Randomised comparison of early versus late induction of labour in post-term pregnancy

KÄRE AUGENSEN, PER BERGSJØ, TORUNN EIKELAND, KJELL ASKVIK, JOHANNES CARLSEN

References

Abstract
In a prospective randomised study of mothers referred for prolonged pregnancy (around the 42nd week) 214 (group 1) were submitted to attempted induction of labour and 195 (group 2) assigned to continue for a further week without intervention. Strict selection criteria were used for the certainty of term. Mothers in group 2 were given regular non-stress tests to ensure fetal wellbeing, as were those in group 1 in whom induction failed. In group 1, 48 (23%) out of 210 first attempted inductions failed. In group 2, 135 (69%) of the births started spontaneously as compared with 38 (18%) in group 1. The mean duration of labour was 7.5 hours in each group. There was no significant difference in incidence of operative delivery, use of analgesics, or signs of perinatal asphyxia. Significantly more children in group 1 needed phototherapy for hyperbilirubinaemia. There was a clustering of births in the late afternoon and evening, which was most pronounced in group 1.

A policy of vigilant non-intervention up to the 44th completed week of pregnancy does not appear to jeopardise mother or fetus.

Introduction
Clinical decisions vary considerably when pregnancy continues beyond the accepted normal term. Systematic induction of labour has been justified as a means of averting placental insufficiency and fetal death. Another argument for induction is that good timing will ensure delivery during hours of optimal staffing and preparedness for emergency treatment. As some 10% of pregnancies go beyond 42 completed weeks, even when strict criteria for reliability of menstrual dates are employed, a rigorous policy of induction after term generates a large extra clinical workload.

In Norway the established practice for mothers who pass the 44th week limit is to refer them to the maternity hospital for evaluation. From then on the course of action depends on local policy. In our hospital the induction policy was fairly liberal. We planned and conducted the following study in order to obtain a more rational basis for future decision making.

Patients and methods
Design of study—The ideal comparison would be induction at the end of 42 weeks (294 days) versus no intervention before spontaneous labour. This, however, was not feasible, as the risk of prolonged pregnancy is ultimately greater than that of any deliberate form of intervention. We decided on 30-31 days as the limit for non-intervention and designed a randomised study of immediate induction of labour in women referred for post-term pregnancy versus induction one week later in those who had not delivered in the interim.

Criteria for inclusion—We included in the study only healthy women with normal pregnancies. Other criteria for inclusion were a single fetus in cephalic presentation; a duration of pregnancy of 290 to 297 days from the first day of the last menstrual period; and reliable dates. Reliable dates were defined as regular menstrual periods (28±4 day intervals) and clear recollection of dates. Use of contraceptive pills during the two months before the last menstrual period was a cause for exclusion. Table I lists other reasons for exclusion.

Examination, randomisation, and management—Pregnant women who had not delivered by about 42 weeks were referred by their doctors. After scrutiny of the menstrual and pregnancy histories one of us performed a clinical examination, and if all criteria for enrolment were fulfilled randomisation (non-stratified) was done. The midwife consulted a list of random numbers, which was inaccessible to the participating physicians. Women in group 1 (immediate induction) were then referred to the delivery department for induction. Those assigned to group 2 (postmenstrual induction) were submitted to cardiocographic non-stress tests on the day of referral (day zero) and again on day 3 or 4 if still undelivered. If birth had not occurred by day 7 labour was induced. In cases of failed induction in group 1 further management was as for group 2. For mothers who were still undelivered after the attempted induction on day 7 management was left to the clinical judgment.

Department of Obstetrics and Gynaecology, University of Bergen, Kvinnel inkulen, Haukeland Sykehus, N-5021 Bergen, Norway
KARE AUGENSEN, MD, senior registrar (reserveveile)
PER BERGSJØ, MD, PhD, professor of obstetrics and gynaecology
TORUNN EIKELAND, MD, junior registrar (assistantveile)
KJELL ASKVIK, MD, senior registrar (reserveveile)
JOHANNES CARLSEN, MD, junior registrar (assistantveile)

Correspondence to: Professor Bergsjo.
Method of induction—Labour was induced with 5 IU oxytocin in 500 ml 5% glucose given by intravenous drip infusion, dose rates being increased stepwise according to response. In exceptional cases amniotomy was performed at the start of induction but otherwise only once labour was established. If labour was not clearly established after six to eight hours of infusion induction was considered unsuccessful. A cardiotocographic recording was obtained before disconnection from the drip in these cases.

Data handling and statistics—Data were transferred to electronic tape consecutively as soon as all information on mother and child was available. We used a predetermined code for the case tabulations. Statistical analyses were by Student’s t test and χ2 tests, as appropriate, with the help of a Minitab package (Statistics Department, Pennsylvania State University).

Execution—Enrolment started on 1 January 1982. We decided to collect 400 cases and estimated that this would take one or possibly two years, given the rate of post-term referrals at the time. The last patient (case 409), however, was enrolled in June 1985. An important reason for the slow progress was the number of women ineligible for the study (table 1). Out of 944 women referred for prolonged pregnancy, 535 (56·7%) were excluded, mostly for unreliable dates. In addition, increasing use of early ultrasoundography during the trial reduced the daily number of post-term referrals.

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>No of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term uncertain for various reasons</td>
<td>373</td>
</tr>
<tr>
<td>Hypertension or growth retarded fetus, or both</td>
<td>55</td>
</tr>
<tr>
<td>Other medical conditions</td>
<td>8</td>
</tr>
<tr>
<td>Various obstetric conditions</td>
<td>43</td>
</tr>
<tr>
<td>Birth starting spontaneously</td>
<td>19</td>
</tr>
<tr>
<td>Geographical, social, other</td>
<td>18</td>
</tr>
<tr>
<td>Duration of pregnancy outside limits for inclusion</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>535</td>
</tr>
</tbody>
</table>

Comparability of groups—In order to ensure strict comparability of the groups a non-stress test would need to have been performed initially on all mothers in the study. This, however, would have entailed a new element of clinical testing before induction, as it was not done routinely before the trial. On the other hand, omitting non-stress tests in group 2 was considered unethical, as a reactive non-stress test result was an added measure of safety in allowing the pregnancy to continue.

Ethical considerations—Seeking informed consent of the mothers would, in all likelihood, have invalidated the trial. Given the short time available between the decision on eligibility and randomisation, we could not expect mothers fully to understand the trial, much less to make a decision on whether to join. Furthermore, as the participating physicians could not support a preference for one type of management over the other we considered that drawing lots without informing the mother was justified. The study was put before the chairman of the ethical board of the Norwegian Medical Association and before Haukeland Hospital's ethical committee and was approved.

Results

OUTCOME OF RANDOMISATION

Of the 409 eligible mothers, 214 were allocated to group 1 and 195 to group 2. Possibly confounding factors turned out to be distributed similarly between the two groups. Group 1 included 137 (46%) nulliparous mothers and group 2, 82 (42%); and Bishop scores (express ripeness of cervix) were equally distributed, scores below six being recorded in 77 (36%) mothers in group 1 and 69 (35%) in group 2.

LABOUR AND DELIVERY

Figure 1 shows the sequence of events in each group. Four women in group 1 went into spontaneous labour before induction could be attempted.

Group 1

- Birth after induction of labour
  - First induction attempted: 214
  - Successful induction: 162
  - Second induction attempted: 10
  - Remaining undelivered: 30
  - Third induction attempted: 4

Group 2

- Birth after induction or elective caesarean section
  - Elective caesarean section: 5
  - Remaining undelivered: 190
  - Induction attempted according to protocol: 52
  - Successful induction: 44
  - Remaining undelivered: 8
  - Second induction attempted: 4

FIG 1—Flow charts of sequence of events in the two groups.
Four of the five initial caesarean sections in group 2 were performed because of abnormal patterns on cardiotocography, two on day zero and two on days 3 and 4. In the fifth case disproportion was found on re-examination.

In both groups a proportion of women were submitted to repeated induction attempts after initial failures, some having up to three courses. In group 1, 48 out of 210 (23%) of the first courses of induction failed, as compared with eight out of 52 (15%) in group 2; these inductions in group 2 were performed on average one week later than in group 1. This difference was not significant. Figure 2 shows the intervals between randomisation and delivery in the two groups. The mean interval was 1-4 days in group 1 and 4-0 days in group 2. Roughly three quarters of the mothers in group 1 were delivered within two days. The 44 women (23%) in group 2 who gave birth on day 7 reflect the effect of induction on the 52 (27%) who were still pregnant.

![Figure 2: Days from randomisation to delivery in the two groups.](image)

Counting from day 280, delivery took place at a mean of 14·8 days (SD 2·9; median 14·0) in group 1 and 17·6 days (SD 3·7; median 17·0) in group 2. Hence in group 2 pregnancy lasted a mean of 2·8 days longer. In group 2, 40 pregnancies continued longer than 300 days, as against 13 in group 1.

In group 1, 38 (18%) of the labours began spontaneously compared with 135 (69%) in group 2. The difference in mode of onset of labour did not influence the duration of labour (mean 7·5 (SD 5·5) hours in group 1, 7·4 (4·8) hours in group 2).

Table II lists the operative deliveries in the two groups. The difference in rates of caesarean section (6·5% vs 10·3%) was due to the five elective operations in group 2.

![Figure 3: Hour of delivery in the two groups.](image)

Table III—Vital measurements of infants

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Group 1 (n=214)</th>
<th>Group 2 (n=195)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (cm)</td>
<td>Mean 51·39</td>
<td>Mean 51·43</td>
</tr>
<tr>
<td></td>
<td>SD 2·04</td>
<td>SD 2·04</td>
</tr>
<tr>
<td>Weight (g)</td>
<td>3804</td>
<td>3856</td>
</tr>
<tr>
<td></td>
<td>449</td>
<td>502</td>
</tr>
<tr>
<td>Ponderal index*</td>
<td>28·01</td>
<td>29·28</td>
</tr>
<tr>
<td></td>
<td>2·50</td>
<td>2·41</td>
</tr>
</tbody>
</table>

*Ponderal index = (weight (g) × 1000)/length (cm)².

FETAL OUTCOME

Analysis of length, weight, and ponderal index of infants in the two groups showed no significant differences (table III). Infants in group 2 were a mean of 52 g heavier at birth, which presumably reflected the slightly longer gestational period.

An important question that can be examined only indirectly in a single hospital trial is whether delaying induction for one week jeopardises the fetus. Green amniotic fluid was noted in 37 cases (17%) in group 1 and 32 cases (16%) in group 2. There were no perinatal deaths. Apgar scores at one and five minutes were equally distributed. Twelve newborn infants in group 1 were transferred to the neonatal intensive care unit, where they stayed for an average of 4-3 days; in group 2, 15 were transferred and had a mean stay of 9-7 days. The difference in length of stay was due to one child in group 2 who stayed 93 days in the unit. The median stay was 3-0 days in each group.

Ten infants in group 1 needed phototherapy for hyperbilirubinaemia compared with one in group 2 (0·01 > p > 0·005).

USE OF RESOURCES

Both the number of hospital admissions and the number of days spent in hospital differed significantly between the two groups (mean SD). There were 40% fewer days in hospital: group 1, 7·05 (1·67); group 2, 6·69 (1·37) (p = 0·02). Women in group 1 had a mean of 1-09 courses of induction compared with 0·34 in group 2.

Figure 3 shows the times of delivery in the two groups. In both groups, but especially group 1, there was clustering from 1400 to 2300. A total of 1400 (65%) of the births in group 1 and 90 (46%) in group 2 occurred between 1400 and 2200.

**Discussion**

In this series postponing induction of labour until one week beyond the 42 week limit had no adverse effects on mothers and children. There was no evidence that the infants had suffered from lack of oxygen or nutrients during this period.

Gestational age specific perinatal mortality in Norway in 1981 was 2·5/1000 at weeks 40-41 (days 280-293), 3·8/1000 at weeks 42-43 (days 294-307), and 9·2/1000 at week 44 (day 308) and beyond.

Using these figures as an argument for or against post-term induction, however, would be fallacious; inductions may contribute to favourable post-term rates, but pregnancies with unreliable menstrual dating tend to distort the "true" picture.

Concern has been raised about the possible risks from the procedure of induction. Non-randomised comparisons of induced versus spontaneous labour have disclosed higher incidences of caesarean section among induced cases. Few prospective controlled studies of post-term induction have been published. In three instances randomised comparisons between induction at about 40 weeks and induction one to two weeks later were done. There were more forceps deliveries in the early induced group in one,
of pethidine among those induced early and more meconium stained amniotic fluid among those induced late in another, and no significant differences between the groups in the third. One post-term trial, very similar in design to ours, was reported while our original manuscript was being considered for publication. In that trial Cardozo et al found that more babies in the active management group were intubated and that the same group had slightly lowered cord blood pH, but otherwise there were no differences in outcome. In our trial there were no significant differences in any of the variables concerning labour and delivery. It is fair to conclude that either type of management appears to be safe, provided that certain conditions in case selection and supervision are met. An intelligent choice must therefore be based on other considerations.

The menstrual and pregnancy histories should be scrutinised to ensure that the pregnancy is truly past term. We excluded 373 (40%) out of 944 women from the trial because of uncertainty about dates. This proportion is not surprising, given that the proportion of pregnancies with uncertain dates increases with time past term. Furthermore, it is common experience that strict adherence to preset selection criteria in any randomised trial leads to many more exclusions than anticipated.

One feature of elective induction is often overlooked and therefore not universally acknowledged—namely, the comparatively high rates of unsuccessful attempts. In each of two studies of post-term induction delivery in 70% of cases occurred within 30 hours. In this trial the success rate was 77% (first attempt in group 1). Success is related to parity and the intensity and duration of oxytocin infusion. The chance of success must be weighed against the inconvenience and risk of prolonged contractions due to oxytocin. In principle the same argument is valid for treatment with prostaglandin. Restricting induction to mothers with rife cervixes (high Bishop scores) naturally increases the success rate. This group, however, has the highest chance of going into spontaneous labour soonest; hence if preventing the postmaturity syndrome is the aim these women may not need induction. The emotional response of some mothers in whom induction attempts fail also deserves consideration.

Oxytocin infusion has been incriminated in raised bilirubin concentrations in the newborn. The same tendency was found in this study, in which significantly more children in the induction group needed phototherapy. Though mildy raised bilirubin concentrations may not have serious implications, treatment is clearly demanding of resources.

Conclusion

With regard to safety the results do not warrant recommending one or other type of management. Nevertheless, we now postpone induction of labour in post-term cases, as the risk in monitoring the natural course, certainly up to day 308, seems minimal. Proper attention to menstrual history combined with ultrasonic and other information discloses some cases that are wrongly labelled as beyond-term. Resources spent on cardiotocography and ultrasonography of gradually dwindling numbers must be weighed against efforts at induction. Recent publications suggest that ultrasound assessment of amniotic fluid volume combined with non-stress tests are the best markers of fetal condition in post-term surveillance.13,14

We thank leading midwife Guro Kvamme and her staff for unfailing help and enthusiasm throughout the project.

References


(Accepted 5 March 1987)

SHORT REPORTS

Effect of negative ion generators in a sick building

Discomfort within office buildings is widespread and has been called the "sick building syndrome." A deficiency of negative ions has been suggested as the cause of the symptoms, though research has not shown a consistent improvement when the ion concentrations are increased. Data from Hawkins suggest that any effect of negative ions is long lasting, thus making a randomised crossover trial ineffective. We conducted a survey in a sick building whose occupants had a high prevalence of symptoms to test the effects of negative ion generators.

Subjects, methods, and results

Twenty six subjects in five rooms agreed to participate. In each room an EC 300 negative ion generator (Medion) was installed. The on/off light indicated that the machine was on throughout the study, but an internal switch activated the ionisers without the subjects knowing. The subjects completed a questionnaire daily for 12 weeks, using linear analogue scales to rate the environment and their personal comfort (see table). Certain specific symptoms were also recorded—for example, headaches, lethargy, nausea, dizziness, and nasal, eye, and throat symptoms.

After four weeks without ionisation the ionisers were activated in three of the rooms. After a further two weeks the ionisers in the other two rooms were activated. Once activated the ionisers remained on for the rest of the study. The air conditioning was not adjusted during the study. The negative ion concentrations in each study room and in a control room were measured every fortnight by an experienced operator with a 134A Atmospheric Ion Analyser (Medion). The temperature and relative humidity were measured each morning and afternoon with a wet and dry bulb hygrometer. A more detailed thermal survey of temperature, relative humidity, radiant heat, and air velocity was performed once in each study room and in the control room during each part of the study.

There were no significant changes between the different stages of the study in the variables in the control room or any of the study rooms (unpaired t test). The mean negative ion concentration was 139/ml before the study and 184/ml after activation (p<0.001). Because factors other than the negative ion concentration might have affected the linear analogue scale result each participant's ratings were adjusted by least squares multiple linear regression for the effects of temperature, relative humidity, radiated heat, and air velocity. These results suggest the presence of an effect of negative ions. The suitability of this linear model was confirmed by plotting residuals against predicted values of the linear analogue scale. The 95% confidence interval for these coefficients was calculated, every case including the null value of zero (table). Similar analyses were performed to determine the effects of temperature and relative humidity. These gave expected results: higher temperatures led to the environment being assessed as being hotter and more stuffy and people feeling hotter, while relative humidity, which varied from 31% to 54%, had no effect on comfort.

For specific symptoms the proportions of positive responses for the low and