Computerised interpretation of the fetal heart rate during labour: a randomised controlled trial (INFANT).
Brocklehurst, Peter; Kenyon, Sara

DOI:
10.1016/S0140-6736(17)30568-8

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial

The INFANT Collaborative Group

Summary
Background Continuous electronic fetal heart-rate monitoring is widely used during labour, and computerised interpretation could increase its usefulness. We aimed to establish whether the addition of decision-support software to assist in the interpretation of cardiotocographs affected the number of poor neonatal outcomes.

Methods In this unmasked randomised controlled trial, we recruited women in labour aged 16 years or older having continuous electronic fetal monitoring, with a singleton or twin pregnancy, and at 35 weeks' gestation or more at 24 maternity units in the UK and Ireland. They were randomly assigned (1:1) to decision support with the INFANT system or no decision support via a computer-generated stratified block randomisation schedule. The primary outcomes were poor neonatal outcome (intrapartum stillbirth or early neonatal death excluding lethal congenital anomalies, or neonatal encephalopathy, admission to the neonatal unit within 24 h for ≥48 h with evidence of feeding difficulties, respiratory illness, or encephalopathy with evidence of compromise at birth), and developmental assessment at age 2 years in a subset of surviving children. Analyses were done by intention to treat. This trial is completed and is registered with the ISRCTN Registry, number 98680152.

Findings Between Jan 6, 2010, and Aug 31, 2013, 47 062 women were randomly assigned (23 515 in the decision-support group and 23 547 in the no-decision-support group) and 46 042 were analysed (22 987 in the decision-support group and 23 055 in the no-decision-support group). We noted no difference in the incidence of poor neonatal outcome between the groups—172 (0·7%) babies in the decision-support group compared with 171 (0·7%) babies in the no-decision-support group (adjusted risk ratio 1·055 in the no-decision-support group). We noted no difference in the incidence of poor neonatal outcomes at age 2 years in a subset of surviving children (178 (0·82–1·25) vs 171 (0·76–1·25)). At 2 years, no significant differences were noted in terms of developmental assessment.

Interpretation Use of computerised interpretation of cardiotocographs in women who have continuous electronic fetal monitoring in labour does not improve clinical outcomes for mothers or babies.

Funding National Institute for Health Research.

Copyright © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY license.

Introduction Continuous electronic fetal heart rate monitoring in labour is widely used but its potential for improving neonatal outcomes has not been realised. The reasons for this are complex, but include difficulty interpreting the fetal heart rate trace correctly during labour. Computerised interpretation could be used to objectively detect abnormalities in fetal heart rate patterns during labour that are associated with asphyxia and bring them to the attention of birth attendants, who could then take action or expedite delivery and potentially prevent stillbirth or exposure to significant asphyxia.

Guardian (K2 Medical Systems, Plymouth, UK) is an electronic information capture system for managing information from labour monitoring. It displays the cardiotocograph on a computer screen alongside other clinical data either from external ultrasonographic transducers, or from fetal scalp electrodes (eg, partographs, maternal vital signs, details of maternal anaesthesia and analgesia) collected as part of routine clinical care. Guardian does not interpret any of the data gathered, but acts as an interface to collect and display data at the bedside, centrally on the labour ward, in consultants' offices, or remotely.

INFANT (K2 Medical Systems) is a decision-support software that was developed to run on the Guardian system. It analyses the quality of fetal heart signals and, if these signals are adequate, displays baseline heart rate; heart-rate variability; accelerations and type and timing of decelerations; the quality of the signal; and the contraction pattern. INFANT then makes an assessment of the overall pattern, which, if necessary, will result in a colour-coded alert (blue is the least severe, yellow is moderate severity, and red is the most severe; appendix). The decision-support software does not provide recommendations for any action that should be taken in response to abnormalities—such decisions are at the discretion of attending clinicians. In the UK’s National Health Service (NHS), all clinicians supervising labour are expected to have been trained in the appropriate response to an abnormal cardiotocographic reading—eg, by completing computer-based training packages every 6–12 months, attending annual lectures, or attending regular cardiotocographic review meetings.
However, training is not standardised, and individual competence is not assessed in most units.9

We did a randomised controlled trial to test the hypotheses that a substantial proportion of substandard care results from failure to correctly identify abnormal fetal heart rate patterns, that improved recognition of abnormality would reduce substandard care and poor outcomes, and that improved recognition of normality would decrease unnecessary intervention.

Methods
Study design and participants
We did a pragmatic, unblinded, randomised controlled trial in maternity units in the UK and Ireland. All 24 sites in the UK and Ireland that used Guardian at the time of the trial took part. Doctors and midwives were able to recruit participants. Eligible women were judged to require continuous electronic fetal heart rate monitoring by the local clinical team on the basis of their existing practice, had a singleton or twin pregnancy, were at 35 weeks’ gestation or further along, had no known gross fetal abnormality, including any known fetal heart arrhythmia such as heart block, aged 16 years or older, and able to give consent to participate as judged by the attending clinicians. Continuous electronic fetal heart rate monitoring during labour is not routine in the UK. Clinical guidance for the NHS recommends that women assessed as having a low risk of complications should be offered intermittent auscultation during labour.9

Pregnant women attending recruiting hospitals were provided with written information about the trial during pregnancy and in labour. For women who met the eligibility criteria, written informed consent was sought by means of a dated signature from the woman and from the person who obtained informed consent. Research ethics committee approval for the study was granted by the National Research Ethics Service—Northern and Yorkshire Research Ethics Committee (09/H0903/31). The study protocol has been published.11

Randomisation and masking
The Guardian system was used to confirm that all necessary eligibility criteria were met and then to randomly allocate women (1:1) to cardiotocography with or without decision support. The allocations were computer generated in Stata (version 10.1) by the trial statistician, who used stratified block randomisation, in which variable block sizes were used to balance between the two trial arms by whether the pregnancy was a singleton (block sizes 12, 14, 16, 18, 20, 22, and 24, allocated in proportion to the elements of Pascal’s triangle—1:6:15:20:15:6:1) or twin (block sizes 2, 4, and 6, allocated in proportion to the elements of Pascal’s triangle 1:2:1), and within each participating centre. The trial was not blinded, which allowed indirect measurement of any changes in clinician behaviour, such as how much time the attending midwife spent with the woman on the basis of the knowledge that the decision-support system was active or not.

Procedures
Clinicians in participating centres were trained in the use of the decision-support software by staff from the trial office. A training team at each site was responsible for cascading training among the local clinicians. Women in the trial were managed according to standard procedures. No additional training was provided to clinical staff in how to respond to fetal heart rate abnormalities.

Labour data and outcomes were stored contemporaneously on the Guardian system, which were then sent electronically to the trial office. Data were extracted from notes of babies admitted to the neonatal unit and for all neonatal deaths. All children surviving were
flagged at the NHS Information Centre for those born in England and NHS Greater Glasgow and Clyde Safe Haven for those born in Scotland, which meant that all deaths occurring after discharge in these countries could be identified. Ireland does not have a similar system for monitoring deaths, so Irish data had to be excluded from the denominator for these calculations. A sample of surviving children were followed up to age 2 years via a parent-completed questionnaire to assess the child’s health, development, and wellbeing (appendix).

All babies with an adverse outcome potentially associated with intrapartum asphyxia (ie, the trial primary outcome plus cord artery pH < 7·05 with base deficit 12 mmol/L or more) and all neonatal deaths and intrapartum stillbirths had their care in labour assessed by review of de-identified case notes by a panel comprising a senior obstetrician, neonatologist, and midwife, to see if care was suboptimal—ie, if it was possible or probable that different management would have prevented the adverse outcome.1,2

Outcomes
The trial had two primary outcomes. The first was a binary (present or not present) composite of poor neonatal outcome, including deaths (intrapartum stillbirths plus neonatal deaths up to 28 days after birth, except for congenital anomalies deaths) and significant morbidity (moderate or severe neonatal encephalopathy, defined as the use of whole-body cooling or admission to the neonatal unit within 48 h of birth for 48 h or more with evidence of feeding difficulties or respiratory illness, with evidence of compromise at birth suggesting mild asphyxia or mild encephalopathy, or both). The second was a continuous outcome of developmental progress measured by the Parent Report of Children’s Abilities—Revised (PARCA-R) composite score3,4 at age 2 years for a subset of children.

Infant secondary outcomes, all of which are binary unless specified, were intrapartum stillbirth (excluding deaths from congenital anomalies); neonatal deaths up to 28 days after birth (excluding deaths from congenital anomalies); moderate or severe encephalopathy; admission to the neonatal unit within 48 h of birth for 48 h or more with evidence of feeding difficulties or respiratory illness, with evidence of compromise at birth suggesting mild asphyxia or mild encephalopathy, or both). The second was a continuous outcome of developmental progress measured by the Parent Report of Children’s Abilities—Revised (PARCA-R) composite score3,4 at age 2 years for a subset of children.

Statistical analysis
A sample size of 46 000 births was needed.1 We postulated an incidence of the primary outcome of three per 1000 births by summing the previous reported rates of intrapartum stillbirth, neonatal death, moderate and severe encephalopathy, and mild encephalopathy (reliable data for significant asphyxial morbidity were not available and so could only be estimated).16–18 The effect size that could be detected with 46 000 women (23 000 in each group), assuming a 5% level of significance and 90% power, was a 50% reduction in poor neonatal outcomes from three to 1·5 per 1000. In a study of preterm infants,14 the mean PARCA-R composite score at 2 years was 80 (SD 33) and the mean Mental Development Index (Bayley Scales of Infant Development II) was around half an SD below the mean of 100. Assuming that a healthy group of term infants would have a PARCA-R composite score half an SD higher than this sample of preterm infants, we estimated a mean 2 year score of 96 (SD 33). A follow-up sample of size 7000 (3500 per arm) had more than 90% power to detect a difference of 3 points in the PARCA-R component score with a two-sided 5% significance level. The incidence of severe metabolic acidosis (cord artery pH < 7·05) is ten per 1000.19–22 A sample of 46 000 women enabled us to detect a 28% relative risk reduction in this incidence with more than 80% power, assuming a 5% level of significance, in babies in whom cord artery pH was measured.
During the early part of the trial, and with advice from the data-monitoring committee, the primary outcome definition was refined to ensure that it captured babies who were likely to have experienced hypoxia during labour. The original component of the primary outcome—admission to neonatal unit within 48 h of birth for 48 h or more with evidence of feeding difficulties, respiratory illness, or encephalopathy—resulted in the inclusion of many babies with a range of disorders, many of which were unlikely to be related to hypoxia. Each case fulfilling this component of the primary outcome was reviewed by an independent panel of neonatologists (who were blinded to allocation) to ascribe it as fulfilling the revised definition or not (appendix).

The trial steering committee approved the statistical analysis plan before the analysis (appendix). For the main comparative analysis, participants were analysed in the groups into which they were randomly allocated, irrespective of allocation received. All women and babies with available data were included, except for women for whom a valid signed consent form could not be located or women who withdrew consent. The numbers and percentages of babies in whom the primary outcomes were noted are for each group, and the risk ratios plus 95% CIs were calculated. Risk ratios were estimated with generalised linear models with a binomial distribution and a log link (or a Poisson distribution with a log link if convergence could not be achieved). Hazard ratios were estimated with Cox regression and rate ratios with Poisson regression. We adjusted for the stratification factors used in the randomisation (centre and singleton or twin pregnancy), and used robust variance estimators in all models to account for the correlation in outcomes between twins and siblings delivered in a subsequent...
pregnancy during the trial period. The mean (SD) PARCA-R composite score was calculated for each group, and the mean difference between groups plus 95% CI were calculated and compared with a Gaussian model with identity link. For secondary outcomes including the components of the primary outcome, a 1% level of significance was employed.

We did prespecified subgroup analyses with the statistical test for interaction for singletons versus twins, suspected fetal growth restriction at labour onset versus no growth restriction, body-mass index group, and centre. These analyses were done for the trial primary outcomes, all neonatal outcomes, instrumental vaginal deliveries, and caesarean section. Additionally, we did a subgroup analysis of all process outcomes by centre. Major disability at 2 years was classified in terms of neuromotor function, seizures, auditory function, communication, visual function, cognitive function, and other physical disability.

We used Stata/SE for Windows (version 13.1) for all analyses. The trial was overseen by an independent trial steering committee and an independent data-monitoring committee. The data-monitoring committee used the Haybittle-Peto approach for interim analyses, with three SEs as the cutoff for consideration of early cessation, preserving the type-1 error across the trial. The trial is registered with ISRCTN, number 98680152.

Role of the funding source
The funder of the study had no role in the trial design; data collection, analysis, or interpretation; or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
Between Jan 6, 2010, and Aug 31, 2013, 47 062 women were recruited to the INFANT trial (appendix). 1020 women (2.2%) were excluded from the analysis of
the primary outcome (figure; appendix), mostly because of missing or incomplete consent forms. Data at the time of birth were available for 100% of women and babies eligible to be analysed. Follow-up data at 2 years were available for 7066 of those contacted; data were sufficiently complete for 6707 children.

Baseline characteristics were similar between the two groups (table I). Median maternal age was 29 years.
(IQR 25–33). Around 60% of women were having their first baby, and most women had a gestational age between 38 and 41 completed weeks (table 1). Very few women had a previous stillbirth (1%) and around 6% had previously had a previous caesarean section (table 1). Almost 60% of women had their labour induced.

The incidence of the primary outcome—poor neonatal outcome—did not differ significantly between the groups...
Articles

<table>
<thead>
<tr>
<th>Decision support (n=22517)</th>
<th>No decision support* (n=22564)</th>
<th>Adjusted effect measure (99% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epidural analgesia</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2770 (27.3%)</td>
<td>2689 (26.5%)</td>
</tr>
<tr>
<td>No</td>
<td>7383 (72.7%)</td>
<td>7453 (73.5%)</td>
</tr>
<tr>
<td>Unknown†</td>
<td>12364</td>
<td>12422</td>
</tr>
<tr>
<td><strong>Labour augmentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2705 (30.9%)</td>
<td>2750 (31.3%)</td>
</tr>
<tr>
<td>No</td>
<td>6047 (69.1%)</td>
<td>6042 (68.7%)</td>
</tr>
<tr>
<td>Unknown†</td>
<td>13765</td>
<td>13772</td>
</tr>
<tr>
<td><strong>Presence of meconium</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>440 (4.5%)</td>
<td>469 (4.8%)</td>
</tr>
<tr>
<td>No</td>
<td>9316 (95.5%)</td>
<td>9346 (95.2%)</td>
</tr>
<tr>
<td>Unknown†</td>
<td>12761</td>
<td>12749</td>
</tr>
<tr>
<td>At least one blue, yellow, or red level of concern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21950 (97.3%)</td>
<td>22021 (97.6%)</td>
<td>Risk ratio 1.00 (99% CI 1.00-1.00)</td>
</tr>
<tr>
<td>At least one blue level of concern (mild abnormality)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21863 (97.1%)</td>
<td>21913 (97.1%)</td>
<td>Risk ratio 1.00 (99% CI 1.00-1.00)</td>
</tr>
<tr>
<td>At least one yellow level of concern (moderate abnormality)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16765 (74.5%)</td>
<td>16722 (74.4%)</td>
<td>Risk ratio 1.00 (99% CI 0.99-1.02)</td>
</tr>
<tr>
<td>At least one red level of concern (severe abnormality)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2335 (10.8%)</td>
<td>2413 (11.1%)</td>
<td>Risk ratio 0.97 (99% CI 0.90-1.04)</td>
</tr>
<tr>
<td>Unknown†</td>
<td>822</td>
<td>833</td>
</tr>
<tr>
<td><strong>Blue, yellow, or red levels of concern in women with at least one level of concern</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>9 (5-15)</td>
<td>9 (5-15)</td>
</tr>
<tr>
<td>Rate (per h)</td>
<td>1.37</td>
<td>1.40</td>
</tr>
<tr>
<td>Unknown§</td>
<td>765</td>
<td>824</td>
</tr>
<tr>
<td><strong>Blue levels of concern in women with a blue level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7 (4-11)</td>
<td>7 (4-11)</td>
</tr>
<tr>
<td>Rate (per h)</td>
<td>1.06</td>
<td>1.05</td>
</tr>
<tr>
<td>Unknown§</td>
<td>740</td>
<td>800</td>
</tr>
<tr>
<td><strong>Yellow levels of concern in women with a yellow level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>2 (1-4)</td>
<td>2 (1-5)</td>
</tr>
<tr>
<td>Rate (per h)</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>Unknown§</td>
<td>354</td>
<td>421</td>
</tr>
<tr>
<td><strong>Red levels of concern in women with a red level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>1 (1-1)</td>
<td>1 (1-1)</td>
</tr>
<tr>
<td>Rate (per h)</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Unknown†§</td>
<td>41</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 3 continues on next page

Just over half of all births were spontaneous vaginal births and the frequency did not differ significantly between the two groups (adjusted RR 0.99, 99% CI 0.97–1.01; table 2). More women underwent fetal blood sampling in the decision-support group than in the no-decision-support group (2366 [10.3%] vs 2187 [9.5%]; adjusted RR 1.08, 99% CI 1.01–1.16). No other significant differences were noted between the two groups from trial entry to birth in terms of clinical outcomes (table 2).

For babies with an adverse outcome and cord metabolic acidosis who underwent expert review, the overall proportion of babies judged to have received suboptimal care likely to have affected the outcome was 38%—14 of 35 babies in the decision-support group and 13 of 36 babies in the control group—which is similar to that reported previously. We could not investigate whether in all cases not reviewed appropriate action was taken in response to recognised abnormality.

In women with any level of concern as measured by INFANT (table 3), blue levels of concern were most frequent (median seven alerts during labour—roughly 1.1 per h), followed by yellow alerts (median two alerts per labour), and then red alerts (median one per labour). A lower rate of yellow levels of concern was noted in the decision-support group compared with the no-decision-support group (adjusted rate ratio 0.87, 99% CI 0.84–0.89; table 3). Frequency of blue and red alerts did not differ significantly between groups (table 3).

Although there was a worry that women in the decision-support group would be left alone more frequently during labour than those in the no-decision-support group, the frequency of thumbprint entries on the Guardian did not differ significantly between groups (4.22 per h vs 4.21; adjusted rate ratio 0.99, 99% CI 0.95–1.03; table 3).

Time from the last red level of concern to birth was similar in both groups (median 58 min; table 3). In a subgroup of 473 readable traces from a sample of 500 taken as a similar number of consecutive cases from each contributing centre, the last red level of concern was judged (by expert investigator, PS) to be a valid fetal concern for 276 (58%) traces. Maternal heart rate triggered the red level of concern in 128 (27%) cases, misclassified accelerations in 36 (8%) cases, and other reasons in 33 (7%) cases.

Families were contacted when their surviving child or children born reached age 2 years. Nearly 7000 families returned a questionnaire. The characteristics of the mothers who responded differed significantly from those of the entire trial cohort and from those of mothers who did not respond (appendix). Compared with the entire
trial cohort, responders were more likely to be white, to have given birth at a later gestational age, and to have been having their first baby (appendix).

Of the 7066 infants for whom a questionnaire was returned, data could be analysed for 6707 (95%). We noted no significant differences between the two groups for any of the 2 year outcomes, including the primary outcome, PARCA-R score (table 4). Nearly 6% of children for whom data were available had a major disability. The classification of disability used meant that large numbers of children were assigned a major disability as a result of having poor growth (between 2·8% and 3·0% of all children) and cognitive difficulties (between 1·2% and 1·5% of all children). Other major disabilities such as physical disability, blindness, and deafness were all very uncommon (appendix).

We noted no evidence that the decision-support software performed significantly differently between any of the pre-specified subgroups for either the primary outcome or a range of secondary outcomes (appendix). Furthermore, no differences were noted in the distribution of cord blood pH measurements (appendix). The number of alerts differed significantly by centre, but no other significant differences were noted by centre (appendix).

Discussion
In this trial of more than 46 000 women, we found no evidence that the use of decision-support software in conjunction with cardiotocography reduced the likelihood of poor neonatal outcomes compared with cardiotocography alone.

In another randomised trial, which also recruited in the UK, the use of decision support was assessed in women monitored during labour with fetal electrocardiographic monitoring. This study also showed no evidence that decision support improved the primary outcome of cord blood metabolic acidosis in 7730 women. The results of one small trial of 220 women in Bulgaria have suggested that decision support is associated with benefits with respect to cord blood metabolic acidosis. In the UK continuous electronic fetal heart-rate monitoring is not routine, making generalisability to settings in which it is routine less certain.

The strengths of this study are its contemporaneous data collection and size. Potential weaknesses include the challenges of use of a composite primary outcome measure, the potential for staff to learn from exposure to the decision-support arm of the trial, resulting in improved outcomes in the control arm, and the issue of accounting for multiple comparisons.

Use of a composite primary outcome might not always be helpful if different components of the outcome respond differently to the intervention. We initially hypothesised that the components of the outcome would have similar incidences, with each component likely to contribute around a third to the composite. Estimates of the incidence of these components for eligible women were difficult to find before the trial began. The perinatal mortality in our study (13 per 46 614 babies [0·3 per 1000]) was lower than the previous estimate (1·05 per 1000), and the incidence of neonatal encephalopathy requiring cooling was also lower than previous estimates (0·8 per 1000 vs 1·3 per 1000). However, the incidence of prolonged admission to neonatal units with evidence of compromise at birth, for which we had no good data when planning the trial, occurred more frequently (291 per 46 614 [six per 1000]) and contributed substantially more to the higher-than-anticipated overall primary event rate of seven per 1000 compared with our estimated three per 1000. This frequency afforded us power to detect smaller differences in the composite outcome than we had originally planned.

The potential weakness of staff learning from exposure to the decision-support system was identified when planning the trial. We acknowledged that passive learning from the decision-support system might be possible. However, part of our previous hypothesis was that, although some poor cardiotocographic interpretation is due to a lack of training, some clinicians have poor intrinsic pattern-recognition abilities, which, by definition, would not be affected by training, and the performance of such clinicians would be particularly improved by assistance from automatic interpretation. We collected a range of process outcomes to measure the impact on clinician behaviour during the trial, and these data suggested some evidence for behaviour change in the decision-support arm: fetal blood sampling was more frequent and the incidence of repeated yellow
alerts lower than in the control group. Perhaps different action was taken in response to the alerts in the decision-support arm of the trial—eg, clinicians might have reduced the dose of an oxytocin infusion in women having their labour augmented if the infusion was leading to very frequent contractions. Such actions could have prevented further yellow alerts, leading to a decrease in this group, although we do not have any direct evidence for this scenario. Even if this effect did occur, it did not result in any significant change in clinical outcomes. Although the median time from last level of red concern to birth might have increased, this was not significant and no major differences were seen in important clinical outcomes, such as duration of labour, the rate of labour progress, and clustering of meconium in the amniotic fluid. The only two significant findings relate to behaviour change in clinicians favouring decision support. These findings were in the expected direction of effect and are mutually supportive, suggestive of a real effect.

In this trial, the only two significant findings relate to behaviour change in clinicians favouring decision support. These findings were in the expected direction of effect and are mutually supportive, suggestive of a real effect.

**Detection of abnormalities in the fetal heart rate can improve outcome only if caregivers respond appropriately to the alerts.** A review of all severe adverse outcomes in the trial showed no evidence of differences in suboptimal care between the two groups. Therefore our hypothesis that substandard care is largely related to failure to identify pathological fetal heart-rate patterns is not supported. Most adverse outcomes associated with preventable substandard care seemed to involve failure to take appropriate management decisions once the cardiotocographic abnormality had been recognised. This aspect will be reported in detail in a follow-up paper. Our hypothesis that unnecessary intervention would be reduced was also not supported.

The decision-support software used in this trial identifies fetal heart-rate abnormalities. However, the alerts do not take into account other information about the labour, such as duration of labour, the rate of labour progress, and presence of meconium, all of which could modify the way a clinician interprets the fetal heart rate and acts on this information. Further development of decision-support software could improve the quality of feedback that the system provides to clinicians to make a difference to outcomes. In view of the importance for parents, clinicians, and health services of the consequences of intrapartum hypoxia, identification of signs of early compromise during labour so that timely intervention can be used to reduce poor outcomes is an urgent unmet need.

**Contributors**

All members of the writing committee contributed to the development of the protocol and management and running of the trial. LL did the analyses. PB wrote the Article, and revised it with input from all authors. All authors read and approved the final Article.

**Writing committee**

Peter Brocklehurst, David Field, Keith Greene, Ed Juszczak, Sara Kenyon, Louise Linsell, Christopher Mabey, Mary Newburn, Rachel Plachcinski, Maria Quigley, Elizabeth Schroeder, Philip Steer.

**Co-investigator group**

Peter Brocklehurst, David Field, Keith Greene, Ed Juszczak, Robert Keith, Sara Kenyon, Louise Linsell, Christopher Mabey, Mary Newburn, Rachel Plachcinski, Maria Quigley, Elizabeth Schroeder, Philip Steer.

**Recruiting centres**

Birmingham Women’s Hospital (Nina Johns, Tracey Johnston, Gemma Barnfield, Karen Davies), Chelsea and Westminster Hospital (Mark Johnson, Holly Patterson), Derriford Hospital, Plymouth (Imogen Montague, Sally Watmore, Alison Stolton), Homerton University Hospital (Maryam Parisaei, Natasha McGhee, Silvia Segovia), Lancashire Women & Newborn Centre (Elizabeth Martindale, Hilary Jackson, Josephine Holleran), Liverpool Women’s Hospital (Devender Roberts, Siobhan Holt), Northwick Park Hospital

---

**Table 4: Health and development outcomes at 2 years in a sample of surviving infants without the primary outcome selected for follow-up.**

<table>
<thead>
<tr>
<th>Decision support (n=3556)</th>
<th>No decision support (n=3510)</th>
<th>Adjusted effect measure (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARCA-R composite score</td>
<td>Mean (SD) 98 (33.8)</td>
<td>97.2 (33.4)</td>
</tr>
<tr>
<td>Median (IQR) 98 (73-126)</td>
<td>97 (72-125)</td>
<td>17</td>
</tr>
<tr>
<td>Unknown</td>
<td>17</td>
<td>184</td>
</tr>
</tbody>
</table>

**PARCA-R non-verbal cognition scale**

| Mean (SD) 27.7 (3.7) | 28.0 (3.6) | Mean difference 0.22 (99% CI 0.4 to 0.01) |
| Median (IQR) 28 (26-30) | 28 (26-31) |

**PARCA-R vocabulary sub-scale**

| Mean (SD) 57.4 (27.8) | 56.5 (27.7) | Mean difference 0.82 (99% CI 0.1 to 1.54) |
| Median (IQR) 58 (36-81) | 56 (35-80) |

**PARCA-R sentence complexity sub-scale**

| Mean (SD) 12.4 (5.4) | 12.3 (5.3) | Mean difference 0.07 (99% CI 0.1 to 0.6) |
| Median (IQR) 12 (9-16) | 12 (9-16) |

**Cerebral palsy**

| n (%) 4 (0.1%) | 4 (0.1%) | RR 0.99 (99% CI 0.16 to 6.10) |
| Unknown (n) 111 | 114 |

**Non-major or major disability**

| n (%) 942 (40.4%) | 840 (37.4%) | RR 1.08 (99% CI 0.98 to 1.18) |
| Unknown (n) 1225 | 1266 |

**Major disability**

| n (%) 134 (5.7%) | 135 (6.0%) | RR 0.95 (99% CI 0.70 to 1.29) |
| Unknown (n) 1225 | 1266 |

(Missing data are <3% unless otherwise presented; there were no apparent differences in missing data between trial arms. Effect measures were adjusted for stratification factors used in the randomisation (centre and twin birth) and clustering as a result of twins and multiple birth episodes. Stratification factors were not adjusted for in the analysis of infant deaths and cerebral palsy because of the small number of events. Crude effect measures were not presented as identical to one decimal place (two decimal places for most outcomes). At 2 years, 70 (13%) of 5208 infants had died in the decision-support group and 35 (0.16%) of 21597 infants had died in the no-decision-support group (adjusted RR 0.81 [99% CI 0.44-1.59]). All deaths were reported to age 2 years excluding stillbirths (one in the decision-support group and two in the no-decision-support group). Data from Ireland were not included in the numerator (n=1 in the decision-support group) or denominator (n=1754 in decision-support group and n=1752 in the no-decision-support group) because data for deaths after discharge were not available. CI-confidence interval. PARCA-R-Parent Report of Children’s Abilities—Revised. RR-risk ratio. *Disability in any of neuromotor function, seizures, auditory function, communication, visual function, cognitive function, or other physical disability.
Bosko Dragovic, Miriam Willmott-Powell, Laura Hutchinson, Benedek Toth, Gemma Chandler, Suzanne Ridley), Nottingham City Hospital (George Bugg, Anna Molnar, Denise Lochrie), Princess Anne Hospital (Jillian Connor, David Howe, Katie Head, Sue Wellstead), Princess Royal Hospital (Alan Mathers, Laura Walker, Isobel Crawford), Queen Alexandra Hospital (David Davies, Zoe Garner, Lucy Galloway), Queen’s Medical Centre (George Bugg, Yvette Davies, Carys Smith, Gill Perkins), Rotunda Hospital (Mike Geary, Fiona Walsh, Ursula Nagle), Royal Blackburn Hospital (Elizabeth Martindale, Hilary Jackson, Louise O’Malley), Royal Bolton Hospital (Narmada Katakam, Heather White, Emma Tanton), Royal Derby Hospital (Rosie Hamilton, Hilary Glanowski, Ethel Forde), Southern General Hospital (Alan Mathers, Christina MacDonald, Lorna McKay), St Mary’s Hospital (Lenny Edizoom, Paula Doran, Julie Dillon, Cara Taylor, Paula Evans), Stoke Mandeville Hospital (Veronica Miller, Christopher Wayne, Julie Tebbutt, Ellie Hendy), University College London Hospital (Patrick O’Brien, Seni Subair, Helen Dent, Camille Mallet), University Hospital of Coventry and Warwickshire (Sidohan Quenby, Jane Hillen), University Hospital of North Staffordshire (Peter Young, Tracey Harrison, Louise Wood), Warrington Hospital (Rita Arya, Lindsay Roughley), Warwick Hospital (Olanrewaju Sorinola, Carole Rogers, Janet Phipps).

Trials steering committee
Bob Arndtz, Denis Azzopardi, Zoe Chivers, Andy Cole, Max Parmar (chair until November 2015), Tracey Roberts, Julia Sanders, Derek Tuffnell (chair from November 2015).

Data-monitoring committee
Deborah Ashby (chair), Jane Norman, Andy Shennan, Helen Spiby, Win Tin.

Declaration of interests
KG is a founding shareholder of K2 Medical Systems, and was clinical director for development of the INFANT and Guardian systems until 2005. CM is an employee and shareholder of K2 Medical Systems. All other authors declare no competing interests.

Acknowledgments
This study was funded by the UK National Institute for Health Research (NIHR) Health Technology Assessment programme (project number 06/38/03). SK was partly funded by the NIHR Collaboration for Leadership in Applied Health Research and Care West Midlands. The quality-of-care review panel consisted of an obstetrician (PS), neonatologist (Ilya Koval), and midwife (Christina McKenzie) for all sessions except one, when CM was replaced by Mary Griffin. Previous funding from the UK Medical Research Council enabled the initial development of Guardian and the INFANT software. The views and opinions expressed by authors in this publication are those of the authors and do not reflect those of the UK NHS, NIHR, the UK Medical Research Council, Central Commissioning Facility, NIHR Evaluation, Trials and Studies Coordinating Centre, the Health Technology Assessment programme, or the Department of Health.

References