

**Evidence evaluation report —  
Prolonged pregnancy**

Consultation draft — October 2018



Prepared by Ampersand Health Science Writing for the  
Australian Government Department of Health

## Contents

Key messages.....	3
1 Process of the review .....	4
1.1 Research questions .....	4
1.1.1 Definition .....	4
1.1.2 Harms and benefits .....	4
1.1.3 Intervention .....	4
1.1.4 Additional considerations .....	4
1.2 Evidence evaluation .....	4
1.3 Outcomes for GRADE analysis.....	4
1.4 Grading of the certainty of the body of evidence.....	4
2 Definition .....	5
2.1 Q1: What is the definition of post-term pregnancy?.....	5
2.1.1 Background information.....	5
3 Harms and benefits.....	6
3.1 Q2: What are the maternal risks and/or benefits associated with post-term pregnancy? .....	6
3.2 Q3: What are the fetal risks and/or benefits associated with post-term pregnancy?.....	6
3.3 Advice to the Expert Working Group .....	6
4 Intervention .....	7
4.1 Q4: What options are available for women to prevent post-term pregnancy?.....	7
4.1.1 Background information.....	7
4.1.2 Labour induction versus expectant management.....	7
4.1.3 Advice to the Expert Working Group.....	7
4.1.4 Summary of findings tables.....	8
4.1.5 Evaluation of quality of randomised controlled trial .....	13
5 Additional considerations .....	14
5.1 Q6: What are the additional needs of Aboriginal and Torres Strait Islander women?.....	14
5.2 Q7: What are the additional considerations for migrant and refugee women? .....	14
References.....	15

## Key messages

- Compared with a policy of expectant management, a policy of labour induction for prolonged pregnancy was associated with fewer (all-cause) perinatal deaths (risk ratio (RR) 0.37, 95% confidence interval (CI) 0.17 to 0.80; 21 trials, 16,056 infants; moderate-quality evidence). There were fewer stillbirths in the induction group (RR 0.38, 95% CI 0.14 to 1.01; 21 trials, 16,056 infants; moderate-quality evidence). Rates of neonatal intensive care unit (NICU) admission were lower (RR 0.89, 95% CI 0.81 to 0.98; 14 trials, 14,627 infants; moderate-quality evidence) and fewer babies had Apgar scores less than seven at 5 minutes in the induction groups compared with expectant management (RR 0.70, 95% CI 0.50 to 0.98; 16 trials, 9,047 infants; moderate-quality evidence). There was no evidence of a difference for neonatal trauma (RR 1.00, 95%CI 0.65 to 1.54; 4 trials, 10,351 infants; low-quality evidence), for induction compared with expectant management.
- For women in the policy of induction arms of trials for prolonged pregnancy, there were fewer caesarean sections compared with expectant management (RR 0.89, 95% CI 0.83 to 0.94; 28 trials, 17,834 women; moderate-quality evidence); and a corresponding marginal increase in operative vaginal births with induction (RR 1.02, 95% CI 0.95 to 1.10; 19 trials, 15,337 women; moderate-quality evidence). There was no evidence of a difference between groups for perineal trauma (RR 1.21, 95%CI 0.95 to 1.55; 5 trials; 9,394 women; low-quality evidence), postpartum haemorrhage (RR 1.09 95%CI 0.92 to 1.30, 5 trials; 3,315 women; low-quality evidence), or length of maternal hospital stay (average mean difference [MD] -0.34 days, 95% CI -1.00 to 0.33; 5 trials; 1,146 women; very low quality evidence).
- In subgroup analyses, no clear differences between timing of induction (<41 weeks versus ≥41 weeks) or by state of cervix were seen for perinatal death, stillbirth, NICU admission, caesarean section, perineal trauma or instrumental vaginal birth.

# 1 Process of the review

## 1.1 Research questions

### 1.1.1 Definition

Q1 What is the definition of post-term pregnancy?

### 1.1.2 Harms and benefits

Q2 What are the maternal risks and/or benefits associated with post-term pregnancy?

Q3 What are the fetal risks and/or benefits associated with post-term pregnancy?

### 1.1.3 Intervention

Q4 What options are available for women to prevent post-term pregnancy?

### 1.1.4 Additional considerations

Q5 What are the additional needs of Aboriginal and Torres Strait Islander women?

Q6 What are the additional considerations for migrant and refugee women

## 1.2 Evidence evaluation

As a Cochrane review on induction of labour for improving birth outcomes for women at or beyond term has recently been published (Middleton et al 2018), it was agreed that the review be used as the basis for this evaluation of the evidence on prolonged pregnancy. Subsequent to the publication of the Cochrane review a large randomised controlled trial relevant to the topic was published (Grobman et al 2018). This review provides a summary of the findings of the Cochrane review updated with the Grobman trial results against the pre-specified research questions.

## 1.3 Outcomes for GRADE analysis

Outcome	Importance	Inclusion
Perinatal death	9	<input checked="" type="checkbox"/>
Stillbirth	9	<input checked="" type="checkbox"/>
Neonatal death	9	<input checked="" type="checkbox"/>
Admission to neonatal intensive care unit	9	<input checked="" type="checkbox"/>
Apgar score less than seven at 5 minutes	9	<input checked="" type="checkbox"/>
Neonatal trauma	9	<input checked="" type="checkbox"/>
Caesarean section	9	<input checked="" type="checkbox"/>

Key: 1 - 3 less important; 4 - 6 important but not critical for making a decision; 7 - 9 critical for making a decision

## 1.4 Grading of the certainty of the body of evidence

Assessing the certainty of a body of evidence using GRADE involves consideration of the following five domains: risk of bias, inconsistency, indirectness, imprecision and publication bias.

For an evidence base drawn from RCTs, the grading of the certainty of the body of evidence starts at 'high'. An evidence base drawn from observational studies starts as 'low'. In both cases, the evidence can be downgraded for each of the five domains depending on whether the limitation is considered serious (downgrade one level) or very serious (downgrade two levels). Evidence can also be upgraded when the effect is large (upgrade one level) or very large (upgrade two levels), where confounders would reduce the effect or where there is a dose-response effect.

Diagnostic accuracy studies start as high quality evidence. However, these studies are vulnerable to limitations and often lead to low quality evidence, mostly owing to indirectness of evidence associated with diagnostic accuracy being only a surrogate for patient outcomes.

## 2 Definition

### 2.1 Q1: What is the definition of post-term pregnancy?

#### 2.1.1 Background information

For the purposes of this review, 'term' is defined as 37<sup>0</sup> to 41<sup>6</sup> weeks gestation and 'post-term' as  $\geq 42^0$  weeks (AIHW 2018b).

In Australia in 2016, 91.7% of babies were born at 37 to 41 weeks and 0.6% of babies were born post-term (AIHW 2018b).

While the aetiology of post-term birth is not well elucidated (Mandrizzato et al 2010), risk factors such as obesity, nulliparity and maternal age greater than 30 years have been associated with an increased risk of post-term birth (Arrowsmith et al 2011; Caughey et al 2009; Heslehurst et al 2017; Roos et al 2010). Placental senescence may play a role in the pathophysiology of post-term birth (Mandrizzato et al 2010), and genetic/epigenetic factors have also been implicated (Schierding et al 2014).

## 3 Harms and benefits

### 3.1 Q2: What are the maternal risks and/or benefits associated with post-term pregnancy?

Potential risks for the mother associated with post-term pregnancy include prolonged labour, postpartum haemorrhage and perineal tears. It is likely that some of these outcomes result from intervening when the uterus and cervix are not ready for labour (Caughey & Musci 2004).

### 3.2 Q3: What are the fetal risks and/or benefits associated with post-term pregnancy?

In a study from the Norwegian Birth Registry (Heimstad et al 2008), the perinatal death rate was 0.018% at day 287 (41 weeks) and 0.51% at day 302+ (>43 weeks). These findings are important in that, even in a setting where early booking allows accurate assessment of gestational age and antenatal services are accessible for most women, post-term pregnancy constitutes a high-risk situation, especially for the baby. In another Norwegian study of nearly two million births from 1967 to 2006, the risk of post-term infant death was strongly associated with growth restriction (Morken et al 2014).

In Australia in 2016, <1% of babies were born post-term (AIHW 2018b). Among babies born post-term (AIHW 2018a):

- 52% were born vaginally, 31% were born by caesarean section and 17% were instrumental vaginal births (among babies born at 37-41 weeks, 54% were born vaginally, 33% by caesarean section and 13% were instrumental vaginal births)
- 5.9% weighed more than 4,500 g (compared with 1.3% of babies born at 37 to 41 weeks)
- 1.9% had Apgar scores lower than seven at 5 minutes (compared with 1.4% of babies born at 37 to 41 weeks).

The perinatal death rate was 2.2 per 1,000 births (compared to 1.5 per 1,000 births for babies born at term) (AIHW 2018a).

### 3.3 Advice to the Expert Working Group

Include the above information in the narrative.

## 4 Intervention

### 4.1 Q4: What options are available for women to prevent post-term pregnancy?

#### 4.1.1 Background information

Induction of labour is widely practised with the aim of preventing stillbirth and reducing perinatal morbidity (eg shoulder dystocia in large babies). It may also be practised to prevent outcomes such as caesarean section, prolonged labour, postpartum haemorrhage and traumatic birth (Caughey & Musci 2004).

Variation in rates of post-term births suggests that different policies and practices for managing post-term pregnancies (especially timing of inductions) have been used in Europe (Zeitlin et al 2007) and elsewhere. Concern about the high and increasing induction rate in many countries, and increasing caesarean rates despite increasing induction rates has been expressed (Keirse 2010).

Earlier versions of the Cochrane review included interventions involving monitoring, such as early pregnancy ultrasound, that may have an effect on the outcomes of pregnancies for women at or beyond term. This topic is now addressed in the Cochrane review *Ultrasound for fetal assessment in early pregnancy* (Whitworth 2015). In the updated review, the effects of timing of labour induction at or beyond term compared with expectant management (which may include various intensities and forms of monitoring) were evaluated.

When the cervix is favourable (usually a Bishop score of six or more), induction is often carried out using intravenous oxytocin and artificial rupture of amniotic membranes. If the cervix is not favourable, usually a prostaglandin gel or tablet is placed in the vagina or cervix to ripen the cervix and initiate the uterine contractions and labour. Many protocols are used with varying repeat intervals and transition to oxytocin and amniotomy depending on the onset of uterine contractions and progress of cervical dilatation.

Determining the threshold for induction of post-term pregnancies has been described as ‘the 41 week to 42 week dilemma’ (Kortekaas et al 2014), with many hospitals now adopting a policy of induction from 41<sup>0</sup> weeks rather than a policy of waiting to induce at 42 completed weeks if spontaneous labour has not occurred. This 41-week policy may substantially increase numbers of inductions - in the Netherlands this policy would mean that 18% of all pregnant women would be induced compared with 1.5% if a 42-week policy was adopted (Kortekaas et al 2014). If numbers of inductions of labour increase, busy obstetrical centres will need to find new ways to accommodate larger numbers of women with longer lengths of stay in the labour and maternity unit (Greene 2018). It is important to assess whether improved outcomes such as reduced perinatal death and fewer caesarean sections can be achieved with earlier inductions and to determine the optimal gestational threshold for induction.

#### 4.1.2 Labour induction versus expectant management

##### *Discussion*

A recent Cochrane review assessed the effects of a policy of labour induction at or beyond term compared with a policy of awaiting spontaneous labour indefinitely (until a later gestational age or until a maternal or fetal indication for induction of labour is identified) on pregnancy outcomes for the infant and the mother (Middleton et al 2018).

The updated review comprised 30 trials that randomised 12,479 women and their babies and three ongoing studies. Since a risk factor at this stage of pregnancy would normally require an intervention, only trials including women at low risk for complications were eligible, with the trialists’ definition of ‘low risk’ accepted. Trials of induction of labour in women with prelabour rupture of membranes at or beyond term were not considered.

A subsequent multicentre trial randomly assigned low-risk nulliparous women who were at 38<sup>0</sup> to 38<sup>6</sup> weeks gestation to labour induction at 39<sup>0</sup> weeks to 39<sup>4</sup> weeks (n=3,062) or to expectant management (n=3,044) (Grobman et al 2018). The findings of this study have been incorporated into the Cochrane analysis for the following outcomes – perinatal mortality, stillbirth, admission to neonatal intensive care unit, neonatal trauma, neonatal convulsions, meconium aspiration syndrome, caesarean section, instrumental vaginal birth, perineal trauma.

#### 4.1.3 Advice to the Expert Working Group

Based on moderate quality evidence of fewer perinatal deaths, stillbirths, admission to neonatal intensive care unit, Apgar scores <7 and caesarean sections with induction for prolonged pregnancy compared to expectant management, suggest include a recommendation that options in prolonged pregnancy, including induction of labour, be discussed with women.

#### 4.1.4 Summary of findings tables

##### Summary of findings: Labour induction versus expectant management (infant/child outcomes)

**Patient or population:** Pregnant women at  $\geq 39$  weeks gestation

**Setting:** Norway, China, Thailand, USA, Austria, Turkey, Canada, UK, India, Tunisia, Finland, Spain, Sweden and the Netherlands

**Intervention:** Labour induction

**Comparison:** Expectant management

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with expectant management	Risk with labour induction				
<b>Perinatal death</b>	2 per 1,000	1 per 1,000 (0 to 2)	RR 0.37 (0.17 to 0.80)	16,056 (21 RCTs)	⊕⊕⊕○ MODERATE <sup>1</sup>	
Stillbirth	1 per 1,000	1 per 1,000 (0 to 1)	RR 0.38 (0.14 to 1.01)	16,056 (21 RCTs)	⊕⊕⊕○ MODERATE <sup>1</sup>	
<b>Admission to neonatal intensive care unit</b>	104 per 1,000	92 per 1,000 (684 to 102)	RR 0.89 (0.81 to 0.98)	14,627 (14 RCTs)	⊕⊕⊕○ MODERATE <sup>1</sup>	
Neonatal encephalopathy	—	—	—	(0 RCTs)	—	No RCTs reported data for this outcome
<b>Apgar scores less than 7 at 5 minutes</b>	17 per 1,000	12 per 1,000 (7 to 17)	RR 0.70 (0.50 to 0.98)	9,047 (16 RCTs)	⊕⊕⊕○ MODERATE <sup>1</sup>	
Neonatal trauma	8 per 1,000	8 per 1,000 (5 to 12)	RR 1.00 (0.65 to 1.54)	10,351 (4 RCTs)	⊕⊕○○ LOW <sup>1,2</sup>	
Neurodevelopment at childhood follow-up	—	—	—	(0 RCTs)	—	No RCTs reported data for this outcome

\*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; RR: Risk ratio

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Studies contributing data had some design limitations. (-1)

2. Wide confidence intervals crossing the line of no effect. (-1)



Summary of findings 2: Labour induction versus expectant management (maternal outcomes)

**Patient or population:** Pregnant women at  $\geq 39$  weeks gestation

**Setting:** Norway, China, Thailand, the USA, Austria, Turkey, Canada, UK, India, Tunisia, Finland, Spain, Sweden and the Netherlands

**Intervention:** Labour induction

**Comparison:** Expectant management

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with expectant management	Risk with labour induction				
Caesarean section	197 per 1,000	176 per 1,000 (164 to 185)	RR 0.89 (0.83 to 0.94)	17,834 (28 RCTs)	⊕⊕⊕○ MODERATE <sup>1</sup>	
Instrumental vaginal birth (forceps or ventouse)	149 per 1,000	152 per 1,000 (142 to 164)	RR 1.02 (0.95 to 1.10)	15,337 (19 RCTs)	⊕⊕○○ LOW <sup>1,2</sup>	
Perineal trauma	24 per 1,000	29 per 1,000 (22 to 37)	RR 1.21 (0.95 to 1.55)	9,394 (5 RCTs)	⊕⊕○○ LOW <sup>1,2</sup>	
Postpartum haemorrhage	122 per 1,000	133 per 1,000 (112 to 159)	RR 1.09 (0.92 to 1.30)	3,315 (5 RCTs)	⊕⊕○○ LOW <sup>1,2</sup>	
Breastfeeding at discharge	—	—	—	(0 RCTs)	—	No RCTs reported data for this outcome
Postnatal depression	—	—	—	(0 RCTs)	—	No RCTs reported data for this outcome
Length of maternal hospital stay (days)			Average MD 0.34 days shorter for women who were induced (1 day shorter to 0.33 days longer)	1,146 (5 RCTs)	⊕⊕○○ LOW <sup>1,2,3</sup>	

\*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval;

**GRADE Working Group grades of evidence**

**High certainty:** We are very confident that the true effect lies close to that of the estimate of the effect

**Moderate certainty:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

**Low certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

**Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

1. Studies contributing data had some design limitations. (-1)
2. Wide confidence intervals crossing the line of no effect. (-1)
3. Statistical heterogeneity ( $I^2 = 95\%$ ). Variation in size and direction of effect. (-2)

## Primary outcome

### Perinatal death

There were fewer perinatal deaths in the labour induction groups (a relative risk reduction of 63%) than in the expectant management groups: four perinatal deaths occurred in the induction group compared with nineteen in the expectant group (RR 0.37, 95%CI 0.17 to 0.80; 21 trials; 16,056 infants; moderate-quality evidence) (Augensen et al 1987; Bergsjø et al 1989; Chanrachakul & Herabutya 2003; Cole et al 1975; Dyson et al 1987; Egarter et al 1989; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Henry 1969; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Martin et al 1978; Martin et al 1989; NICHD 1994; Sahraoui et al 2005; Sande et al 1983; Suikkari et al 1983; Walker et al 2016).

**Subgroup interaction:** There were no significant differences between the timing of induction subgroups for perinatal deaths ( $\text{Chi}^2=0.16$ ,  $P=0.69$ ,  $I^2=0\%$ ) or for subgroups according to state of cervix ( $\text{Chi}^2=0.08$ ,  $P=0.96$ ,  $I^2=0\%$ ).

**Sensitivity analysis:** On sensitivity analysis, the point estimate of 56% relative risk reduction was similar to the overall analysis (RR 0.44, 95% CI 0.14 to 1.44; six trials, 12,789 infants).

### Secondary outcomes for the infant/child

#### Stillbirth

Thirteen of the 19 perinatal deaths were stillbirths. Two stillbirths occurred in the induction group and eleven stillbirths occurred in the expectant management groups (RR 0.38, 95% CI 0.14 to 1.01; 21 trials; 16,056 infants; moderate-quality evidence) (Augensen et al 1987; Bergsjø et al 1989; Chanrachakul & Herabutya 2003; Cole et al 1975; Dyson et al 1987; Egarter et al 1989; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Henry 1969; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Martin et al 1978; Martin et al 1989; NICHD 1994; Sahraoui et al 2005; Sande et al 1983; Suikkari et al 1983; Walker et al 2016).

**Subgroup interaction:** There were no significant differences between the timing of induction subgroups ( $\text{Chi}^2=0.08$ ,  $P=0.78$ ,  $I^2=0\%$ ) or for subgroups according to state of cervix ( $\text{Chi}^2=0.01$ ,  $P=0.95$ ,  $I^2=0\%$ ) for the outcome of stillbirth.

**Sensitivity analysis:** There were two stillbirths in the induction group and five stillbirths in the expectant management group. On sensitivity analysis, the point estimate of a relative risk reduction of 54% was similar to the overall analysis (RR 0.46, 95% CI 0.10 to 2.03; six trials, 12,789 infants).

#### Neonatal death

There were 10 live birth deaths (all occurring before 7 days of life). Two were in the induction group and eight were in the expectant management groups (RR 0.39, 95%CI 0.12 to 1.25; 20 trials; 15,872 infants) (Augensen et al 1987; Bergsjø et al 1989; Chanrachakul & Herabutya 2003; Cole et al 1975; Dyson et al 1987; Egarter et al 1989; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Henry 1969; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Martin et al 1989; NICHD 1994; Sahraoui et al 2005; Sande et al 1983; Suikkari et al 1983; Walker et al 2016).

#### Birth asphyxia

Rates of birth asphyxia were not clearly different between the induction and expectant management groups (RR 1.66, 95% CI 0.61 to 4.55; four trials; 1,456 infants) (Chanrachakul & Herabutya 2003; Heimstad et al 2007; Tylleskar et al 1979; Walker et al 2016).

#### Admission to neonatal intensive care unit (NICU)

Rates of NICU admissions were lower when labour induction was compared with expectant management (RR 0.89, 95% CI 0.81 to 0.98; 14 trials; 14,627 infants; moderate-quality evidence) (Augensen et al 1987; Brane et al 2014; Chanrachakul & Herabutya 2003; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Herabutya et al 1992; Kortekaas et al 2014; Miller et al 2015; Nielsen et al 2005; Ocon et al 1997; Roach & Rogers 1997; Walker et al 2016).

**Subgroup interaction:** There were no significant differences between the timing of induction subgroups ( $\text{Chi}^2=0.51$ ,  $P=0.78$ ,  $I^2=0\%$ ) or for subgroups according to state of cervix ( $\text{Chi}^2=0.86$ ,  $P=0.65$ ,  $I^2=0\%$ ).

**Sensitivity analysis:** On sensitivity analysis, results were similar to the overall analysis (RR 0.89, 95% CI 0.80 to 0.98; seven trials, 12,818 infants).

#### Neonatal convulsions

There was no significant difference in instance of neonatal convulsions when labour induction was compared with expectant management (RR 1.38, 95%CI 0.62 to 3.04; four trials, 10,467 infants) (Grobman et al 2018; Hannah et al 1992; NICHD 1994; Walker et al 2016).

#### Use of anticonvulsants

No clear differences between induction and expectant groups were evident for use of anticonvulsants in a single trial (RR 0.34, 95% CI 0.01 to 8.17; 1 trial; 349 infants) (NICHD 1994).

### Meconium aspiration syndrome

There was a 24% relative reduction in the risk of meconium aspiration syndrome in the induction groups compared with the expectant management groups (RR 0.76 95% CI 0.62 to 0.93; 12 trials; 13,877 infants)(Bergsjö et al 1989; Dyson et al 1987; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; James et al 2001; Kortekaas et al 2014; NICHD 1994; Roach & Rogers 1997; Sahraoui et al 2005; Witter & Weitz 1987).

### Apgar score less than seven at five minutes

Fewer babies had Apgar scores less than seven at five minutes in the induction groups compared with the expectant management groups (RR 0.70, 95% CI 0.50 to 0.98; 16 trials; 9047 infants; moderate-quality evidence) (Brane et al 2014; Chanrachakul & Herabutya 2003; Dyson et al 1987; Gelisen et al 2005; Hannah et al 1992; Heimstad et al 2007; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Miller et al 2015; NICHD 1994; Nielsen et al 2005; Ocon et al 1997; Roach & Rogers 1997; Walker et al 2016; Witter & Weitz 1987).

### Birthweight (g)

On average, infants born to mothers in the induction group had lower birthweights than those born to mothers in the expectant management group (mean difference (MD) -69.43 g, 95% CI -96.83 to -42.02; 14 trials; 3,799 infants) (Augensen et al 1987; Brane et al 2014; Chanrachakul & Herabutya 2003; Cole et al 1975; Dyson et al 1987; Heimstad et al 2007; Herabutya et al 1992; Miller et al 2015; NICHD 1994; Nielsen et al 2005; Roach & Rogers 1997; Tylleskar et al 1979; Walker et al 2016; Witter & Weitz 1987).

### Birthweight greater than 4,000 g

There was a 28% relative reduction in the rate of macrosomia (greater than 4,000 g) in the labour induction groups (average RR 0.72, 95% CI 0.54 to 0.96; eight trials; 5,593 infants;  $\text{Tau}^2 = 0.09$ ;  $\text{Chi}^2 = 20.84$ ,  $P = 0.004$ ;  $I^2 = 66\%$ ) (Chanrachakul & Herabutya 2003; Dyson et al 1987; Gelisen et al 2005; Hannah et al 1992; Heimstad et al 2007; James et al 2001; NICHD 1994; Ocon et al 1997). (Hannah 1992 used a cutoff-of 4,500 g rather than 4,000 g for this outcome.)

### Neonatal trauma

On meta-analysis of data from three trials no clear difference in rates of birth trauma in newborns was seen between labour induction and expectant management (RR 1.00, 95% CI 0.65 to 1.54; 10,351 infants; low-quality evidence) (Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; NICHD 1994).

### Unreported outcomes

No trials reported on neonatal encephalopathy, pneumonia, or neurodevelopment at childhood follow-up (although Bergsjö 1989 reported no signs of neurological impairment in children at two years of age).

### Secondary outcomes for the mother

#### Caesarean section

There were fewer caesarean sections (a relative reduction of 11%) in the induction groups compared with the expectant management groups (RR 0.89, 95% CI 0.83 to 0.94; 28 trials; 17,834 women; moderate-quality evidence) (Augensen et al 1987; Bergsjö et al 1989; Brane et al 2014; Breart et al 1982; Chakravarti & Goenka 2000; Chanrachakul & Herabutya 2003; Cole et al 1975; Dyson et al 1987; Egarter et al 1989; Gelisen et al 2005; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Henry 1969; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Martin et al 1978; Martin et al 1989; Miller et al 2015; NICHD 1994; Nielsen et al 2005; Ocon et al 1997; Roach & Rogers 1997; Sahraoui et al 2005; Tylleskar et al 1979; Walker et al 2016; Witter & Weitz 1987).

**Subgroup interaction tests:** There were no clear differences according to timing of induction ( $\text{Chi}^2=1.87$ ,  $P=0.39$ ,  $I^2=0\%$ ) or by state of cervix ( $\text{Chi}^2 = 1.06$ ,  $P = 0.59$ ,  $I^2 = 0\%$ ).

**Sensitivity analysis:** On sensitivity analysis, results were very similar to the overall analysis (RR 0.89, 95% CI 0.83 to 0.95; eight trials, 13,176 women).

#### Instrumental vaginal birth (forceps or ventouse)

On meta-analysis of data from the 19 trials that reported this outcome, the rate of operative vaginal birth was higher in the policy of labour induction groups compared with expectant management (RR 1.02, 95%CI 0.95 to 1.10; 15,337 women; moderate-quality evidence) (Augensen et al 1987; Bergsjö et al 1989; Brane et al 2014; Breart et al 1982; Cole et al 1975; Egarter et al 1989; Grobman et al 2018; Hannah et al 1992; Heimstad et al 2007; Henry 1969; Herabutya et al 1992; James et al 2001; Kortekaas et al 2014; Martin et al 1978; Martin et al 1989; Nielsen et al 2005; Ocon et al 1997; Tylleskar et al 1979; Walker et al 2016).

**Subgroup interaction:** No clear differences were seen in subgroup analyses for timing of induction ( $\text{Chi}^2=2.19$ ,  $P=0.33$ ,  $I^2=8.7\%$ ) or state of cervix ( $\text{Chi}^2= 0.45$ ,  $P = 0.80$ ,  $I^2 = 0\%$ ) for this outcome.

**Sensitivity analysis:** On sensitivity analysis, results were very similar to the overall analysis (RR 0.98, 95% CI 0.91 to 1.06; six trials, 12,666 women).

#### Analgesia used

In nine trials with 3,724 women, there was substantial variation in type of analgesia/anaesthesia used and so data were not pooled. In general, there were few differences seen in need for analgesia between the induction and expectant management groups.

#### Perineal trauma

On meta-analysis of data from five trials, no clear differences in perineal trauma were seen between induction and expectant management (RR 1.21, 95%CI 0.95 to 1.55; 9,394 women; low-quality evidence) (Brane et al 2014; Grobman et al 2018; Heimstad et al 2007; Kortekaas et al 2014; Walker et al 2016).

**Subgroup interaction:** No clear differences were seen in subgroup analyses for timing of induction ( $\text{Chi}^2=2.95$ ,  $P=0.23$ ,  $I^2=32.3\%$ ) or state of cervix (tests for subgroup differences: not applicable) for this outcome.

**Sensitivity analysis:** On sensitivity analysis, results were similar to the overall analysis (RR 1.26, 95% CI 0.98 to 1.62; four trials, 9,303 women).

#### Prolonged labour

The outcome of prolonged labour was reported in several different ways by three trials with 869 women (Chanrachakul & Herabutya 2003; Heimstad et al 2007; Henry 1969), with none of the four comparisons showing clear differences between the induction and expectant management groups.

#### Postpartum haemorrhage

No clear difference in rates of postpartum haemorrhage was seen between induction and expectant management groups (RR 1.09 95% CI 0.92 to 1.30; five trials, 3,315 women; low-quality evidence) (Brane et al 2014; Chanrachakul & Herabutya 2003; Heimstad et al 2007; Kortekaas et al 2014; Walker et al 2016).

#### Other measures of satisfaction with the approach

In one trial of 496 women, more women in the induction group said that they would choose the same arm in a future trial compared with women in the expectant management group (RR 1.93, 95% CI 1.62 to 2.30) (Heimstad et al 2007). However, in another trial of 184 women, similar numbers of women indicated that they preferred the group they had been allocated to (RR 0.90, 95% CI 0.72 to 1.13) (Martin et al 1978). Due to the high heterogeneity ( $I^2=96\%$ ) the results of these two trials were not pooled.

#### Secondary outcomes relating to health service use

##### Length of maternal hospital stay (days)

No clear overall differences between induction and expectant management were observed for duration of maternal hospital stay (average MD -0.34 days 95% CI -1.00 to 0.33; five trials; 1,146 women; very low-quality evidence) (Augensen et al 1987; Dyson et al 1987; James et al 2001; Miller et al 2015; Witter & Weitz 1987). There was, however, substantial heterogeneity ( $\text{Tau}^2=0.49$ ;  $\text{Chi}^2=77.02$ ,  $P<0.00001$ ;  $I^2=95\%$ ) between the trials for this outcome.

##### Length of neonatal hospital stay (days)

In one trial of 302 babies, there was a slightly shorter mean hospital stay for the induction group compared with the expectant management group (MD -0.30 day, 95% CI -0.61 to 0.01) (Dyson et al 1987).

##### Length of labour (hours)

Overall, length of labour was slightly shorter for women undergoing induction compared with expectant management (average MD -1.01 hours, 95% CI -1.72 to -0.31; nine trials; 1,980 women;  $\text{Tau}^2 = 0.97$ ;  $\text{Chi}^2=34.04$ ,  $P<0.0002$ ;  $I^2=71\%$ ) (Augensen et al 1987; Brane et al 2014; Cole et al 1975; Dyson et al 1987; Egarter et al 1989; Herabutya et al 1992; Martin et al 1978; Nielsen et al 2005; Tylleskar et al 1979).

#### 4.1.5 Evaluation of quality of randomised controlled trial

Study limitation	Judgement	Support for judgement
<b>(Grobman et al 2018)</b>		
Random sequence generation	Low risk	The randomisation sequence, prepared by an independent data coordinating centre, used the simple urn method, with stratification according to clinical site.
Allocation concealment	High risk	Unmasked trial.
Blinding	High risk	Blinding was not feasible.
Incomplete outcome data	Low risk	Loss to follow-up reported.
Selective reporting	Low risk	Pre-specified outcomes reported.
Other limitations	Low risk	The two groups were similar at baseline, except that fewer women in the induction group than in the expectant-management group had had a previous pregnancy loss (22.8% vs. 25.6%, P = 0.01)

## 5 Additional considerations

### 5.1 Q6: What are the additional needs of Aboriginal and Torres Strait Islander women?

No studies were identified to answer this question.

### 5.2 Q7: What are the additional considerations for migrant and refugee women?

No studies were identified to answer this question.

## References

- AIHW (2018a) [National Perinatal Data Collection](#). Accessed: 14 August 2018.
- AIHW (2018b) [Australia's Mothers and Babies 2016 – In Brief](#). Canberra: Australian Institute of Health and Welfare.
- Arrowsmith S, Wray S, Quenby S (2011) Maternal obesity and labour complications following induction of labour in prolonged pregnancy. *BJOG* 118(5): 578-88.
- Augensen K, Bergsjø P, Eikeland T et al (1987) Randomised comparison of early versus late induction of labour in post-term pregnancy. *Br Med J (Clin Res Ed)* 294(6581): 1192-5.
- Bergsjø P, Huang GD, Yu SQ et al (1989) Comparison of induced versus non-induced labor in post-term pregnancy. A randomized prospective study. *Acta Obstet Gynecol Scand* 68(8): 683-7.
- Brane E, Olsson A, Andolf E (2014) A randomized controlled trial on early induction compared to expectant management of nulliparous women with prolonged latent phases. *Acta Obstet Gynecol Scand* 93(10): 1042-9.
- Breart G, Goujard J, Maillard F et al (1982) Comparison of two obstetrical policies with regard to artificial induction of labour at term. A randomised trial. *Journal de Gynecologie, Obstetrique et Biologie de la Reproduction* 11: 107-12.
- Caughey AB & Musci TJ (2004) Complications of term pregnancies beyond 37 weeks of gestation. *Obstet Gynecol* 103(1): 57-62.
- Caughey AB, Sundaram V, Kaimal AJ et al (2009) Systematic review: elective induction of labor versus expectant management of pregnancy. *Ann Intern Med* 151(4): 252-63, W53-63.
- Chakravarti S & Goenka B (2000) Conservative policy of induction of labor in uncomplicated postdated pregnancies. XVI FIGO World Congress of Obstetrics & Gynecology; 2000 Sept 3-8, Washington DC, USA,
- Chanrachakul B & Herabutya Y (2003) Postterm with favorable cervix: is induction necessary? *Eur J Obstet Gynecol Reprod Biol* 106(2): 154-7.
- Cole RA, Howie PW, Macnaughton MC (1975) Elective induction of labour. A randomised prospective trial. *Lancet* 1(7910): 767-70.
- Dyson DC, Miller PD, Armstrong MA (1987) Management of prolonged pregnancy: induction of labor versus antepartum fetal testing. *Am J Obstet Gynecol* 156(4): 928-34.
- Egarter C, Kofler E, Fitz R et al (1989) Is induction of labor indicated in prolonged pregnancy? Results of a prospective randomised trial. *Gynecol Obstet Invest* 27(1): 6-9.
- Gelisen O, Caliskan E, Dilbaz S et al (2005) Induction of labor with three different techniques at 41 weeks of gestation or spontaneous follow-up until 42 weeks in women with definitely unfavorable cervical scores. *Eur J Obstet Gynecol Reprod Biol* 120(2): 164-9.
- Greene MF (2018) Choices in Managing Full-Term Pregnancy. *N Engl J Med* 379(6): 580-81.
- Grobman WA, Rice MM, Reddy UM et al (2018) Labor Induction versus Expectant Management in Low-Risk Nulliparous Women. *N Engl J Med* 379(6): 513-23.
- Hannah ME, Hannah WJ, Hellmann J et al (1992) Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. A randomized controlled trial. The Canadian Multicenter Post-term Pregnancy Trial Group. *N Engl J Med* 326(24): 1587-92.
- Heimstad R, Skogvoll E, Mattsson LA et al (2007) Induction of labor or serial antenatal fetal monitoring in postterm pregnancy: a randomized controlled trial. *Obstet Gynecol* 109(3): 609-17.
- Heimstad R, Romundstad PR, Salvesen KA (2008) Induction of labour for post-term pregnancy and risk estimates for intrauterine and perinatal death. *Acta Obstet Gynecol Scand* 87(2): 247-9.
- Henry GR (1969) A controlled trial of surgical induction of labour and amnioscopy in the management of prolonged pregnancy. *J Obstet Gynaecol Br Commonw* 76(9): 795-8.
- Herabutya Y, Prasertsawat PO, Tongyai T et al (1992) Prolonged pregnancy: the management dilemma. *Int J Gynaecol Obstet* 37(4): 253-8.
- Heslehurst N, Vieira R, Hayes L et al (2017) Maternal body mass index and post-term birth: a systematic review and meta-analysis. *Obes Rev* 18(3): 293-308.
- James C, George SS, Gaunekar N et al (2001) Management of prolonged pregnancy: a randomized trial of induction of labour and antepartum foetal monitoring. *Natl Med J India* 14(5): 270-3.
- Keirse MJ (2010) Elective induction, selective deduction, and cesarean section. *Birth* 37(3): 252-6.
- Kortekaas JC, Bruinsma A, Keulen JK et al (2014) Effects of induction of labour versus expectant management in women with impending post-term pregnancies: the 41 week - 42 week dilemma. *BMC Pregnancy Childbirth* 14: 350.
- Mandruzzato G, Alfrevic Z, Chervenak F et al (2010) Guidelines for the management of postterm pregnancy. *J Perinat Med* 38(2): 111-9.
- Martin DH, Thompson W, Pinkerton JH et al (1978) A randomized controlled trial of selective planned delivery. *Br J Obstet Gynaecol* 85(2): 109-13.
- Martin JN, Jr., Sessums JK, Howard P et al (1989) Alternative approaches to the management of gravidas with prolonged-postterm-postdate pregnancies. *J Miss State Med Assoc* 30(4): 105-11.
- Middleton P, Shepherd E, Crowther CA (2018) [Induction of labour for improving birth outcomes for women at or beyond term](#). *Cochrane Database Syst Rev*(5): CD004945.
- Miller NR, Cypher RL, Foglia LM et al (2015) Elective Induction of Labor Compared With Expectant Management of Nulliparous Women at 39 Weeks of Gestation: A Randomized Controlled Trial. *Obstet Gynecol* 126(6): 1258-64.
- Morken NH, Klungsoyr K, Skjaerven R (2014) Perinatal mortality by gestational week and size at birth in singleton pregnancies at and beyond term: a nationwide population-based cohort study. *BMC Pregnancy Childbirth* 14: 172.

- NICHHD (1994) A clinical trial of induction of labor versus expectant management in postterm pregnancy. The National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units. *Am J Obstet Gynecol* 170(3): 716-23.
- Nielsen PE, Howard BC, Hill CC et al (2005) Comparison of elective induction of labor with favorable Bishop scores versus expectant management: a randomized clinical trial. *J Matern Fetal Neonatal Med* 18(1): 59-64.
- Ocon L, Hurtado R, Coteron JJ et al (1997) Prolonged pregnancy: procedure guidelines [Gestacion prolongada: pautas de actuacion]. *Progresos de Obstetricia y Ginecologia* 40: 101-06.
- Roach VJ & Rogers MS (1997) Pregnancy outcome beyond 41 weeks gestation. *Int J Gynaecol Obstet* 59(1): 19-24.
- Roos N, Sahlin L, Ekman-Ordeberg G et al (2010) Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. *Acta Obstet Gynecol Scand* 89(8): 1003-10.
- Sahraoui W, Hajji S, Bibi M et al (2005) [Management of pregnancies beyond forty-one week's gestation with an unfavorable cervix]. *J Gynecol Obstet Biol Reprod (Paris)* 34(5): 454-62.
- Sande HA, Tuveng J, Fonstelien T (1983) A prospective randomized study of induction of labor. *Int J Gynaecol Obstet* 21(4): 333-6.
- Schierding W, O'Sullivan JM, Derraik JG et al (2014) Genes and post-term birth: late for delivery. *BMC Res Notes* 7: 720.
- Suikkari AM, Jalkanen M, Heiskala H et al (1983) Prolonged pregnancy: induction or observation. *Acta Obstet Gynecol Scand* 116(Suppl): 58.
- Tylleskar J, Finnstrom O, Leijon I et al (1979) Spontaneous labor and elective induction--a prospective randomized study. I. Effects on mother and fetus. *Acta Obstet Gynecol Scand* 58(6): 513-8.
- Walker KF, Bugg GJ, Macpherson M et al (2016) Randomized Trial of Labor Induction in Women 35 Years of Age or Older. *N Engl J Med* 374(9): 813-22.
- Witter FR & Weitz CM (1987) A randomized trial of induction at 42 weeks gestation versus expectant management for postdates pregnancies. *Am J Perinatol* 4(3): 206-11.
- Zeitlin J, Blondel B, Alexander S et al (2007) Variation in rates of postterm birth in Europe: reality or artefact? *BJOG* 114(9): 1097-103.