a 1 mIU and increasing 1 mIU every 15 minutes. Artificial rupture of membranes done as soon as clinically feasible. 2) Misoprostol (n = 53) Protocol: 50-µg tablet placed in posterior vaginal fornix every 4 hours to maximum of 6 doses. Dosing suspended in the event of onset of labor, uterine tachysystole, non-reassuring FHR, or rupture of membranes. Oxytocin started (as above) 4 hours after last dose of misoprostol in women	No. of subjects at start: 114 Dropouts: 3 (2 for protocol violations; 1 for failure to meet inclusion criteria) Loss to follow-up: NA No. of subjects at end: 111 Inclusion criteria: Admitted for labor induction; single gestation; vertex presentation; > 28 weeks' gestation; Bishop score < 6 Exclusion criteria: Rupture of membranes; antepartum bleeding; active genital herpes infection; fetal death; placenta previa; previous induction or preinduction agent during pregnancy; known allergy to misoprostol Age (mean ± SD): Catheter, 25.1 ± 6.9; misoprostol, 25.9 ± 6.9 Race: NR Gestational age at entry: NR Dating criteria: NR Parity: Catheter, 70.6% nulliparous; misoprostol, 71.7% nulliparous Bishop score (median): Catheter, 3.0; misoprostol, 2.0 Other: Indications for induction: Preeclampsia: 32% Oligohydramnios: 25% Postterm: 14% Growth restriction: 8%	1) Birthweight 2) C-sections 3) Delivery within 24 hours 4) Vaginal delivery within 24 hours	1) Birthweight (mean ± SD): Catheter: 2979.5 ± 619.9 g Misoprostol: 2969.8 ± 743.7 g p = 0.94 2) C-sections: *Overall:* Catheter: 31.8% Misoprostol: 37.8% p = 0.46 *For nonreassuring FHR tracing::* Catheter: 12% Misoprostol: 24% p = 0.09 3) Delivery within 24 hours: Catheter: 54.5% Misoprostol: 67.9% p = 0.31 4) Vaginal delivery within 24 hours: Catheter: 73% Misoprostol: 84% p = 0.23	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: - Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (14% of total study population). Results not stratified by parity. Sample size estimates based on change in Bishop score. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	not in active labor, but with Bishop scores > 5, or after 6 doses. Artificial rupture of membranes done as soon as clinically feasible. Dates: July 1997 - July 1999 Location: Newark, DE Setting: Community hospital Type(s) of providers: General OB/GYN; residents Length of follow-up: None	Elective: 5% Chronic hypertension: 3% Diabetes: 3% Macrosomia: 3% Other: 8%			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Srisomboon, Piyamongkol, and Aiewsakul, 1997	Design: RCT, blocked randomization Interventions: 1) Intracervical misoprostol (n = 50) Protocol: 100 µg misoprostol pill crushed in 3 ml sterile jelly. Mixture instilled in endocervical canal with assistance of speculum visualization. 2) Intravaginal misoprostol (n = 50) Protocol: Same mixture as above, but placed in posterior vaginal fornix. Patients in both groups were left in supine position for 1 hour after administration of gel. Vital signs and side effects monitored every 2 hours. Continuous external cardiotocography performed. Patients re-examined at 12 hours. If cervix unfavorable, then 2nd dose of gel given. If cervix became favorable (≥ 6), then amniotomy performed and oxytocin infusion started, if needed. Oxytocin also started if no cervical change occurred or no uterine contractions occurred after 2nd dose. Infusion started at 1-2 mU/min, increased 1-2 mU/min at 30-min intervals. Dates: Aug 1994 - Sep 1995 Location: Chiang Mai, Thailand	No. of subjects at start: 100 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 100 Inclusion criteria: Singleton pregnancy; parity ≤ 3; vertex presentation; obstetric or medical indication for delivery; intact membranes with no prior stripping; Bishop score ≤ 4; gestational age > 35 weeks; no previous C-section or other uterine surgery; no labor or fetal distress; no evidence of cephalo-pelvic disproportion; no placenta previa, forelying cord, or vasa previa; no contraindication to the use of prostaglandins Exclusion criteria: None specified Age (mean ± SD): Intracervical, 25.8 ± 5.3; intravaginal, 28.1 ± 5.8 Race: NR Gestational age at entry (mean ± SD): Intracervical, 39.7 ± 2.2; intravaginal, 39.2 ± 2.2 Dating criteria: NR Parity (mean ± SD): Intracervical, 1.3 ± 0.5; intravaginal, 1.4 ± 0.5 Bishop score (mean ± SD): Intracervical, 2.6 ± 0.8; intravaginal, 2.6 ± 0.9	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Forceps delivery 5) Vacuum delivery 6) C-sections 7) Post-partum hemorrhage 8) Time to delivery	1) Apgar score < 7 at 1 minute: Intracervical: 3/50 (6%) Intravaginal: 0/50 p = not significant 2) Apgar score < 7 at 5 minutes: Intracervical: 0/50 Intravaginal: 0/50 p = not significant 3) Birthweight (mean ± SD): Intracervical: 2823 ± 426 g Intravaginal: 2833 ± 505 g p = not significant 4) Forceps delivery: Intracervical: 2/50 (4%) Intravaginal: 5/40 (10%) p = not significant 5) Vacuum delivery: Intracervical: 8/50 (16%) Intravaginal: 9/50 (18%) p = not significant 6) C-sections: Intracervical: 3/50 (6%) Intravaginal: 5/50 (10%) p = not significant 7) Post-partum hemorrhage: Intracervical: 0/50 Intravaginal: 1/50 (2%) p = not significant 8) Time to delivery (mean ± SD): Intracervical: 17.0 ± 8.6 hours Intravaginal: 16.4 ± 8.6 hours p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: - Statistical tests: - Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (34% of total study population). Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	Other: Indications for induction were as follows: Postterm: 34% (40% intra-cervical, 28% intravaginal) Pregnancy-induced hyper-tension: 31% IUGR: 26% Other: 9%			

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Stenlund, Ekman, Aedo, et al., 1999	Design: RCT, randomization by table of random numbers and sealed envelope Interventions: 1) Mifepristone 400 mg (n = 24) Protocol: Bishop score, U/S, and, "in some cases," Doppler performed before starting treatment. Mifepristone 400 mg given as two tablets. If labor did not start, patients returned to hospital at 24 and 48 hours for assessment of Bishop score and FHR monitoring (30 minutes). If Bishop score ≥ 6 at 48 hours and no labor, then labor induced by amniotomy and oxytocin infusion. If Bishop score < 6, then patient given PGE_2 (0.5 mg) intracervically, repeated 12 hours later, if necessary. 2) Placebo (n = 12) Protocol: Same as above, except that identical placebo substituted for mifepristone. Dates: NR Location: Stockholm, Sweden Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 36 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 36 Inclusion criteria: Indication for induction; induction deferrable for 48 hours; Bishop score ≤ 5; single pregnancy in vertex presentation; intact membranes Exclusion criteria: Contraindication to vaginal delivery; oligohydramnios; prior uterine surgery; parity > 4; renal failure; hepatic disorder; adrenal insufficiency; blood-clotting disorder; anticoagulant or corticosteroid therapy during pregnancy Age (mean ± SD): Mifepristone, 27.4 ± 4.6; placebo, 30.3 ± 5.8 Race: NR Gestational age at entry (mean ± SD): Both groups, 295 ± 4 days Dating criteria: U/S performed in week 16 or 17 Parity: Mifepristone, 79% nulliparous; placebo, 58% nulliparous Bishop score (median, with range): Mifepristone, 3 (0 to 5); placebo, 3 (1 to 5)	1) Apgar scores 2) Birthweight 3) Umbilical pH 4) Seizure requiring anticonvulsant treatment 5) Time to onset of labor 6) Percent in labor by 48 hours 7) Labor or ripe cervix within 48 hours 8) Need for PGE_2: 9) C-sections 10) Vacuum extraction 11) Duration of labor	1) Apgar scores: Median Apgar scores were significantly ($p < 0.05$) lower at 1 minute in the mifepristone group, but did not differ between the two treatment groups at 5 or 10 minutes. (Actual scores NR.) 2) Birth weigh (mean ± SD): Mifepristone: 3881 ± 323 g Control: 3779 ± 438 (no p-value reported) 3) Umbilical pH (mean ± SD): Mifepristone (N = 21/24): 7.12 ± 0.15 Control: 7.19 ± 0.09 p = 0.08 4) Seizure requiring anticonvulsant treatment: Mifepristone: 1/24 (4%) Control: 0 (no p-value reported) 5) Time to onset of labor (median, with range): Mifepristone: 24 hrs, 10 min (1 hr, 50 min to 94 hrs, 45 min) Control: 52 hrs (11 hrs, 15 min to 94 hrs, 45 min) (no p-value reported) 6) Percent in labor by 48 hours (with 95% CI): Mifepristone: 81.8% (65.7% to 97.9%) Control: 27.3% (1.0% to 53.6%) $p < 0.05$ 7) Labor or ripe cervix within 48 hours: Mifepristone: 83.3% Control: 41.7% p = 0.008	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size discussed for primary outcome, but not for secondary outcomes Sample size estimates based on proportion of women delivering within 48 hours and on change in Bishop score. Large discrepancy in parity between two groups (more multiparas in mifepristone group). Results not stratified by parity. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				8) Need for PGE_2: Mifepristone: 17% Control: 58% $p < 0.05$ 9) C-sections (all for fetal distress): Mifepristone: 17% Control: 25% p = not significant 10) Vacuum extraction: Mifepristone: 33% Control: 8% p = not significant 11) Duration of labor (median): Mifepristone: 13 hrs, 39 min Control: 8 hrs, 9 min p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Surbek, Boesiger, Hoesli, et al., 1997	Design: RCT, randomization performed by pharmacy using random-numbers table Interventions: 1) Misoprostol (n = 50) Protocol: 50-µg misoprostol gelatin capsule placed in posterior vaginal fornix. If adequate contraction pattern not achieved, then further doses given at 6 hours, 24 hours, and 30 hours. Patients not in labor at 48 hours received IV oxytocin. 2) Oxytocin (n = 50) Protocol: Same as above, except that PGE_2 3-mg capsules used instead of misoprostol. Dates: Jan-Nov 1995 Location: Basel, Switzerland Setting: University hospital Type(s) of providers: Unspecified OB/GYN, residents, and midwives Length of follow-up: None	No. of subjects at start: 103 Dropouts: 3 (excluded due to protocol violations) Loss to follow-up: NA No. of subjects at end: 100 Inclusion criteria: Bishop score ≤ 5; reactive stress test; singleton vertex presentation; no labor Exclusion criteria: Fetal mal-presentation; C-section or other prior uterine surgery; contra-indications to prostaglandins Age (mean ± SD): Misoprostol, 28.8 ± 5.4; PGE_2, 30.4 ± 4.7 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 40 ± 1.63 weeks; PGE_2, 40 ± 2.0 Dating criteria: NR Parity: Misoprostol, 60% nulliparous; PGE_2, 50% nulliparous Bishop score (mean ± SD): Misoprostol, 2.4 ± 1.35; PGE_2, 3.0 ± 1.64 Other: Indications for induction: PROM: 37% Postterm: 32% IUGR/oligohydramnios: 14% Hypertensive disorder: 6% Diabetes mellitus: 6% Psychosocial: 5%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Cord arterial pH 5) Admission to NICU 6) Vaginal operative delivery 7) C-sections	1) Apgar score < 7 at 1 minute: Misoprostol: 4/50 (8%) PGE_2: 6/50 (12%) p = not significant 2) Apgar score < 7 at 5 minutes: Misoprostol: 0/50 PGE_2: 0/50 p = not significant 3) Birthweight (mean ± SD): Misoprostol: 3360 ± 602 g PGE_2: 3419 ± 659 g p = not significant 4) Cord arterial pH (mean ± SD): Misoprostol: 7.25 ± 0.09 PGE_2: 7.23 ± 0.09 p = not significant 5) Admission to NICU: Misoprostol: 0/50 PGE_2: 3/50 (6%) p = not significant 6) Vaginal operative delivery: Misoprostol: 10/50 (20%) PGE_2: 6/50 (12%) p = not significant 7) C-sections: Misoprostol: 6/50 (12%) PGE_2: 7/50 (14%) p = not significant	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (32% of total study population). Results not stratified by parity. Tachysystole less common in PGE_2 group (8% vs. 14%), but difference not significant. Sample size estimates based on proportion of patients delivering within 24 hours.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Toppozada, Anwar, Hassan, et al., 1997	Design: RCT, randomization by computer-generated table of random numbers Interventions: 1) Vaginal misoprostol (n = 20) Protocol: 100-µg tablet applied intravaginally. If positive response (3 contractions/10 minutes, each lasting 45 seconds and inducing changes in the Bishop score), then dose repeated every 3 hours until cervix ≥ 5 cm. If no response to first dose, then 100-µg dose repeated at 3 hours, and 200-µg dose given every 3 hours thereafter until positive response achieved (up to max of 1000 µg). 2) Oral misoprostol (n = 20) Protocol: Same as above, except that tablets administered orally and second dose (rather than third) doubled if no response to first. In both groups, AROM performed and oxytocin given when cervix ≥ 5 cm. Dates: NR Location: Alexandria, Egypt Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 40 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 40 Inclusion criteria: Indication for induction (diabetes, pregnancy-induced hypertension, or postdates); gestational age 37-42 weeks; single viable pregnancy; vertex presentation; Bishop score ≤ 4 Exclusion criteria: Contra-indication to induction or prostaglandins Age (mean ± SD): Vaginal, 27.5 ± 4.51; oral, 29.15 ± 5.40 Race: NR Gestational age at entry (mean ± SD): Vaginal, 40.30 ± 1.87 weeks; oral, 40.85 ± 1.57 weeks Dating criteria: NR Parity (mean ± SD): Vaginal, 0.80 ± 0.95; oral, 1.25 ± 1.16 Bishop score (mean ± SD): Vaginal, 2.25 ± 1.69; oral, 1.85 ± 1.39 Other: Indications for induction were diabetes, pregnancy-induced hypertension, or postdates. Proportion of patients in each category not reported.	1) Forceps deliveries 2) Vacuum deliveries 3) C-sections	1) Forceps deliveries: Vaginal: 1/20 (5%) Oral: 0/20 (no p-value reported) 2) Vacuum deliveries: Vaginal: 3/20 (15%) Oral: 2/20 (10%) (no p-value reported) 3) C-sections: Vaginal: 2/20 (10%) Oral: 4/20 (20%) (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: - Mode of delivery: + Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Proportion of patients who were induced for postterm pregnancy not reported. No separate results reported for this subgroup. Significantly higher incidence of uterine activity and FHR tracing abnormalities in vaginal group.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Varaklis, Gumina, and Stubble-field, 1995	Design: RCT, randomization by table of random numbers and sealed envelopes Interventions: 1) Misoprostol (n = 36) Protocol: 25 µg given intravaginally every 2 hours for a maximum of 6 doses or until patient experience 3 contractions per 10 minutes. 2) PGE_2 gel (n = 33) Protocol: 0.5 mg placed intracervically. Second dose given after 6 hours if patient not having 3 contractions per 10 minutes. In both groups, no further agents were administered once contraction rate reached 3 per 10 minutes. Oxytocin started 12 hours after first dose of induction agent if patient not in active labor. AROM performed at 3 cm. Dates: NR Location: Portland, ME Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 80 Dropouts: 11 Loss to follow-up: NA No. of subjects at end: 6 Inclusion criteria: Medical indication for induction Exclusion criteria: Severe oligohydramnios; nonreactive stress test; prior uterine surgery; malpresentation; multiple gestation; > 3 contractions per 10 minutes; Bishop score > 5 Age (mean ± SD): Misoprostol, 26.75 ± 5.95; PGE_2, 38.96 ± 1.89 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.52 ± 2.4 weeks; PGE_2, 38.96 ± 1.89 weeks Dating criteria: Last menstrual period Parity (mean ± SD): Misoprostol, 0.44 ± 0.70; PGE_2, 0.67 ± 1.34 Bishop score: Median, 3 in both groups Other: Reasons for induction not described in detail. Investigators stated that "the reasons for induction, most frequently prolonged pregnancy, were similar in both groups."	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Cord arterial pH 5) Assisted vaginal deliveries 6) C-sections 7) Time to vaginal delivery	1) Apgar score < 7 at 1 minute: Misoprostol: 7/36 (19%) PGE_2 gel: 7/33 (21%) p = 0.855 2) Apgar score < 7 at 5 minutes: Misoprostol: 1/36 (3%) PGE_2 gel: 1/33 (3%) p = 1.000 3) Birthweight (mean ± SD): Misoprostol: 3.2 ± 0.84 kg PGE_2 gel: 3.33 ± 0.72 kg p = 0.505 4) Cord arterial pH (mean ± SD): Misoprostol: 7.31 ± 0.05 PGE_2 gel: 7.30 ± 0.08 p = 0.632 5) Assisted vaginal deliveries: Misoprostol: 6/36 (17%) PGE_2 gel: 11/33 (33.3%) (no p-value reported) 6) C-sections: Misoprostol: 8/36 (22%) PGE_2 gel: 3/33 (9%) (no p-value reported) 7) Time to vaginal delivery (mean ± SD): Misoprostol: 15.7 ± 8.1 hours PGE_2 gel: 20.7 ± 8.1 hours p = 0.023	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Proportion of patients who were induced for postterm pregnancy not reported. No separate results reported for this subgroup. Results not stratified by parity. Study underpowered to detect differences in categorical outcomes.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Voss, Cumminsky, Cook, et al., 1996	Design: RCT, randomization by computer-generated random number tables Interventions: 1) PGE_2 gel (0.125 mg) (n = 79) Protocol: FHR and contractions monitored for 30 min before treatment, and Bishop score assessed. Gel (2 ml) inserted into cervix at level of internal cervical os. Monitoring continued for 4 hours after insertion. If no labor and Bishop score ≤ 6 at end of 4-hour monitoring period, then second dose of gel instilled, followed by 4 more hours of monitoring. Subsequent management of labor by attending physician and resident staff. 2) PGE_2 gel (0.25 mg) (n = 70) Protocol: Same as above, but with 0.25-mg dosage. 3) PGE_2 gel (0.5 mg) (n = 80) Protocol: Same as above, but with 0.5-mg dosage. Dates: July 1991 - May 1993 Location: Louisville, KY Setting: University hospital and community hospital (2 sites) Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 291 Dropouts: 62 (excluded due to protocol violations) Loss to follow-up: NA No. of subjects at end: 229 Inclusion criteria: Bishop score ≤ 4; induction required Exclusion criteria: Noncephalic presentation; previous vertical C-section; heavy vaginal bleeding; placenta previa; spontaneous labor; abnormal FHR tracing; maternal asthma or glaucoma; history of hypersensitivity to prostaglandin Age (mean, with 95% CI): 0.125 mg: 25.3 (24.1 to 26.6) 0.25 mg: 25.4 (23.9 to 27.0) 0.5 mg: 26.2 (24.6 to 27.8) Race: NR Gestational age at entry (mean, with 95% CI): 0.125 mg: 39.3 weeks (38.8 to 39.9); 29/79 (37%) “postdates” 0.25 mg: 38.5 weeks (37.3 to 39.6); 21/70 (30%) “postdates” 0.5 mg: 39.4 weeks (38.8 to 40.0); 21/80 (26%) Dating criteria: NR Parity: 0.125 mg: 61% nulliparous 0.25 mg: 60% nulliparous 0.5 mg: 69% nulliparous Bishop score: NR	1) FHR abnormality 2) C-sections 3) Change in Bishop score 4) Hyperstimulation 5) Time to a) active phase of labor, b) complete dilatation, and c) delivery (survival analysis)	1) FHR abnormality: 0.125 mg: 21.8% 0.25 mg: 29.9% 0.5 mg: 24.7% p = not significant 2) C-sections: 0.125 mg: 40.8% 0.25 mg: 40.8% 0.5 mg: 36.8% p = not significant 3) Change in Bishop score (mean): 0.125 mg: 2.08 0.25 mg: 1.43 0.5 mg: 1.94 p = not significant 4) Hyperstimulation: 0.125 mg: 7.7% 0.25 mg: 11.9% 0.5 mg: 10.4% p = not significant 5) Time to a) active phase of labor, b) complete dilatation, and c) delivery: Survival analysis showed no significant differences among the three groups for these outcomes.	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: - Sample size: - Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not stratified by parity.

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wing, Fassett, and Mishell, 2000	Design: RCT, randomization by computer-generated random number sequence and sealed envelopes Interventions: 1) Mifepristone (n = 97) Protocol: Mifepristone 200 mg given by mouth. Patient re-examined in 24 hours. If Bishop score ≥ 7, then labor induced using oxytocin. If Bishop score < 7, FHR tracing reactive, and no contractions, then patient given 25 µg misoprostol intravaginally. Misoprostol repeated every 4 hours until adequate labor established or 24 hours elapsed (maximum 6 doses or 150 µg). Oxytocin used if no active labor after maximum misoprostol dose and for failure to progress in active phase of labor. Oxytocin infused by pump at an initial dose of 1 mU/minute, with incremental increases every 30 minutes to a maximum dose of 22 mU/minute. 2) Placebo (n = 83) Protocol: Same as above, but with placebo rather than mifepristone Dates: Mar 1997 - Jan 1999 Location: Los Angeles, CA Setting: University hospital (2 sites) Type(s) of providers:	No. of subjects at start: 180 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 180 Inclusion criteria: Singleton pregnancy; vertex presentation; reactive NST; intact membranes; gestational age > 41 weeks; maternal age > 18 Exclusion criteria: Bishop score ≥ 7; cervix > 3 cm dilated; > 9 contractions per hour; estimated fetal weight < 2000 g or > 4500 g; evidence of cephalopelvic disproportion; placenta previa; unexplained vaginal bleeding; active genital herpes simplex; previous C-section or uterine surgery; chorioamnionitis; parity ≥ 6; pre-existing moderate or severe disease; contraindications to prostaglandins Age (mean ± SD): Mifepristone, 27.2 ± 5.9; placebo, 25.8 ± 5.4 Race: NR Gestational age at entry (mean ± SD): Mifepristone, 41.4 ± 0.4 weeks; placebo, 41.4 ± 0.4 weeks Dating criteria: 1) LMP confirmed by physical exam at 20 weeks or U/S no later than 26 weeks; *or* 2) U/S no later than 26 weeks Parity (mean ± SD): Mifepristone, 1.5 ± 1.4, 26% nulliparous;	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Abnormal FHR pattern 4) Birthweight 5) Admission to NICU 6) Length of stay in NICU 7) Plasma glucose, day 1 8) Plasma glucose, day 2 9) C-sections 10) Chorioamnionitis 11) Vaginal delivery in 24 hours 12) Vaginal delivery in 48 hours 13) Time to delivery 14) Time to active labor	1) Apgar score < 7 at 1 minute: Mifepristone: 15/97 (15.5%) Placebo: 7/83 (8.4%) p = 0.44 2) Apgar score < 7 at 5 minutes: Mifepristone: 2/97 (2%) Placebo: 0 p = 0.54 3) Abnormal FHR pattern: Mifepristone: 18/97 (18.6%) Placebo: 6/83 (7.2%) p = 0.34 4) Birthweight (mean ± SD) Mifepristone: 3676.57 ± 417.5 g Placebo: 3693.34 ± 501.8 p = 0.81 5) Admission to NICU: Mifepristone: 13/97 (13.4%) Placebo: 11/83 (13.3%) p = 0.98 6) Length of stay in NICU (mean ± SD): Mifepristone (n = 13): 5.5 ± 3.5 days Placebo (n = 11): 6.0 ± 4.1 days p = 0.78 7) Plasma glucose, day 1 (mean ± SD): Mifepristone: 64.8 ± 19.5 mg/dL Placebo: 66.5 ± 21.1 mg/dL p = 0.68 8) Plasma glucose, day 2 (mean ± SD): Mifepristone: 66.4 ± 19.5 mg/dL Placebo: 71.3 ± 23.1 mg/dL p =0.28 9) C-sections: Mifepristone: 9/97 (9.3%) Placebo: 18/83 (21.7%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Sample size based on proportion of patients delivering within 48 hours. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Maternal and family medicine Length of follow-up: None	placebo, 1.1 ± 1.2, 40% nulliparous Bishop score (median, with range): Mifepristone, 2 (0 to 6); placebo, 3 (0 to 6)		p = 0.02 10) Chorioamnionitis: Mifepristone: 15/97 (15.5%) Placebo: 18/83 (21.7%) p = 0.28 11) Vaginal delivery in 24 hours: Mifepristone: 12/88 (13.6%) Placebo: 7/65 (10.8%) p = 0.60 12) Vaginal delivery in 48 hours: *Overall:* Mifepristone: 77/88 (87.5%) Placebo: 46/65 (70.8%) p = 0.01 *Among nulliparas:* Mifepristone: 15/25 (60.0%) Placebo: 10/34 (29.4%) (no p-value reported) *Among multiparas:* Mifepristone: 62/72 (86.1%) Placebo: 36/49 (73.5%) (no p-value reported) 13) Time to delivery (mean ± SD): *Overall:* Mifepristone: 2209 ± 698 minutes Placebo: 2671 ± 884 minutes p < 0.001 *Among nulliparas:* Mifepristone (n = 25): 2426 ± 804 minutes Placebo (n = 34): 3169 ± 875 minutes p = 0.002 *Among multiparas:* Mifepristone (n = 72): 2129 ± 644 minutes Placebo (n = 49): 2326 ± 714 minutes	*(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				p = 0.16 14) Time to active labor (mean ± SD): Mifepristone: 1890 ± 668 minutes Placebo: 2303 ± 806 minutes p = 0.002	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wing, Jones, Rahall, et al., 1995	Design: RCT, randomization by table of random numbers and sealed envelopes Interventions: 1) Misoprostol (n = 68) Protocol: Misoprostol 50 µg applied intravaginally to posterior fornix. Dose repeated every 3 hours until adequate contraction pattern established (3 contractions in 10 minutes), Bishop score ≥ 8, dilation ≥ 3 cm, or SROM occurred. Maximum dose 300 µg or 6 doses 2) PGE_2 gel (n = 67) Protocol: PGE_2 gel (0.5 mg) applied intracervically every 6 hours as necessary to a maximum of 3 doses. In both groups, artifical rupture of the membranes generally performed when the cervix was 80% effaced and 3 cm dilated. If patient did not enter active labor after receiving maximum dose, had SROM without ensuing adequate contractile pattern, or had an arrest of dilatation, then IV oxytocin augmentation given (3 hours after last dose of misoprostol or ≥ 6 hours after last dose of PGE_2). Dates: Oct –Nov 1993 Location: Los Angeles, CA Setting: University hospital	No. of subjects at start: 135 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 135 Inclusion criteria: Medical or obstetric indication for induction; singleton gestation; cephalic presentation; intact membranes; Bishop score ≤ 4; reactive NST; < 4 spontaneous uterine contractions per hour Exclusion criteria: Abnormal FHR patterns; malpresentation; estimated fetal weight > 4500 g or other evidence of cephalopelvic disproportion; ruptured membranes; placenta previa or other unexplained vaginal bleeding; vasa previa; active herpes simplex infection; contraindication to prostaglandins; renal or hepatic dysfunction; suspected chorioamnionitis; previous C-section or history of uterine surgery; parity > 5 Age (mean ± SD): Misoprostol, 24.9 ± 6.9; PGE_2, 26.4 ± 6.9 Race: NR Gestational age at entry (mean ± SD): Misoprostol, 39.9 ± 2.3 weeks; PGE_2, 40.3 ± 1.9 weeks Dating criteria: NR Parity: Misoprostol / PGE_2 Nullip 52% / 48%	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Meconium aspiration syndrome 5) Admission to NICU 6) Neonatal resuscitation 7) Forceps delivery 8) Vacuum delivery 9) C-sections (overall and by indication) 10) Time to delivery 11) Vaginal delivery in 24 hours 12) Tachysystole 13) Hyperstimulation	1) Apgar score < 7 at 1 minute: Misoprostol: 9/68 (13.2%) PGE_2: 6/67 (9.0%) p = not significant 2) Apgar score < 7 at 5 minutes: Misoprostol: 1/68 (1.5%) PGE_2: 0/67 p = not significant 3) Birthweight (mean ± SD): Misoprostol: 3273.5 ± 522.4 g PGE_2: 3356.0 ± 523.0 g p = not significant 4) Meconium aspiration syndrome: Misoprostol: 3/68 (4.4%) PGE_2: 1/67 (1.5%) $p < 0.05$ 5) Admission to NICU: Misoprostol: 13/68 (9.6%) PGE_2: 11/67 (8.1%) p = not significant 6) Neonatal resuscitation: Misoprostol: 15/68 (22.1%) PGE_2: 5/67 (7.5%) $p < 0.05$ 7) Forceps delivery: Misoprostol: 2/68 (2.9%) PGE_2: 2/67 (3.0%) p = not significant 8) Vacuum delivery: Misoprostol: 5/68 (7.4%) PGE_2: 6/67 (8.9%) p = not significant 9) C-sections: *Overall:* Misoprostol: 10/68 (14.7%) PGE_2: 13/67 (19.4%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (10% of total study population). Results not stratified by parity. Sample size estimates based on proportion of patients achieving “adequate labor pattern” and proportion undelivered at 24 hours. Study underpowered to detect differences in categorical variables (e.g., tachysystole). *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: NR Length of follow-up: None	Primip 16% 19% Multip 32% 33% Bishop score (median, with range): Misoprostol, 2 (0-4); PGE2, 2 (0-4) Other: Indications for induction: Oligohydramnios: 58% Preeclampsia: 14% Postterm: 10% Macrosomia: 7% Abnormal antepartum testing: 3% Rh sensitization: 2% IUGR: 1% Diabetes mellitus: 1% Chronic hypertension: 1% Other: 3%		p = not significant *For arrest disorder:* Misoprostol: 6/68 (8.8%) PGE_2: 7/67 (10.4%) (no p-value reported) *For failed induction:* Misoprostol: 3/68 (4.4%) PGE_2: 5/67 (7.5%) (no p-value reported) *For fetal distress:* Misoprostol: 1/68 (1.5%) PGE_2: 1/67 (1.5%) (no p-value reported) 10) Time to delivery (mean ± SD): *Any delivery:* Misoprostol: 1100.9 ± 751.4 minutes PGE_2: 1592.6 ± 927.5 minutes $p < 0.001$ *Vaginal delivery:* Misoprostol: 903.3 ± 482.1 minutes PGE_2: 1410.9 ± 869.1 minutes $p < 0.001$ 11) Vaginal delivery in 24 hours: Misoprostol: 48/68 (70.6%) PGE_2: 32/67 (47.8%) $p < 0.01$ 12) Tachysystole: Misoprostol: 25/68 (36.7%) PGE_2: 8/67 (11.9%) $p < 0.001$ 13) Hyperstimulation Misoprostol: 5/68 (7.4%) PGE_2: 2/67 (3.0%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wing, Ortiz-Omphroy, and Paul, 1997	Design: RCT, randomization by computer-generated random numbers and sealed envelopes Interventions: 1) PGE_2 (n = 98) Protocol: 10-mg vaginal insert place in posterior fornix. Drug released at rate of 0.3 mg per hour. Insert removed if active labor (dilation ≥ 4 cm), SROM, Bishop score ≥ 8, cervical dilation ≥ 3 cm, uterine contraction abnormality (tachysystole, hypertonus, or hyperstimulation), abnormal FHR activity, or after 24 hours. 2) Misoprostol (n = 99) Protocol: 25 µg placed in posterior vaginal fornix every 4 hours until adequate contraction pattern established (3 contractions in 10 minutes), Bishop score ≥ 8, dilation ≥ 3 cm, SROM occurred, or 24 hours passed. Maximum dose 150 µg, or 6 doses. In both groups, AROM generally performed when cervix 80% effaced and 3 cm dilated, or when dilatation > 4 cm regardless of effacement. Patients who did not enter labor after maximum dose, or had SROM without adequate labor pattern, or arrest of dilatation received oxytocin augmentation. Dates: Oct 1995 - June 1996	No. of subjects at start: 200 Dropouts: 3 (excluded from analysis due to protocol violation) Loss to follow-up: NA No. of subjects at end: 197 Inclusion criteria: Medical or obstetric indication for induction; singleton gestation; cephalic presentation; intact membranes; Bishop score ≤ 4; reactive FHR pattern; < 8 spontaneous uterine contractions per hour Exclusion criteria: Abnormal FHR pattern; malpresentation; estimated fetal weight > 4500 g or other evidence of cephalopelvic disproportion; ruptured membranes; placenta previa or other unexplained vaginal bleeding; vasa previa; active herpes simplex infection; contraindications to prostaglandins; renal or hepatic dysfunction; suspected chorioamnionitis; previous C-section or other uterine surgery; parity > 5 Age: "Similar" in two groups Race: 97% Hispanic, equally distributed between the two groups Gestational age at entry (mean ± SD): PGE_2, 39.2 ± 2.3 weeks; misoprostol, 29.5 ± 2.4 weeks Dating criteria: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Neonatal resuscitation 5) Admission to NICU 6) C-sections 7) Cost of study medication (per dose) 8) Vaginal delivery within 12 and 24 hours	1) Apgar score < 7 at 1 minute: PGE_2: 11/98 (11.2%) Misoprostol: 9/99 (9.1%) p = 0.29 2) Apgar score < 7 at 5 minutes: PGE_2: 0/98 Misoprostol: 0/99 p = not significant 3) Birthweight (mean ± SD): PGE_2: 3264.6 ± 592.3 g Misoprostol: 3305.8 ± 549.3 g p = 0.61 4) Neonatal resuscitation: PGE_2: 25/98 (25.5%) Misoprostol: 29/99 (29.3%) p = 0.55 5) Admission to NICU: PGE_2: 27/98 (27.6%) Misoprostol: 30/99 (30.3%) p = 0.67 6) C-sections: PGE_2: 20/98 (20.4%) Misoprostol: 18/99 (18.2%) p = not significant 7) Cost of study medication (per dose): PGE_2: $135 per insert Misoprostol: $0.08 per 25-µg dose (no p-value reported) 8) Vaginal delivery: *Within 12 hours:* PGE_2: 19/98 (19.4%) Misoprostol: 20/99 (20.2%) p = not significant *Within 24 hours:* PGE_2: 45/98 (45.9%) Misoprostol: 51/99 (51.5%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (13% of total study population). Results not stratified by parity. Sample size based on proportion delivering within 12 hours. Tachysystole was less frequent with misoprostol than with PGE_2 (7.1% vs. 18.4%, p = 0.02). *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Location: Los Angeles, CA Setting: University hospital Type(s) of providers: Unspecified OB/GYN; senior residents Length of follow-up: None	Parity: "Similar" in the two groups Bishop score (median, with range): 2 (0-4) in both groups Other: Indications for induction: Oligohydramnios: 43% Preeclampsia: 25% Postterm: 13% Macrosomia: 6% Diabetes mellitus: 7.5% IUGR: 3.5% Chronic hypertension: 1% Other: 1%		p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wing and Paul, 1996	Design: RCT, randomization by computer-generated random numbers and sealed envelopes Interventions: 1) Misoprostol, 3-hour dosing regimen (n = 261) Protocol: Misoprostol 25 µg applied in posterior vaginal fornix every 3 hours until adequate contraction pattern established (3 contractions in 10 minutes), Bishop score ≥ 8, dilation ≥ 3 cm, SROM occurred, or 24 hours passed. Maximum dose 200 µg, or 8 doses. 2) Misoprostol, 6-hour dosing regimen (n = 259) Protocol: Same as above except dosing repeated every 6 hours to a maximum of 100 µg, or 4 doses. In both groups, AROM generally performed when cervix 80% effaced and 3 cm dilated, or when dilatation > 4 cm regardless of effacement. Patients who did not enter labor after maximum dose, or had SROM without adequate labor pattern, or arrest of dilatation received oxytocin augmentation. Dates: Oct 1994 - July 1995 Location: Los Angeles, CA Setting: University hospital	No. of subjects at start: 522 Dropouts: 2 (excluded from analysis due to protocol violation) Loss to follow-up: NA No. of subjects at end: 520 Inclusion criteria: Medical or obstetric indication for induction; singleton pregnancy; cephalic presentation; intact membranes; Bishop score ≤ 4; reactive FHR pattern; < 8 spontaneous uterine contractions per hour Exclusion criteria: Abnormal FHR pattern; malpresentation; estimated fetal weight > 4500 g or other evidence of cephalopelvic disproportion; ruptured membranes; placenta previa or other unexplained vaginal bleeding; vasa previa; active herpes simplex infection; contraindications to prostaglandins; renal or hepatic dysfunction; suspected chorioamnionitis; previous C-section or other uterine surgery; parity > 5 Age: "Similar" in two groups Race: 96% Hispanic, equally distributed between the two groups Gestational age at entry (mean ± SD): 3-hour dosing, 39.6 ± 2.3 weeks; 6-hour dosing, 39.5 ± 2.3 weeks	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Neonatal resuscitation 5) Admission to NICU 6) Instrumental vaginal delivery 7) C-sections 8) Maternal adverse events 9) Tachysystole 10) Time to vaginal delivery 11) Vaginal delivery within 24 hours	1) Apgar score < 7 at 1 minute: 3-hour dosing: 31/261 (13%) 6-hour dosing: 34/259 (13%) p = not significant 2) Apgar score < 7 at 5 minutes: 3-hour dosing: 3/261 (1.5%) 6-hour dosing: 4/259 (1.5%) p = not significant 3) Birthweight (mean ± SD): 3-hour dosing: 3273 ± 565.4 g 6-hour dosing: 3267.6 ± 554.1 p = not significant 4) Neonatal resuscitation: 3-hour dosing: 90/261 (34.5%) 6-hour dosing: 83/259 (32.0%) p = not significant 5) Admission to NICU: 3-hour dosing: 61/261 (23.4%) 6-hour dosing: 54/259 (20.8%) p = not significant 6) Instrumental vaginal delivery: 3-hour dosing: 16/261 (6%) 6-hour dosing: 17/259 (6.5%) p = not significant 7) C-sections: 3-hour dosing: 53/261 (20.3%) 6-hour dosing: 55/259 (21.3%) p = not significant 8) Maternal adverse events (treatment groups not specified): One maternal death from amniotic fluid embolism, 2 cesarean hysterectomies performed for vaginal hemorrhage resulting from uterine atony.	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: - Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (13% of total study population). Results not stratified by parity. Sample size estimates based on equivalence in tachysystole. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
	Type(s) of providers: MFM, senior resident Length of follow-up: None	Dating criteria: NR Parity: "Similar" in two groups Bishop score: Median, 2 in both groups (range NR) Other: Indications for induction: Oligohydramnios: 49% Preeclampsia: 17% Postterm: 13% Macrosomia: 5% Abnormal antepartum testing: 5% Diabetes mellitus: 5% IUGR: 2% Chronic hypertension: 0.6% Rh sensitization: 0.2% Other: 3%		9) Tachysystole: 3-hour dosing: 38/261 (14.6%) 6-hour dosing: 29/259 (11.2%) p = not significant 10) Time to vaginal delivery (mean ± SD): 3-hour dosing: 903.3 ± 482.1 minutes 6-hour dosing: 1410.9 ± 869.1 minutes $p < 0.05$ 11) Vaginal delivery within 24 hours: 3-hour dosing: 133/261 (63.9%) 6-hour dosing: 113/259 (55.4%) p = not significant	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wing, Rahall, Jones, et al., 1995	Design: RCT, randomization by table of random numbers and sealed envelopes Interventions: 1) Misoprostol (n = 138) Protocol: Misoprostol 25-µg tablet applied intravaginally to posterior fornix. Dose repeated every 3 hours until adequate contraction pattern established or until cervical ripening or SROM occurred. Maximum dose = 200 µg, or 8 doses. 2) PGE_2 (n = 137) Protocol: PGE_2 gel (0.5 mg) applied intracervically. Dose repeated every 6 hours as necessary for a maximum of 3 doses. Dates: Feb-June 1994 Location: Los Angeles, CA Setting: University hospital Type(s) of providers: Unspecified hospital Length of follow-up: None	No. of subjects at start: 276 Dropouts: 1 (excluded from analysis due to protocol violation) Loss to follow-up: NA No. of subjects at end: 275 Inclusion criteria: Medical or obstetric indication for induction; singleton gestation; cephalic presentation; intact membranes; Bishop score ≤ 4; reactive NST; < 4 spontaneous uterine contractions per hour Exclusion criteria: Abnormal FHR patterns; malpresentation; estimated fetal weight > 4500 g or other evidence of cephalopelvic disproportion; ruptured membranes; placenta previa or other unexplained vaginal bleeding; vasa previa; active herpes simplex infection; contraindication to prostaglandins; renal or hepatic dysfunction; suspected chorioamnionitis; previous C-section or history of uterine surgery; parity > 5 Age (mean ± SD): Misoprostol, 25.8 ± 6.2; PGE_2, 26.2 ± 6.5 Race: Both groups 95% Hispanic Gestational age at entry (mean ± SD): Misoprostol, 39.7 ± 2.3 weeks; PGE_2, 40.0 ± 2.4 weeks Dating criteria: NR	1) Apgar score < 7 at 1 minute 2) Apgar score < 7 at 5 minutes 3) Birthweight 4) Admission to NICU 5) Neonatal resuscitation 6) Forceps delivery 7) Vacuum delivery 8) C-sections (overall and by indication) 9) Cost of study medication per dose 10) Time to vaginal delivery 11) Vaginal delivery within 24 hours	1) Apgar score < 7 at 1 minute: Misoprostol: 15/138 (11%) PGE_2: 9/137 (7%) p = not significant 2) Apgar score < 7 at 5 minutes: Misoprostol: 0/138 PGE_2: 0/137 p = not significant 3) Birthweight (mean ± SD): Misoprostol: 3269.7 ± 587.5 g PGE_2: 3395.0 ± 607.4 g p = not significant 4) Admission to NICU: Misoprostol: 17/138 (12%) PGE_2: 23/137 (17%) p = not significant 5) Neonatal resuscitation: Misoprostol: 44/138 (32%) PGE_2: 43/137 (31%) p = not significant 6) Forceps delivery: Misoprostol: 4/138 (3%) PGE_2: 8/137 (6%) (no p-value reported) 7) Vacuum delivery: Misoprostol: 5/138 (4%) PGE_2: 11/237 (8%) (no p-value reported) 8) C-sections: *Overall:* Misoprostol: 28/138 (20%) PGE_2: 38/137 (28%) p = not significant *For abnormal FHR:* Misoprostol: 9/138 (6.5%) PGE_2: 4/137 (3%)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: + Sample size: + Statistical tests: + Gestational age: + Dating criteria: - Bishop score: + Results not reported separately for subgroup of patients induced for postterm pregnancy (16% of total study population). Results not stratified by parity. Sample size estimate based on proportion delivering within 24 hours. *(continued on next page)*

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				(no p-value reported)	
		Parity: Miso / PGE_2 Nullip 47% 47% Primip 18% 23% Multip 35% 30% Bishop score: NR Other: Indications for induction: Oligohydramnios: 40% Preeclampsia: 23% Postterm: 16% Macrosomia: 10% Diabetes mellitus: 5% Abnormal antepartum testing: 2% Chronic hypertension: 2% IUGR: 2% Other: 1%		*For failed induction:* Misoprostol: 4/138 (3%) PGE_2: 27/137 (20%) (no p-value reported) *For arrest disorder:* Misoprostol: 15/138 (11%) PGE_2: 7/137 (5%) (no p-value reported) 9) Cost of study medication per dose: Misoprostol: $0.08 PGE_2: $75.00 (no p-value reported) 10) Time to vaginal delivery (mean ± SD): Misoprostol: 1323.0 ± 844.4 minutes PGE_2: 1532.4 ± 706.5 minutes $p < 0.05$ 11) Vaginal delivery within 24 hours Misoprostol: 72/138 (65.5%) PGE_2: 41/137 (41.4%) $p < 0.01$	

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
Wiriya-sirivaj, Vutya-vanich, and Ruangsri, 1996	Design: RCT, randomization by table of random numbers Interventions: 1) Membrane stripping (n = 61) Protocol: Membranes stripped by digital separation from lower uterine segment as far as possible. Unfavorable cervixes stretched digitally as far as possible or until stripping could be accomplished. Repeated weekly until labor or 42 completed weeks' gestation. If no labor at 42 weeks, then labor induced with prostaglandin suppository or oxytocin drip. 2) Pelvic exam (control) (n = 59) Protocol: Pelvic exam to assess Bishop score only. Repeated weekly until labor or 42 completed weeks' gestation. If no labor at 42 weeks, then labor induced with prostaglandin suppository or oxytocin drip. Dates: Oct-Nov 1994 Location: Chiang Mai, Thailand Setting: University hospital Type(s) of providers: Unspecified OB/GYN Length of follow-up: None	No. of subjects at start: 120 Dropouts: 0 Loss to follow-up: NA No. of subjects at end: 120 Inclusion criteria: Gestational age 38 weeks; vertex presentation; no size-date discrepancy; no placenta previa or low-lying placenta; ability to attend follow-up visits; intention to deliver at study hospital Exclusion criteria: Previous C-section; known medical or surgical or obstetric complication that would preclude vaginal delivery; high risk Age (mean ± SD): Stripping, 25.6 ± 4.9; control, 26.2 ± 4.9 Race: NR Gestational age at entry: 38 weeks Dating criteria: LMP; early assessment of uterine size; or U/S before 28 weeks Parity: Both groups, 56% primigravidae Bishop score (mean ± SD): Stripping, 2.3 ± 1.5; control, 2.1 ± 1.7	1) Birthweight 2) Apgar scores at 1 minute 3) Apgar scores at 5 minutes 4) Neonatal jaundice 5) Post-partum fever 6) Post-partum hemorrhage 7) Forceps-assisted delivery 8) Vacuum extraction 9) C-section 10) Proportion of patients delivering within 7 days 11) Incidence of postterm pregnancies	1) Birthweight (mean ± SD): Stripping: 3123 ± 364.8 g Control: 3078 ± 320.5 g p = not significant 2) Apgar scores at 1 minute (mean ± SD): Stripping: 9.1 ± 1.1 Control: 9.1 ± 1.2 p = not significant 3) Apgar scores at 5 minutes (mean ± SD): Stripping: 9.9 ± 0.2 Control: 9.9 ± 0.1 p = not significant 4) Neonatal jaundice: Stripping: 4/61 (6.6%) Control: 4/59 (6.8%) p = not significant 5) Post-partum fever: Stripping: 1/61 (1.6%) Control: 0 p = not significant 6) Post-partum hemorrhage: Stripping: 2/61 (3.3%) Control: 2/59 (3.4%) p = not significant 7) Forceps-assisted delivery: Stripping: 2/61 (3.3%) Control: 5/59 (8.5%) (no p-value reported) 8) Vacuum extraction: Stripping: 8/61 (13.1%) Control: 6/59 (10.2%) (no p-value reported)	QUALITY SCORE: Randomized: + Method of randomization: + Similar to likely pt pop: + Interventions described: + Mode of delivery: - Sample size: + Statistical tests: + Gestational age: + Dating criteria: + Bishop score: + Results not stratified by parity.

(continued on next page)

Evidence Table 3: Studies relevant to Key Question 3 (continued)

Study	Design and Interventions	Patient Population	Outcomes Reported	Results	Quality Score/Notes
				9) C-section: Stripping: 6/61 (9.8%) Control: 3/59 (5.0%) (no p-value reported) 10) Proportion of patients delivering within 7 days: Stripping: 25/61 (41.0%) Control: 12/59 (20.3%) p = 0.014 11) Incidence of postterm pregnancies: Stripping: 1/61 (1.6%) Control: 3/59 (5.1%) p = not significant	