www.elsevier.com/locate/midw

Does a policy of earlier induction affect labour outcomes in women induced for postmaturity? A retrospective analysis in a tertiary hospital in the North of England

Anna Proctor, Paul Marshall

PII: S0266-6138(17)30299-1
DOI: http://dx.doi.org/10.1016/j.midw.2017.04.009
Reference: YMIDW2031

To appear in: *Midwifery*

Received date: 17 December 2015
Revised date: 7 April 2017
Accepted date: 23 April 2017

Cite this article as: Anna Proctor and Paul Marshall, Does a policy of earlie induction affect labour outcomes in women induced for postmaturity? A retrospective analysis in a tertiary hospital in the North of England, *Midwifery* http://dx.doi.org/10.1016/j.midw.2017.04.009

Does a policy of earlier induction affect labour outcomes in women induced for postmaturity? A retrospective analysis in a tertiary hospital in the North of England

Anna Proctor, RM, MSc (Professional Development Practitioner)[a,*], Paul Marshall, RN, PhD (Senior Lecturer and Postgraduate Research Tutor) [b]

[a] Women's Clinical Service Unit, Delivery Suite, Level 5 Gledhow Wing, St James' University Hospital, Beckett Street, Leeds, LS9 7TF, United Kingdom

[b] Adult, Child and Mental Health Nursing Academic Unit, School of Healthcare, Room G17, Baines Wing, University of Leeds, LS2 9UT, United Kingdom

[*]Corresponding author. E-mail address: anna.proctor2@nhs.net

Abstract

Objectives:

To investigate whether a change in the management of postmature pregnancy to earlier induction affects the length of labour and the induction process. Secondly, to assess the feasibility of the research process to inform a future larger study.

Design:

A change in management of postmature pregnancy in an NHS hospital in October 2013, from induction at 42 weeks gestation to induction between 41-42 weeks, provided an opportunity to conduct a retrospective analysis. Pre-existing data from the maternity database and casenotes were collected and primary outcomes analysed using the Mann-Whitney test and the Hodges-Lehman confidence interval for differences in medians.

Setting:

A large city based tertiary referral hospital in the North of England.

Participants:

125 women induced before the change in policy were compared with 309 women induced after the change.

Measurements:

Primary outcomes were length of 1st and 2nd stage of labour, overall length of labour, length of induction to established labour and length of induction to birth.

Findings:

The median overall length of labour for women induced at 42 weeks was 6.5 hours, while for women induced at 41-42 weeks this was 5.2 hours. The difference was not statistically significant (p = 0.15, 95% CI for median difference -0.27 to 1.93 hours) with a small effect size (Pearson's r = -0.08). The median length of induction to birth was 13.6 hours for women induced at 42 weeks and 16.5 hours for women induced at 41-42 weeks. This difference was also not statistically significant (p = 0.14, 95% CI for median difference -7.25 to 1.20 hours) with a small effect size (Pearson's r = -0.13).

Key conclusions and implications for practice:

This study demonstrated no statistically significant differences in length of labour and induction following a change in the management of postmature pregnancy to earlier induction. A large study is needed to establish definitively the effects of earlier induction on labour outcomes.

Keywords:

Prolonged pregnancy, Induced labour, Length of labour, Length of induction

Introduction

Postmature or prolonged pregnancy is defined as a pregnancy that continues beyond 42 weeks gestation (Doherty and Norwitz, 2008). It is estimated that up to 10% of pregnancies are prolonged, although this varies considerably between countries according to the accuracy of pregnancy dating and the use of induction of labour: for example, there are 0.4% in Austria but over 7% in Denmark (Zeitlin et al., 2007).

These pregnancies are of concern because there is substantial observational evidence indicating increasing fetal, neonatal and maternal risks as pregnancy continues beyond 40 weeks gestation (Caughey et al., 2008). Perinatal mortality rates have been found to increase significantly from 0.018% at 41 weeks, sharply rising after 42 weeks to 0.51% at 43 weeks, (Heimstad et al., 2008). Furthermore, there is evidence of an increased risk of morbidity including meconium aspiration, low Apgar scores, postpartum haemorrhage and caesarean section (Olesen et al., 2003). In an attempt to minimise these risks and prevent postmature pregnancy, United Kingdom (UK) national and international guidance recommend induction of labour between 41 and 42 weeks gestation (NICE, 2008; WHO, 2011). Induction, rather than expectant management (continuing the pregnancy), has been shown to decrease perinatal mortality, but the exact gestation at which induction is appropriate is not clear (Gülmezoglu et al., 2012). This uncertainty has resulted in variation with some UK hospitals offering induction as soon as women reach 41 weeks, some during week 41 and others delaying until 42 weeks.

Despite the benefits that induction may bring in preventing postmaturity, the intervention itself is not without risk. The most common complication is failed induction leading to caesarean section which has been estimated to occur in 17% of inductions (Wolfe et al., 2011). There is also a 5% risk of excessive uterine activity, known as tachysystole, which can reduce placental blood flow resulting in fetal heart rate abnormalities and eventual fetal hypoxia if untreated (Thomas et al., 2014). Observational studies have suggested that induction compared with spontaneous labour is associated with increased risk of caesarean delivery (Ehrenthal et al., 2010), postpartum haemorrhage (Phillip et al., 2004), increased admission to newborn intensive care and increased use of analgesia/anaesthetics (Guerra et al., 2009). However, there are concerns that using spontaneous labour as a comparison group may be inappropriate as this is not a choice available for clinicians and may lead to exaggerated estimates by excluding planned caesarean sections (Danilock et al., 2016).

Many of the undesirable effects of induction might be considered the likely result of intervening when the cervix is not ready for labour (Gülmezoglu et al., 2012). Spontaneous labour occurs following a gradual process of physiological changes towards the end of pregnancy and it is, therefore, easier to induce labour when a woman is further along in this natural process (NICE, 2008). It might seem logical that the more advanced the gestation at which induction takes place, the less complicated the process might be and also more women will labour spontaneously and avoid induction altogether. It has been estimated that inductions increase by 15-20% when a policy of induction at 41 rather than 42 weeks is in place (Menticoglou and Hall, 2002). These concerns have led to some advocating for later induction at 42 weeks (Mandruzzato, 2010).

Stillbirth, not least in late pregnancy, is a devastating complication and it could be argued that offering women earlier induction, resulting in higher induction rates, is a

small trade-off for preventing more of these deaths. Substantial evidence indicates that when induction is compared to expectant management, rather than spontaneous labour, there does not seem to be an increase in the risk of caesarean section, even when undertaken earlier than 41 weeks (Gulmezoglu et al., 2012, Stock et al., 2012, Walker et al., 2016). This reassuring evidence has led to calls for earlier induction, particularly in the current climate where there is a global priority to reduce perinatal mortality (Unicef & WHO, 2014). This is prudent in the UK where the government has set out an ambition to halve the stillbirth rate by 2030 (O'Connor, 2016). However, the exact gestation at which induction should be offered for the optimum balance of maternal and fetal risks still needs clarity. The WHO (2011) explicitly states that they regard the recommendation for induction after 41 weeks gestation as 'weak' and based on low quality evidence; hence further investigation is warranted.

Many women are healthy and regarded as low risk when their pregnancy continues beyond 40 weeks and they need to make an informed choice about how they wish to manage this. The absolute risk of stillbirth is low and estimates of the number of inductions needed to prevent one perinatal death are high at 416 (Gulmezoglu et al., 2012). Clearly the evidence regarding stillbirth and induction is crucial in this decision but given the rarity of stillbirth, women and the advising health professionals also need information about a range of other induction and labour outcomes which the current evidence base does little to provide. Two outcomes that have important clinical implications are the length of labour and the length of the induction process. Prolonged labour is associated with operative delivery (Adams 2012), postpartum haemorrhage (Sheiner et al 2005) and increased use of analgesia (Lancaster et al 2012). Furthermore, survey and qualitative evidence indicate that a longer labour and induction can result in lower satisfaction rates and negative accounts of birth (Shetty et al., 2005; Nystedt and Hildingsson, 2014; Murtagh and Folan, 2014. There is a dearth of evidence on how these outcomes are affected by the timing of induction and research is needed to inform appropriate counselling of women.

A change in policy in October 2013 at an NHS Trust from induction of labour at 42 weeks gestation to induction between 41 and 42 weeks, enabled a retrospective analysis of whether this change resulted in any differences in labour and induction characteristics. Additionally, this study was undertaken to inform how a much larger study could be conducted and to address the validity of methods for identifying induced women and their outcomes.

Methods

Design and participants

A retrospective analysis was undertaken at a large city hospital in the north of England which is part of an NHS Trust providing maternity care across two hospitals for approximately 10,000 births per annum. The main sample consisted of 434 women who were induced for postmaturity between January 2013 and June 2014. This was divided into those induced before implementation of the new induction policy (late group) and those after (early group):

- Late group – 125 women induced between 1st January 2013 and 30th September 2013 when management involved induction at 42 weeks (294 days).
- Early group – 309 women induced between 1st October 2013 and 30th June 2014 when management involved induction between 41 and 42 weeks (287-294 days).

The main sample size was dictated by the number of eligible women with available data in the study time period. Data were obtained from the hospital's maternity database (Matsys) which stores labour and birth details entered routinely after birth by midwives for administrative purposes. Data on the length of induction were entirely missing from the database and so case notes were accessed for this outcome but only completed for a sub-sample of 188 women from the main sample. The sub-sample consisted of 93 women in the late group and 95 women in the early group.

Exposure

The exposure in the study was induction of labour for postmaturity either before or after the change in policy to earlier induction. Eligibility criteria included women with a singleton pregnancy and a cephalic presentation who were induced for postmaturity. They were identified through the database and eligibility was assessed through the ward induction diary from 2013 and 2014, where midwives recorded the indication for each woman's induction. The diaries from one hospital were in poor condition and unusable; therefore eligible participants were all obtained from the same hospital. Women who were induced for other reasons were excluded and therefore the majority of included women were considered low risk. Nevertheless, a minority did have complications or risk factors which did not necessitate earlier induction or delivery. These included mild hypertension, previous caesarean delivery, parity more than 4, age outside the range of 18-39 and Body Mass Index (BMI) more than 35kg/m^2. Women were not excluded on this basis to ensure the sample was as representative as possible of the target population of women induced for postmaturity. Women who were augmented (labour accelerated) rather than induced were excluded as they were not relevant to the study aims.

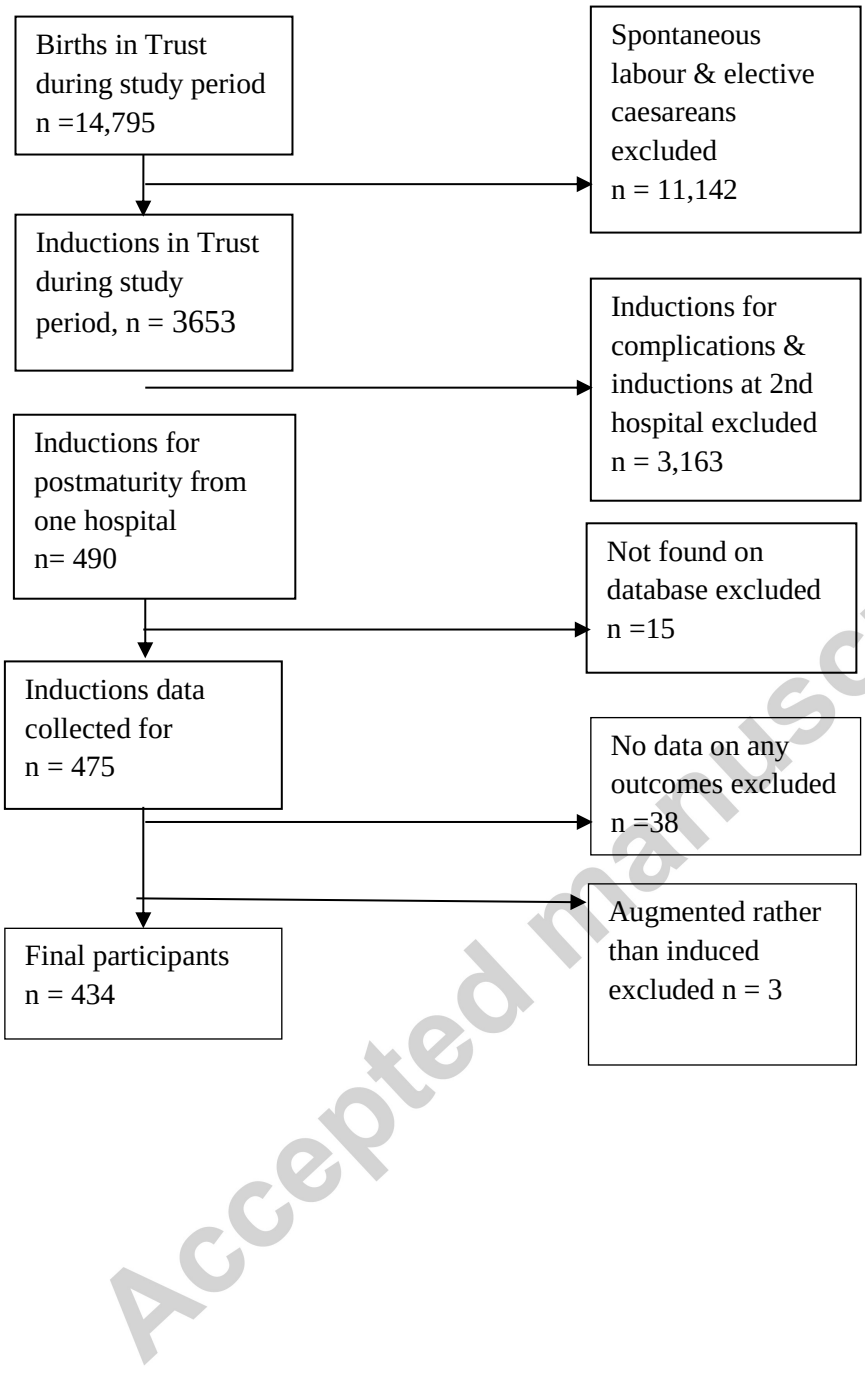
Pregnancies were dated using ultrasound to ensure accurate gestational age and women were offered a membrane sweep at 40 weeks gestation to increase their chances of spontaneous onset of labour. The induction process involved women with an unripe cervix (Bishop score < 7) being given a pessary of 10mg of vaginal prostaglandins (PGE_2) which remained in place for up to 24 hours. Amniotomy was performed when the cervix was ripe (Bishop score ≥ 7) and a syntocinon infusion was commenced if labour was not initiated soon after. These procedures were in accordance with the study hospital guidelines and UK national guidance (NICE, 2008). Figure 1 shows a flow chart of participants into the study.

Measurements

The primary outcomes for the main sample were the length of labour divided into 1st and 2nd stages and overall duration. The 1st stage of labour was defined as **e**stablished labour (regular painful contractions and progressive cervical dilatation from 4cm) to full dilatation of the cervix. The 2nd stage of labour was defined as full dilatation of the cervix to the birth of the baby and overall length of labour was defined as established labour to the birth of the baby (3rd stage of labour was not included). These definitions were in accordance with the study hospital guidelines and national guidance (NICE, 2014). The measurements were calculated from the start time of each labour stage documented in the Matsys database by midwives providing care.

The primary outcomes for the sub-sample were the length of induction and this was divided into induction to established labour and induction to birth. The start time of induction was recorded from the first induction intervention of either insertion of prostaglandins or amniotomy. Secondary outcomes were also investigated in the main sample to consider some adverse maternal and neonatal effects which have been associated with induction (Phillip et al., 2004; Guerra et al., 2009). These included postpartum haemorrhage (estimated blood loss of > 500mls) and admission to newborn intensive care.

Figure 1. Flow chart of participants into the study



Late group: 42wk induction n = 125	Early group: 41-42wk induction n = 309

Statistical analysis

Data were analysed using IBM SPSS Statistics, version 22. Continuous data for primary outcomes and participant baseline characteristics were skewed in distribution and therefore summarised using medians and interquartile ranges (IQR) which are generally more representative (Gosall and Gosall, 2012). The exception was maternal age where the mean and standard deviation (SD) were most appropriate as the data was normally distributed. Categorical data were summarised using frequencies (number of cases) and relative frequencies (percentages). Confidence intervals (CI) for differences in medians across the study groups for primary outcomes were calculated using the Hodges-Lehman method.

The length of labour and induction were examined using the Mann Whitney test which allowed for the non-normal distributions. As parity is known to influence the length of labour and induction, the analysis for the overall length of labour and the length of induction to birth was repeated after stratifying the study groups into primiparous and multiparous women. Sensitivity analysis excluding caesarean section births was undertaken to explore any confounding effects on the length of labour and induction. The categorical secondary outcomes were compared using the Chi-square test for postpartum haemorrhage and Fisher's exact test for admission to newborn intensive care where the expected frequency was less than 5. The significance level for a statistical difference was taken at p value < 0.05 for all the variables. Missing data were excluded from the analysis rather than imputing values using mean substitution which can distort the results (Pallant, 2013).

In this small study it was considered useful to calculate an objective and standardised measure of the size of the observed effect which, unlike hypotheses testing, is independent of the sample size (Sullivan and Feinn, 2012). The effect size, denoted by Pearson's r, was calculated using z values provided from the Mann Whitney test for all the primary outcomes. The following formula was used to provide an approximate effect size: $r = z \div \sqrt{n}$, where n is the total number of cases (Field, 2013). A small effect can be regarded as $r = 0.1$, a medium effect as $r = 0.3$ and a large effect as $r = 0.5$, explaining 1%, 9% and 25% of the total variance respectively (Cohen, 1988).

While the study planned to explore outcomes and control baseline characteristics of the two groups using multivariate linear regression analyses, the nature of the data meant this was considered inappropriate. Transforming data to improve normality can be complex and may be of debatable value (Grayson, 2004; Wilcox, 2012) and was therefore not considered suitable for this study. However, the baseline characteristics were compared to test the hypothesis of no difference between the study groups other than gestational age. The independent samples t-test was appropriate for the normally distributed data of maternal age and the other continuous variables were analysed using the Mann Whitney test. Categorical data were analysed using the Chi-square test,

except caesarean section in a previous pregnancy where the expected frequency was less than 5 making Fisher's exact test more suitable.

Ethical considerations

Retrospective data were anonymised at the point of collection and there was strict adherence to the Data Protection Act of 1998 and the hospital governance policies throughout the process. Ethical approval was obtained from the university Ethics Research Committee (ref no. SHREC/RP/490) and the Research and Innovation department at the NHS Trust (ref no. OG15/125).

Findings

In the main sample, there were 125 inductions in the late group (42 week induction) and this increased to 309 in the early group (41-42 week induction). The late group had a median gestational age of 294 days (42 weeks + 0 days) which decreased by 5 days to 289 days gestation (41 weeks and 2 days) in the early group. Table 1 shows the distribution of baseline characteristics. The baseline characteristics were similar with no significant differences across the groups, except for women in the main sample who had a caesarean in a previous pregnancy ($p = 0.04$), but this involved too few women to be important (4 women in total).

The findings for length of labour outcomes are presented in Table 2. The median overall length of labour for women induced at 42 weeks (late group) was 6.5 hours, compared with 5.2 hours for women induced at 41-42 weeks (early group). This difference was not statistically significant ($p = 0.15$, 95% CI for the median difference -0.27 to 1.93 hours) and there was a small calculated effect size (Pearson's $r = -0.08$). When stratified by parity, the median overall length of labour still demonstrated no statistically significant differences between the study groups. Sensitivity analysis excluding caesarean deliveries yielded similar results with no differences between the groups.

Table 3 presents the findings for length of induction outcomes. There was a small calculated effect size for induction to birth duration (Pearson's $r = -0.13$) with a median duration of 13.6 hours in women induced at 42 weeks (late group) and 16.5 hours in women induced at 41-42 weeks (early group) but this difference was not statistically significant ($p = 0.14$, 95% CI for the median difference -7.25 to 1.20 hours). No statistically significant differences were found when the length of induction to birth was analysed separately for primiparous and multiparous women. The findings were similar when caesarean deliveries were excluded from the analysis.

The results for the secondary outcomes are shown in Table 4 and show no differences between groups. Postpartum haemorrhage occurred in 24.8% of participants in the late group and 25.9% in the early group (p = 0.12). Participants who had a baby admitted to newborn intensive care accounted for 2.4% of cases in the late group compared with 0.6% in the early group (p = 0.88).

Table 1. Baseline characteristics

	Main Sample n = 434			**Sub-sample n=188**		
	Late group n =125	**Early group** n = 309	**P value***	**Late group** n =93	**Early group** n = 95	**P value***
Age (years) mean (SD)	29.8 (5.6)	29.6 (6.0)	0.773	29.7 (5.6)	29.2 (6.0)	0.597
Data missing	0	0		0	0	
BMI (Kg/m^2) median (IQR)	25.5 (8.0)	25.0 (8.0)	0.311	25.0 (8.0)	24.0 (7.0)	0.552
Data missing	1	1		1	1	
Gestational age (days) median (IQR) Range	294 (1.0) 287-296	289 (4.0) 286-294		294 (1.0) 287-296	289(5.0) 287-297	
Data missing	0	0		0	0	
Ethnicity, n (%)			0.540			0.669
White	79 (63.2)	196 (63.4)		58 (62.4)	59 (62.1)	
Black	12 (9.6)	38 (12.3)		10 (10.8)	10 (10.5)	
Asian	15 (12)	36 (11.7)		11 (11.8)	12 (12.6)	
Other	8 (6.4)	24 (7.8)		4 (4.3)	8 (8.4)	
Unknown	11 (8.8)	15 (4.9)		10 (10.8)	6 (6.3)	
Data missing	0	0		0	0	

Table 1: Baseline Characteristics Continued

	Main Sample n = 434			**Sub-sample n=188**		
	Late group n =125	**Early group** n = 309	**P value***	**Late group** n =93	**Early group** n = 95	**P value***
Parity, n (%)			0.726			0.999
Primiparous	62 (49.6)	159 (51.5)		47 (50.5)	48 (50.5)	

Multiparous	63 (50.4)	150 (48.5)		46 (49.5)	47 (49.5)	
Data missing	0	0		0	0	
Regional Analgesia, n (%)	62 (49.6)	155 (50.2)	0.916	46 (49.5)	42 (44)	0.471
Data missing	0	0		0	0	
Previous caesarean, n (%)	3 (2.4)	1 (3)	**0.040**	2 (2.2)	0 (0)	0.151
Data missing	0	0		0	0	
Mode of delivery, n (%)			0.091			0.483
Spontaneous delivery	74 (59.2)	199 (64.4)		56 (60.2)	65 (68.4)	
Assisted vaginal delivery[a]	22 (17.6)	65 (21)		16 (17.2)	14 (14.7)	
Caesarean Section	29 (23.2)	45 (14.6)		21 (22.6)	16 (16.8)	
Data missing	0	0		0	0	

*P value < 0.05 indicates statistical significance by t- test, Mann-Whitney test, Chi-square test or Fisher's exact test

*P value < 0.05 indicates statistical significance by t- test, Mann-Whitney test, Chi-square test or Fisher's exact test

[a] indicates forceps or vacuum extraction delivery

Table 2. Primary outcomes of main sample: length of labour

	Late group Median hours (IQR)	**Early group** Median hours (IQR)	**95% CI for median difference**	**P value***	**r value****
	Primiparous and multiparous women combined				
	n = 125	n = 309			
1st stage labour	3.6 (5.2)	3.4 (3.5)	-0.48 to 1.12	0.49	-0.04
Missing data	55	115			
2nd stage labour	0.6 (2.5)	0.9 (2.1)	-0.18 to 0.15	0.97	-0.01
Missing data	43	63			
Overall length[b]	6.5 (9.6)	5.2 (6.7)	-0.27 to 1.93	0.15	-0.08
Missing data	27	59			

		Primiparous women			
	n = 62	n = 159			
Overall length[b]	8.3 (7.15)	7.0 (5.38)	-0.03 to 3.1	0.054	-0.14
Missing data	11	25			
		Multiparous women			
	n = 63	n = 150			
Overall length[b]	2.8 (6.52)	3.2 (4.18)	-0.87 to 0.97	0.92	0.008
Missing data	16	34			
		Sensitivity analysis excluding caesarean delivery: primiparous and multiparous women combined			
	n = 96	n = 264			
1st stage labour	3.4 (5.1)	3.4 (3.5)	-0.53 to 1.05	0.59	-0.03
Missing data	28	73			
2nd stage labour	0.6 (2.55)	0.9 (1.98)	-0.20 to 0.13	0.86	-0.01
Missing data	16	25			
Overall length[b]	5.0 (6.7)	4.7 (5.5)	-0.83 to 1.15	0.83	-0.01
Missing data	21	52			

*P value <0.05 indicates statistical significance

** r value indicates effect size: 0.1 small effect, 0.3 medium effect, 0.5 large effect

[b] indicates length of established labour to birth excluding 3rd stage of labour

Table 3: Primary outcomes of sub-sample: length of induction

	Late group Median hours (IQR)	**Early group** Median hours (IQR)	**95% CI for median difference**	**P value***	**r value****
		Primiparous and multiparous women combined			
	n = 93	n = 95			
Induction to established labour	11.7 (23.5)	13.3 (22.6)	-5.92 to 3.00	0.50	-0.07

Missing data	47	42			
Induction to birth	13.6 (24.0)	16.5 (23.4)	-7.25 to 1.20	0.14	-0.13
Missing data	24	27			
	Primiparous women				
	n = 47	n = 48			
Induction to birth	25.5 (22.0)	24.8 (20.1)	-8.08 to 7.0	0.72	0.04
Missing data	16	11			
	Multiparous women				
	n = 46	n = 47			
Induction to birth	9.1 (11.1)	12.5 (9.5)	-6.02 to 1.55	0.23	0.15
Missing data	8	16			
	Sensitivity analysis excluding caesarean delivery: primiparous and multiparous women combined				
	n = 72	n = 79			
Induction to established labour	9.2 (22.1)	11.5 (19.19)	-0.497 to 3.67	0.75	-0.04
Missing data	35	37			
Induction to birth	11.9 (21.3)	15.2 (16.8)	-6.05 to 1.82	0.25	-0.11
Missing data	19	26			

*P value <0.05 indicates statistical significance

** r value indicates effect size: 0.1 small effect, 0.3 medium effect, 0.5 large effect

Table 4: Secondary outcomes of main sample

	Late group n =125	**Early group** n = 309	**P values***
Postpartum haemorrhage, n (%)	31 (24.8)	80 (25.9)	0.12

Missing data	6	10	
Admission to newborn intensive care, n (%)	3 (2.4)	2 (0.6)	0.88
Missing data	0	0	

* P value <0.05 indicates statistical significance by the chi-squared test or Fisher's exact test

Discussion

This study found that a change in the management of postmature pregnancy, from induction at 42 weeks gestation to induction between 41 and 42 weeks, did not affect the length of labour or induction. Similar studies comparing two different timings of induction and the effect on these outcomes are limited with contradictory findings. A UK based study found no difference in the length of labour in multiparous women induced at 40 weeks plus 10 days compared with women induced at 42 weeks (Kassab et al., 2011). This finding supports the current study and may reflect the comparable NHS hospital study populations where care is likely to be similar. In contrast, a study in Hong Kong found that women induced for postmaturity at 41 weeks rather than 42 weeks had a longer labour by 1.2 hours ($p<0.001$) (Fok et al., 2006). However, there were more primiparous women in the 41 week group (59%) than the 42 week group (51%) which may have contributed to the longer labour. This was a much larger study than the current study with 449 women induced for postmaturity at 42 weeks and 2043 at 41 weeks so may better represent the true differences in their population. A larger study in the UK may not show the same findings because of different healthcare systems/midwifery practice and particularly if the number of primiparous and multiparous women in each group were similar as they were in the current study.

Some studies evaluating other outcomes have found that a policy of induction during 41 weeks gestation, compared with a policy of later induction at 42 weeks, may be associated with some adverse outcomes including increased caesarean sections (Kassab et al., 2011; Burgos et al., 2012), increased epidural use (Fok et al., 2006) and more newborns with low cord pH (Burgos et al., 2012). However, a recent much larger study has found that induction at 41 weeks and 5 days, rather than 42 weeks, was associated with decreased caesarean sections (Kjeldsen et al., 2015). Importantly, this study was large enough to consider stillbirth rates and found no difference between the

two timings of induction. The main purpose of an earlier induction policy is to reduce the risk of stillbirth whilst balancing the risk of adverse maternal outcomes. Although perinatal mortality increases progressively from 37 weeks (Heimstad et al., 2008), it may be that earlier induction by only a few days has little effect on perinatal deaths.

There is evidence that cervical ripeness and gestational age can predict the length of induction to birth and the success of induction (Elghorori et al., 2006; Braems and Norhausen, 2007). It therefore seems plausible that the more advanced the gestation, the more favorable the cervix and the easier the induction might be. While data were not collected on cervical ripeness, the trends in the data from this study were for a slightly shorter induction when undertaken later at 42 weeks rather than earlier between 41-42 weeks. However, no overall differences were found and it is possible that a few days difference in gestation does not profoundly affect cervical ripeness and subsequent speed of the induction. If the findings of this study were confirmed in a larger study, it would provide some reassurance to women and clinicians that earlier induction during week 41 does not significantly increase the risk of a long arduous birth and a negative experience.

There is clearly a dearth of evidence concerned with the optimal timing of induction for postmaturity and this study contributes to a basis for further research that might provide the evidence needed to guide women and practitioners in making appropriate decisions. In particular, there is very limited research regarding the impact of the timing of induction on the length of labour and, as far as the authors are aware, there are no other studies investigating the effect on the length of induction. Therefore, the strength of this study is in providing a unique analysis of these important outcomes which can affect women's experience of birth and some labour complications. Using retrospective clinical data means the findings are likely to reflect the effect that might occur in clinical practice. A prospective study may have better control over data collection but the outcome may be subject to a Hawthorne effect when staff collecting data are aware of the study purpose. Importantly, this study provides information on the research process that could be used to guide and instigate further research in this much needed area.

A limitation of this study was the size and the lack of power to detect statistically significant differences. Retrospective power analysis indicated that a sample size of 303 in each group was needed to achieve a power of 80% (alpha level of 0.05) for the 1.2 hour difference in the overall length of labour found between the study groups. For a future study, a 'clinically important' difference between study groups needs to be considered and might be regarded as 1 hour. On this premise, sample size calculations to achieve a power of 80% (alpha level of 0.05) estimate that a sample size of 563 in each group (1126 in total) would be needed to detect a 1 hour difference. These figures are based on sample size calculations for the Mann Whitney test using G*Power software which supports the non-normal data distributions of length of labour and induction (Faul et al., 2007). The retrospective power estimates used an effect size of

0.25 calculated from the mean length of labour and standard deviation observed in each group of the current study. The sample size calculation for a future study used an effect size of 0.18 calculated from a 1 hour 'clinically important' difference and the standard deviation of the length of labour observed in the late group of this study.

The exposure classification of induction for postmaturity relied on ward induction diaries completed by midwives for administrative purposes and this meant that inaccuracies were a possibility as the induction indication could not be verified. There was a dramatic increase in the number of inductions following implementation of the new policy for earlier induction which far exceeded the expected rise. It is possible that the rise was exaggerated because midwives were concerned about increasing workloads and so keen to evidence postmature inductions in the diary.

Missing data were problematic for all the primary outcomes. The lack of data on the start time of induction was identified early in the study enabling case notes to be accessed for this information and induction duration calculated. However, the onset of established labour was also poorly recorded in the Matsys database and it was not feasible to access casenotes which resulted in significant missing data on the length of the 1st stage of labour and induction to established labour. This decreased the power of the analysis and could be a source of potential bias if those participants with missing data had a particular type of labour. For example, it is possible that onset of established labour may be poorly recorded for women who experienced a very quick labour due to difficulties in precise diagnosis and this could lead to exaggerated estimates of labour duration. Indeed, it is suggestive that the data were not missing at random because the median 1st and 2nd stage of labour are markedly different from the overall length of labour.

The baseline characteristics included some potential confounders which were unable to be controlled for using robust techniques such as regression analysis. However, some reassurance can be gained from the fact there were no significant differences between the groups for potential confounders, with the exception of previous caesarean which only involved 4 participants. Sensitivity analysis also showed there was no undue influence on the results from participants who experienced a caesarean delivery. Nonetheless, the influence of confounders on the findings cannot be ruled out.

Feasibility of a large study

The methods for classifying the exposure and capturing data would not be feasible for a large study due to potential bias. However, since completion of this study the systems for recording care at the study hospital and many other UK hospitals has changed to a paperless electronic system which includes prompts to ensure appropriate information is recorded. These systems may enable accurate classification of induction of labour for postmaturity and complete data on the length of induction and labour. It would also enable data on a large number of participants to be obtained in order to undertake a

sufficiently powered study. The feasibility of the data storage systems at hospitals that have recently changed their induction policy would need to be investigated.

Conclusions and implications for practice

No significant differences were found in the length of labour and induction when the management of postmature pregnancy changed from induction at 42 weeks to induction between 41 and 42 weeks. A large definitive study would provide evidence to confirm or refute these findings and establish if earlier induction does affect labour outcomes. Such a study would only be feasible with substantial improvements in the completeness of recorded data.

Conflict of interest

The authors were not aware of any conflict of interest

Funding

This report is independent research supported by the National Institute for Health Research / Chief Nursing Officer Clinical Academic Training Programme Masters in Clinical Research Scheme. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

Acknowledgements

Thanks to the administrative and informatic hospital staff for assisting with access to the data.

References

Adams, S., Eberhard-Gran, M., Eskild, A. 2012. Fear of childbirth and duration of labour: a study of 2206 women with intended vaginal delivery. BJOG: An International Journal of Obstetrics and Gynaecology. **119**, pp. 1238-1246. [Online]. [Accessed 3rd October 2016]. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1471-0528.2012.03433.x

Bowers, D., House, A., Owens, D. and Bewick, B. 2014. Understanding Clinical Papers (3rd Edition). Wiley, Chichester.

Braems, G. and Norhausen, I. 2007. Induction of labor with prostaglandins for medical reasons: Determining explanatory variables of the induction to delivery time interval for vaginal deliveries and caesarean section. European Journal of Obstetrics and Gynecology and Reproductive Biology. **135**, pp.164-169.

Burgos, J., Rodriguez, L., Otero, B., Cobos, P., Osuna, C., Centeno, M.D.M., Melchor, J.C., Fernandez-Llebrez, L. and Martinez-Astorquiza, T. 2012. Induction at 41 weeks increases the risk of caesarean section in a hospital with a low rate of caesarean sections. Journal of Maternal-Fetal and Neonatal Medicine. **25**(9), pp.1716-1718.

Caughey, A.B., Snegovskikh, V.V. and Norwitz, E.R. 2008. Postterm pregnancy: how can we improve outcomes? Obstetrical & Gynecological Survey. **63**(11), pp.715-724.

Cohen, J. 1988. Statistical power and analysis for the behavioural sciences. 2nd ed. Academic press, New York.

Danilack, V.A., Triche, E.W., Dore, D.D., Muri, J.H., Phipps, M.G., Savitz, D.A. 2016. Comparing expectant management and spontaneous labor approaches in studying the effect of labor induction on cesarean delivery. Annals of Epidemiology. **26**(6), pp. 405-411.[f]

Doherty, L. and Norwitz, E.R. 2008. Prolonged pregnancy: when should we intervene? Current Opinion in Obstetrics & Gynecology. **20**(6), pp.519-527.

Ehrentha, D.B., Jiang, X., Strobino, D.M. 2010. Labor induction and the risk of a cesarean delivery among nulliparous women at term. Obstet Gynecol. **116**(1), pp. 35-42.

Elghorori, M.R., Hassan, I., Dartey, W., Abdel-Aziz, E. and Bradley, M. 2006. Comparison between subjective and objective assessments of the cervix before induction of labour. Journal of Obstetrics & Gynaecology. **26**(6), pp.521-526.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. **39**, pp. 175-191.

Field, A. 2013. Discovering Statistics Using IBM Statistics. 4th ed. Sage, London.

Fok, W.Y., Chan, L.Y.S., Tsui, M.H.Y., Leung, T.N., Lau, T.K. and Chung, T.K.H. 2006. When to induce labor for post-term?. A study of induction at 41 weeks versus 42 weeks. European Journal of Obstetrics Gynecology and Reproductive Biology. **125**(2), pp.206-210.

Gosall, N. and Gosall, G. 2012. The Doctor's Guide to Critical Appraisal. PasTest, Cheshire.

Grayson, D. 2004. Some myths and legends in quantitative psychology. Understanding Statistics. **3**(1), pp.101-134.

Guerra, G., Cecatti, J., Souza, J., Faundes, A., Morais, S., Gulmezoglu, A., Parpinelli, M., Passini, R. and Carroli, G. 2009. Factors and outcomes associated with the induction of labour in Latin America. BJOG : an international journal of obstetrics and gynaecology. **116**(13), pp.1762-1772.

Gülmezoglu, A.M., Crowther, C.A., Middleton, P. and Heatley, E. 2012. Induction of labour for improving birth outcomes for women at or beyond term. Cochrane Database of Systematic Reviews. (6).

Harrison, J.M. 2000. The initiation of labour: physiological mechanisms. British Journal of Midwifery. **8**(5), pp.281-286.

Heimstad, R., Romundstad, P.R. and Salvesen, K.A. 2008. Induction of labour for post-term pregnancy and risk estimates for intrauterine and perinatal death. Acta Obstetricia et Gynecologica Scandinavica. **87**, pp.247-249.

Kassab, A., Tucker, A., El-Bialy, G., Mustafa, M., Macaulay, J. and Fox, R. 2011. Comparison of two policies for induction of labour postdates. Journal of Obstetrics & Gynaecology. **31**(1), pp.32-36.

Lancaster, S.M., Schick, U.M., Oman, M.M., Enquobahrie, D.A. 2012. Risk factors associated with epidural use. Journal of Clinical Medicine Research. 4 (2), pp. 119-126.

Kjeldsen, L.L., Sindberg, M., Maimburg, R.D. 2015. Earlier induction of labour in post term pregnancies - A historical cohort study. Midwifery. **31**(5), pp.526-531.

Lancaster, S.M., Schick, U.M., Osman, M.M., Enquobahrie, D.A. 2012. Risk factors associated with epidural use. Journal of Clinical Medicine. **4**(2), pp.119-126.

Mandruzzato, G., Alfirevic, Z., Chervenak, F., Gruenebaum, A., Heimstad, R., Heinonen, S, et al. 2010. Guidelines for the management of postterm pregnancy. Journal of Perinatal Medicine. **38**, pp. 111-119.

McCarthy, F.P. and Kenny, L.C. 2014. Induction of labour. Obstetrics, Gynaecology and Reproductive Medicine. **24**(1), pp.9-15.

Menticoglou, S.M. and Hall, P.F. 2002. Routine induction of labour at 41 weeks gestation: nonsensus consensus. BJOG: An International Journal of Obstetrics & Gynaecology. **109**(5), pp.485-491.

Murtagh, M., Folan, M. 2014. Women's experiences of induction of labour for post-date pregnancy. British Journal of Midwifery. **22**(2), pp. 105-109.

NICE. 2008. Induction of labour. [Online]. [Accessed 20th June 2015]. Available from: http://www.nice.org.uk/guidance/cg70/evidence/cg70-induction-of-labour-full-guideline2

NICE. 2014. Intrapartum care: care of healthy women and their babies during childbirth. [Online]. [Accessed 4th July 2015]. Available from: http://www.nice.org.uk/guidance/cg190/resources/guidance-intrapartum-carecare-of-healthy-women-and-their-babies-during-childbirth-pdf

Nystedt, A., Hildingsson, I. 2014. Diverse definitions of prolonged labour and it's consequences with sometimes subsequent inappropriate treatment. BMC Pregnancy Childbirth. 14 (233), pp. 1-11. [Online]. [Accessed 1st October 2016]. Available from: http://www.biomedcentral.com/1471-2393/14/233

O'Connor, D. 2016. Saving Babies' Lives; a care bundle for reducing stillbirths. NHS England, Leeds.

Olesen, A.W., Westergaard, J.G. and Olsen, J. 2003. Perinatal and maternal complications related to postterm delivery: a national register-based study, 1978-1993. American Journal of Obstetrics & Gynecology. **189**, pp.222-227.

Pallant, J. 2013. SPSS Survival Manual. 5th ed. Open University Press, Berkshire.

Phillip, H., Fletcher, H. and Reid, M. 2004. The impact of induced labour on postpartum blood loss. Journal of Obstetrics & Gynaecology. **24**(1), pp.12-15.

Sheiner, E., Sarid, L., Levy, A., Seidman, D.S., Hallak, M. 2005. Obstetric risk factors and outcome of pregnancies complications with early postpartum haemorrhage: a population-based study. Journal of Maternal, Fetal and Neonatal Medicine. **18**, pp. 149-154.

Shetty, A., Burt, R., Rice, P., Templeton, A. 2005. Women's perceptions, expectations and satisfaction with induced labour - A questionaire-based study. European Journal of Obstetrics and Gynecology. **123**(1), pp. 56-61.

Stock, S.J., Ferguson, E., Duffy, A., Ford, I., Chalmers, J., Norman, J.E. 2012. Outcomes of election induction of labour compared with expectant management: population based study. BMJ. **344** (2838), pp. 1-13

Sullivan, G.M. and Feinn, R. 2012. Using effect size - or why the p value is not enough. Journal of Graduate Medical Education. **4**(3), pp.279-282.

Thomas, J., Fairclough, A., Kavanagh, J. and Kelly, A.J. 2014. Vaginal prostaglandin (PGE2 abd PGF2a) for induction of labour at term. Cochrane Database of

Systematic Reviews. [Online]. (6). [Accessed 10th March 2015]. Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003101.pub3/pdf

UNICEF & WHO. 2014. Every Newborn: an action plan to end preventable deaths. World Health Organisation, Geneva.

Walker, K.F., Bugg, G.J., Macpherson, M., McCormick, C., Grace, N., Wildsmith, C., Bradshaw, L., Smith, G.C.S., Thornton, J.G. 2016. Randomized Trial of Labor Induction in Women 35 Years of Age or Older. The New England Journal of Medicine. **374**, pp. 813-822.

WHO. 2011. WHO recommendations for induction of labour. [Online]. [Accessed 24th March 2014]. Available from: http://whqlibdoc.who.int/publications/2011/9789241501156_eng.pdf?ua=1

Wilcox, R.R. 2012. Introduction to robust estimation and hypothesis testing. 3rd ed. M.A: Elsevier, Burlington.

Wolfe, K.B., Rossi, R.A. and Warshak, C.R. 2011. The effect of maternal obesity on the rate of failed induction of labor American Journal of Obstetrics & Gynecology. **205**(2), pp.128.e121-128.e127.

Zeitlin, J., Blondel, B., Alexander, S., Breart, G. and GROUP, P. 2007. Variation in rates of postterm birth in Europe: reality or artefact? BJOG : an international journal of obstetrics and gynaecology. **114**, pp.1097-1103.