Computerised interpretation of the fetal heart rate during labour: a randomised controlled trial (INFANT)

Peter Brocklehurst,[1]* David Field,[2] Keith Greene,[3] Edmund Juszczak,[4] Sara Kenyon,[5] Louise Linsell,[4] Chris Mabey,[6] Mary Newburn,[7] Rachel Plachcinski,[8] Maria Quigley,[4] Philip Steer,[9] Liz Schroeder[10] and Oliver Rivero-Arias[4]

[1]Birmingham Clinical Trials Unit, University of Birmingham, Birmingham, UK
[2]Department of Health Sciences, University of Leicester, Leicester, UK
[3]University College London, London, UK
[4]National Perinatal Epidemiology Unit Clinical Trials Unit (NPEU CTU), Nuffield Department of Population Health, University of Oxford, Oxford, UK
[5]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[6]K2 Medical Systems, Plymouth, UK
[7]Collaboration for Leadership in Applied Health Research and Care (CLAHRC) South London, King's College London, London, UK
[8]National Childbirth Trust, London, UK
[9]Imperial College London, London, UK
[10]Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW, Australia

*Corresponding author p.brocklehurst@Bham.ac.uk

Declared competing interests of authors: Keith Greene is the founder and shareholder of K2 Medical Systems (Plymouth, UK) and Clinical Director for the development of the INFANT system. Christopher Mabey is employed by, and is a shareholder of, K2 Medical Systems, the technology provider for the study. Edmund Juszczak reports grants from the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) and Efficacy and Mechanism Evaluation programmes during the conduct of the study, and is a member of the NIHR HTA Commissioning Board. Peter Brocklehurst reports grants and personal fees from the Medical Research Council and grants from the National Institute for Health and Care Excellence, NIHR Health Services and Delivery Research, NIHR HTA and Wellcome Trust, outside the submitted work, and is chairperson of the NIHR HTA Women and Children's Health panel and is a member of the HTA Prioritisation Group. Sara Kenyon is a member of the NIHR HTA Women and Children's Health panel and received NIHR funding to undertake the HOLDS (High Or Low Dose Syntocinon® for delay in labour) trial, and was part funded by the NIHR Collaboration for Leadership in Applied Health Research and Care West Midlands.

Published February 2018
DOI: 10.3310/hta22090

Scientific summary

The INFANT RCT

Health Technology Assessment 2018; Vol. 22: No. 9
DOI: 10.3310/hta22090

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Continuous electronic fetal monitoring (EFM) in labour is widely used and has the potential to improve neonatal outcomes. The benefits of EFM have been limited so far. The reasons for this appear to be complex, but include difficulties with interpreting the fetal heart rate trace correctly during labour. Computerised interpretation of the fetal heart rate has the potential to objectively detect abnormalities of the fetal heart rate pattern that are associated with asphyxia but not recognised as abnormal by the birth attendants, bringing to their attention the need to act to prevent stillbirth or exposure to significant asphyxia.

The electronic information capture system used in both arms of the trial, Guardian® (version 2.050.038.001, K2 Medical Systems, Plymouth, UK), is a system for managing information from labour monitoring. It displays the cardiotocograph (CTG) on a computer screen alongside other clinical data (e.g. the partogram and maternal vital signs) that are collected as part of routine clinical care. It can display CTG data obtained from external ultrasound transducers or from fetal scalp electrodes. Guardian acts as an interface to collect and display data at the bedside, centrally on the labour ward, in consultants' offices or remotely. The decision support software (INFANT®; version 2.050.035.001, K2 Medical Systems, Plymouth, UK) analyses the quality of the fetal heart signals and, if adequate, displays baseline fetal heart rate, heart rate variability, accelerations, type and timing of decelerations, the quality of the signal and the contraction pattern. It then makes an assessment of the overall pattern, which can result in a colour-coded alert, depending on the severity of any abnormality detected. The decision support software does not provide recommendations for any action that should be taken in response to these abnormalities. This was left to the discretion of the attending clinicians.

Objectives

Our hypotheses were that:

- A substantial proportion of substandard care results from a failure to correctly identify abnormal fetal heart rate patterns.
- Improved recognition of abnormality would reduce substandard care and poor outcomes.
- Improved recognition of normality would reduce unnecessary intervention.

The aim of the INFANT trial was to determine whether or not the addition of computerised interpretation of the intrapartum CTG to current clinical care could improve the management of labour for women who were judged to require EFM, and also to determine whether or not the use of the decision support software is cost-effective.

Methods

This was a two-arm pragmatic individually randomised controlled trial in labour wards in England, Scotland and the Republic of Ireland.

Eligibility

Eligible women were those:

1. who were judged to require EFM by the local clinical team based on their existing practice, and who had consented to EFM
2. who were pregnant with a single fetus or twin fetuses
3. who were at ≥ 35 weeks' gestation
4. had a fetus with no known major abnormality, including any known heart arrhythmia such as heart block
5. who were aged ≥ 16 years
6. who were able to give consent to participate as judged by the attending clinicians.

Randomisation

Women were allocated, in the ratio of 1 : 1, to (1) CTG monitoring with decision support or (2) CTG monitoring with no decision support. The allocations were computer-generated using stratified block randomisation employing variable block sizes to balance between the two trial arms by singleton or twin pregnancy, and within each participating centre. The trial was not masked.

Primary outcome measures

Primary short-term outcome

A composite of poor neonatal outcome including deaths [intrapartum stillbirths plus neonatal deaths (i.e. deaths up to 28 days after birth) except deaths as a result of congenital anomalies], significant morbidity [moderate or severe neonatal encephalopathy (NNE), defined as the use of whole-body cooling] and admissions to a neonatal unit within 48 hours of birth for ≥ 48 hours (with evidence of feeding difficulties or respiratory illness and when there was evidence of compromise at birth suggesting that the condition was the result of mild asphyxia and/or mild NNE).

Primary long-term outcome

Developmental progress as measured by the Parent Report of Children's Abilities-Revised (PARCA-R) composite score at the age of 2 years for a subset of children.

Secondary outcome measures

Secondary short-term outcomes

Neonatal

- Intrapartum stillbirth (excluding deaths as a result of congenital anomalies).
- Neonatal deaths up to 28 days after birth (excluding deaths as a result of congenital anomalies).
- Moderate or severe NNE.
- Admission to neonatal unit within 48 hours of birth for ≥ 48 hours with evidence of feeding difficulties, respiratory illness or NNE (when there was evidence of compromise at birth).
- Admission to a higher level of care.
- An Apgar score of < 4 at 5 minutes after birth.
- The distribution of cord blood gas data for cord artery pH.
- Metabolic acidosis (defined as a cord artery pH of < 7.05 and a base deficit in extracellular fluid of ≥ 12 mmol/l).
- Resuscitation interventions.
- Seizures.
- Destination immediately after birth.
- Length of hospital stay.

Maternal

- Mode of delivery.
- Operative intervention (caesarean section or instrumental delivery) for:
  - fetal indication or
  - failure to progress or
  - a combination of fetal distress and failure to progress or
  - other reason.
- Grade of caesarean section.
- Episiotomy.
- Any episode of fetal blood sampling.
- Length of:
  - first stage of labour from trial entry
  - second stage of labour from trial entry
  - labour from trial entry (total).
- Destination immediately after birth.
- Admission to a higher level of care.

Secondary long-term outcomes (infant)

Health and development outcomes at 24 months

- Non-verbal cognition scale (PARCA-R).
- Vocabulary subscale (PARCA-R).
- Sentence complexity subscale (PARCA-R).
- Late deaths up to 24 months (after the neonatal period).

Diagnosed with cerebral palsy:

- Non-major disability.
- Major disability.
- Breastfeeding (collected at 12 and 24 months).

Quality-of-care outcomes

In the case of any adverse infant outcomes potentially associated with intrapartum asphyxia (trial primary outcome based on the baby's condition after birth, plus a cord artery pH of < 7.05 with a base deficit of ≥ 12 mmol/l) and all neonatal deaths and intrapartum stillbirths, care during labour was assessed by a panel comprising a senior obstetrician, neonatologist and midwife, to determine if it could be considered to be suboptimal (possible or likely that different management would have prevented the adverse outcome).

Process outcomes

- Total number of CTG abnormalities (blue, yellow and red levels of concern) detected by the decision support software.
- Number of blue levels of concern on the decision support software, indicating a mild abnormality on the CTG.
- Number of yellow levels of concern on the decision support software, indicating a moderate abnormality on the CTG.
- Number of red levels of concern on the decision support software, indicating a severe abnormality on the CTG.

- Number of women with at least one yellow level of concern on the decision support software, indicating an abnormality on the CTG.
- Number of women with at least one red level of concern on the decision support software, indicating a severe abnormality on the CTG.
- Time from first red level of concern to birth.

Data collection schedule

Labour data and immediate outcomes were stored on the Guardian system and sent electronically to the co-ordinating centre. Data were extracted from the notes of babies admitted to the neonatal unit and for all neonatal deaths. All surviving children who were discharged were 'flagged' at the NHS Information Centre (England) and NHS Greater Glasgow & Clyde Safe Haven (Scotland), allowing all deaths to be identified. A sample of surviving children and mothers was followed up at 2 years by means of a parent-completed questionnaire to assess the child's health, development and (health service) resource use and to assess the mother's well-being and resource use.

Sample size and analysis

The required sample size was 46,000 births. We assumed an incidence rate of the primary outcome of 3 per 1000 births. This was calculated using reported rates of intrapartum stillbirth, neonatal death, moderate and severe NNE in broadly similar populations, and mild NNE (reliable data on significant asphyxial morbidity resulting in transfer to neonatal care were not available and so could only be estimated). 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 the poor neonatal outcome rate from 3 to 1.5 per 1000 births.

A statistical analysis plan was developed and approved by the Trial Steering Committee prior to analysis. Participants were analysed in the groups into which they were randomly allocated, regardless of allocation received. All women and babies with available data were included, except women with a missing consent form and women who withdrew consent to use their data. The number of babies with the composite primary outcome was presented for each group, and the risk ratio (RR) plus 95% confidence interval (CI) calculated. The mean [standard deviation (SD)] PARCA-R composite score was calculated and the mean difference plus 95% CI was calculated and compared using linear regression. Hazard ratios were estimated using a Cox proportional hazards model and rate ratios were estimated using Poisson regression. We adjusted for the stratification factors used in the randomisation, and robust variance estimators were used 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 and the mean difference plus 95% CI was calculated and compared. For secondary outcomes including the components of the primary outcome, a 1% level of statistical significance was employed.

The following prespecified subgroup analyses were undertaken, using the statistical test of interaction, for all neonatal outcomes, instrumental vaginal deliveries and caesarean section:

1. singletons versus twins
2. suspected growth restriction at labour onset versus no growth restriction
3. centre
4. body mass index (BMI) group [underweight (i.e. a BMI of 12–18.5 kg/m^2), normal (a BMI of 18.5–24.9 kg/m^2), overweight (a BMI of 25–29.9 kg/m^2), obese (a BMI of 30–70 kg/m^2), unrecorded].

For the economic evaluation, health-care resource use was compared using RRs for binomial variables and mean differences for continuous covariates. Parametric methods were used to estimate mean resource use, cost and maternal health-related quality of life EuroQol-5 Dimensions, three-level version (EQ-5D-3L), scores. Differences between treatment arms were adjusted using a random intercept binomial (for RRs) or linear (for mean differences) model adjusting for the stratification factors at randomisation (centre and twin birth) and clustering as a result of twins and multiple-birth episodes. A 95% significance level was

used for all comparisons. A multiple imputation framework with a chained equation was used for estimating resource use and EQ-5D-3L scores at 12 and 24 months.

The INFANT trial was registered with Current Controlled Trials ISRCTN98680152.

Results

Between 6 January 2010 and 31 August 2013, 47,062 women were randomised to the INFANT trial. A total of 1020 women (2.2%) were excluded from the analysis of the primary outcome. 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 56% of those contacted, although data were sufficiently complete for the analysis for 6707 children (53%).

There was no evidence of a difference in the incidence of the primary outcome of poor neonatal outcome between the groups, with a poor outcome being experienced by 0.7% of babies in both the decision support group ($n = 172$) and the no decision support group ($n = 171$) [adjusted risk ratio (aRR) 1.01, 95% CI 0.82 to 1.25]. Similarly, there was no evidence of a difference in any component of the composite primary outcome between the groups.

There was no evidence of any differences in any of the trial's secondary outcomes for the baby, including Apgar score, admission to a neonatal unit, metabolic acidosis of cord blood samples, the need for neonatal resuscitation and duration of hospital stay.

Just over half of all births were spontaneous vaginal births and there was no statistically significant difference in this outcome between the two groups. Half of the operative births were caesarean sections and half were instrumental. The proportion of women who underwent fetal blood sampling was higher in the decision support group (10.3%) than in the no decision support group (9.5%) (aRR 1.08, 99% CI 1.01 to 1.16). No other statistically significant differences in clinical outcomes were found between the two groups from trial entry to birth.

The overall proportion of cases with poor outcome in which babies were judged to have suboptimal care likely to have affected the outcome was 38% (21/71 cases).

There was evidence of a lower rate of yellow levels of concern in the decision support group (adjusted rate ratio 0.87, 99% CI 0.84 to 0.89).

There was no evidence of a difference between the two groups for any of the 2-year outcomes, including the long-term primary outcome of the PARCA-R, with a mean composite score of 98.0 points (SD 33.8 points) in the decision support group and 97.2 points (SD 33.4 points) in the no decision support group (mean difference 0.63, 95% CI –0.98 to 2.25).

There was no evidence that the decision support software produced different outcomes in any of the prespecified subgroups.

No evidence of a difference was detected in any category of resource use assessed, in categories of costs, or in total costs at 24 months for the infant (£104, 95% CI –£174 to £382) or for the mother (–£149, 95% CI –£314 to £16).

Conclusions

There is no evidence of a difference in the risk of a poor neonatal outcome, or intervention in labour, when using CTG interpretation software to support decision-making versus not using CTG interpretation in 46,000 women.

The strength of this study lies in its contemporaneous data collection and its size. Weaknesses include the potential for staff to learn from exposure to the decision support arm of the trial, resulting in improved outcomes in the control arm. This was identified when the trial was being planned. Part of our prior hypothesis was that, although some poor CTG interpretation is because of a lack of training, some clinicians may have a poor intrinsic pattern recognition ability that is not susceptible to improvement by training. This would not be affected by training and the performance of such clinicians would be particularly improved by assistance from automatic interpretation. There was some evidence that clinician behaviour was changed in the decision support arm of the trial. It may be that different action was taken in response to the alerts in the decision support arm of the trial, for example the clinicians might have reduced the dose of an oxytocin infusion in women having their labour augmented or changed maternal position if the CTG abnormality resulted from vena caval compression. Such actions could have prevented further yellow alerts, leading to a decrease in the incidence of repeat yellow alerts in this group, but we do not have any direct evidence that this was the case.

Detecting abnormalities in the fetal heart rate can improve outcome only if caregivers respond appropriately to the alerts. An expert panel reviewed all severe adverse outcomes in the trial and found no evidence that there were differences in levels of suboptimal care between the two groups. Therefore, we conclude that our hypothesis, that substandard care is largely related to a failure to identify pathological fetal heart rate patterns, is not supported.

There is currently no evidence to support the use of computerised interpretation of the CTG in women who have EFM in labour to improve clinical outcomes for mothers or babies.

Trial registration

This trial is registered as ISRCTN98680152.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research (NIHR). Sara Kenyon was part funded by the NIHR Collaboration for Leadership in Applied Health Research and Care West Midlands.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 06/38/01. The contractual start date was in September 2009. The draft report began editorial review in September 2016 and was accepted for publication in June 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).