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

The INFANT Collaborative Group†

†Investigators listed at end of paper
**Background**

Continuous electronic fetal heart rate monitoring in labour is widely used and computerised interpretation has the potential to deliver improvements in care. This trial aimed to find out whether the addition of decision support software to aid the interpretation of the cardiotocograph (CTG) reduced the number of ‘poor neonatal outcomes’.

**Methods**

This randomised controlled trial recruited women in labour having continuous electronic fetal monitoring, with a singleton or twin pregnancy, and 35 weeks’ gestation or more. They were allocated decision support or no decision support. Primary outcomes were (i) a composite of ‘poor neonatal outcome’ (stillbirth or early neonatal death excluding lethal congenital anomalies), significant morbidity (neonatal encephalopathy, admission to the neonatal unit within 48 hours for > 48 hours with evidence of feeding difficulties, respiratory illness or encephalopathy where there was evidence of compromise at birth); and (ii) developmental assessment at the age of two years in a subset of surviving children.

Registered with Current Controlled Trials: ISRCTN98680152

**Findings**

Between 6 January 2010 and 31 August 2013, 47062 women were randomised and 46042 included in the primary analysis (22987 in the decision support and 23055 in the no decision support groups). There was no evidence of a difference in the incidence of poor neonatal outcome between the groups: 0.8% (172) babies in the decision support group compared with 0.8% (171) babies in the no decision support group (adjusted RR 1.01, 95% CI 0.82 to 1.25). No evidence of differences were found for most of the secondary maternal, neonatal or long term outcomes.

**Interpretation**

Currently there is no evidence to support the use of computerised interpretation of the CTG in women who have continuous electronic fetal monitoring in labour to improve clinical outcomes for mothers or babies.
**Funding**

This project was funded by the NIHT HTA programme (project number 06.38.01). This report presents independent research commissioned 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, MRC, CCF, NETSCC, the HTA programme or the Department of Health.

**Introduction**

Continuous electronic fetal heart rate monitoring (EFM) in labour is widely used but its potential for improving neonatal outcomes has not been realised [1]. The reasons for this appear to be complex, but include difficulties with interpreting the fetal heart rate trace correctly during labour [2, 3, 4]. Computerised interpretation of the fetal heart rate has the potential to objectively detect abnormalities of the fetal heart rate pattern during labour that are associated with asphyxia so that action can be taken to expedite delivery and prevent stillbirth or the development of encephalopathy.

The electronic information capture system used in the trial, Guardian®, is a system for managing information from labour monitoring [5]. It displays the cardiotocogram (CTG) on a computer screen alongside other clinical data which are collected as part of routine clinical care. The system does not interpret any of the data being collected, it acts as an interface to collect and display data at the bedside, centrally on the labour ward, or in consultants offices and remotely. The decision-support software (INFANT®) has been developed to run on the Guardian® system. It extracts the important features of baseline heart rate, heart-rate variability, accelerations, type and timing of decelerations, the quality of the signal and the contraction pattern from the CTG [6, 7, 8]. The decision-support software then analyses these data along with the quality of the signals. The system’s assessment of the CTG is presented as a series of colour-coded alerts depending on the severity of the abnormality detected, with blue being the least severe, followed by yellow, and then red as the most severe alert (appendix).
The aim of the INFANT trial was to determine whether this computerised interpretation of the intrapartum CTG could improve the management of labour for women who were judged to require continuous electronic fetal heart rate monitoring.

Methods
This was a pragmatic randomised controlled trial, run in UK maternity units, in women who were judged by the local clinical team to require continuous electronic fetal monitoring (EFM) in labour. Research ethics committee approval for the study was granted by the National Research Ethics Service - Northern and Yorkshire Research Ethics Committee on the 30 September 2009 (reference number 09/H0903/31). The study protocol has been published [9].

Participants
Eligible women were those who: were judged to require continuous EFM by the local clinical team based on their existing practice, and who had consented to have EFM (continuous EFM in 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 [10]); had a singleton or twin pregnancy; were 35 weeks’ gestation or more; had no known gross fetal abnormality, including any known fetal heart arrhythmia such as heart block; were 16 years of age or older, and able to give consent to participate as judged by the attending clinicians.

All pregnant women attending recruiting hospitals were provided with written information about the trial during pregnancy and again when they presented 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.

Randomisation and masking
When the health care professional indicated that a woman was eligible, the Guardian® system confirmed that the necessary eligibility criteria were met and then randomly allocated women in the ratio 1:1 to either “CTG with decision support” or “CTG with no decision support”. The allocations were computer generated in Stata software (release 10.1)
using stratified block randomisation employing variable block sizes to balance between the
two trial arms by whether the pregnancy was a singleton (block sizes 12, 14, 16, 18, 20, 22,
24 allocated in proportion to the elements of Pascal’s triangle 1:6:15:20:15:6:1) or twin
(block sizes 2, 4, 6 allocated in proportion to the elements of Pascal’s triangle 1:2:1), and
within each participating centre. The trial was not masked. This was to measure any changes
in clinician behaviour, such as how much time the attending midwife spent with the woman,
based on knowledge that the decision support system was active or not.

**Procedures**

Clinicians in all participating centres were initially trained in the use of the decision support
software by staff from the trial coordinating centre. This process included developing a
“training team” at each site who were responsible for cascading training amongst the local
clinicians. The clinical management of women in the trial was not altered by their
participation, apart from when abnormalities of the CTG prompted the system in the
decision support arm of the trial to issue a series of alerts or alarms, which increased in
urgency with the severity of the abnormality [9].

Labour data and outcomes were stored automatically and contemporaneously onto the
Guardian® system, which were then sent electronically to the trial co-ordinating centre.
Data were extracted from the notes of babies admitted to the neonatal unit and for all
neonatal deaths. All children surviving to be discharged home from hospital following their
birth were ‘flagged’ at the NHS Information Centre for those born in England and NHS
Greater Glasgow & Clyde Safe Haven for those born in Scotland allowing all deaths occurring
after discharge to be identified. A sample of surviving children were followed up at two
years of age by means of a parent-completed questionnaire to assess the child’s health,
development and well-being (appendix).

**Outcomes**

The two primary outcomes of this trial were: (i) a composite of ‘poor neonatal outcome’ to
include deaths (intrapartum stillbirths plus neonatal deaths i.e. deaths up to 28 days after
birth, except deaths due to congenital anomalies); significant morbidity (moderate or severe
neonatal encephalopathy); admissions to the neonatal unit within 48 hours of birth for 48
hours or more with evidence of feeding difficulties or respiratory illness, where there was evidence of compromise at birth suggesting they were the result of mild asphyxia and/or mild encephalopathy, (ii) developmental progress as measured by the PARCA-R composite score [11, 12] at the age of two years for a subset of children.

Infant secondary outcomes were: intrapartum stillbirth (excluding deaths due to congenital anomalies); neonatal deaths (i.e. up to 28 days after birth but excluding deaths due to congenital anomalies); moderate or severe encephalopathy; admissions to the neonatal unit within 48 hours of birth for 48 hours or more with evidence of feeding difficulties or respiratory illness (where there was evidence of compromise at birth suggesting they were the result of mild asphyxia and/or mild encephalopathy); admission to a higher level of care; Apgar score less than 4 at 5 minutes; the distribution of cord blood gas data for cord artery pH; metabolic acidosis (defined as cord artery pH <7.05 and base deficit of 12 mmol/l or more); resuscitation interventions; seizures; destination immediately after birth; length of hospital stay; health and development outcomes at 24 months; non-verbal cognition scale (PARCA-R); vocabulary sub-scale (PARCA-R); sentence complexity sub-scale (PARCA-R); late deaths up to 24 months; major disability and non-major disability at 2 years; cerebral palsy.

Maternal secondary outcomes were: mode of delivery; operative intervention (caesarean section and instrumental delivery) for (i) fetal indication, or (ii) failure to progress, or (iii) combination of fetal distress and failure to progress, or (iv) other reason; grade of Caesarean section [13]; episiotomy; any episode of fetal blood sampling; length of (i) first stage, (ii) second stage and (iii) total length of labour from trial entry; destination immediately after birth; admission to a higher level of care.

Quality of care: all babies with an adverse outcome potentially associated with intrapartum asphyxia (trial primary outcome based on the baby’s condition after birth, plus cord-artery pH less than 7.05 with base deficit 12 mmol/l or more) and all neonatal deaths and intrapartum stillbirths had their care in labour assessed by panel review to see if care was considered to be suboptimal [9, 14].

As trial allocation was not masked it was important to measure any change in clinical care that could result from clinicians being aware of whether the decision support system was in
operation. Process outcomes were therefore measured to assess this: proportion of women with a CTG abnormality; number of CTG abnormalities identified in the two arms; time taken between last red alert and delivery (for these first three outcomes this was achieved in the ‘no decision-support’ arm by using the decision-support software to analyse the CTG trace after the trial was over and using this to determine when the alert would have occurred); number of routine measurements recorded during labour, including the number of vaginal examinations, use of epidural analgesia, use of labour augmentation, presence of meconium; number of thumb entries (comparable to a signature in paper notes) per hour from time of trial entry to first yellow level of concern or until the cervix was fully dilated (as a proxy measure to assess presence of a health professional in the delivery room during the labour).

**Statistical analysis**

A sample size of 46,000 births was needed [9]. We assumed an incidence of the primary outcome of 3 per 1,000 births. This was calculated by summing the anticipated rate of intrapartum stillbirth, neonatal death, moderate and severe encephalopathy, and mild encephalopathy [15, 16, 17]. The effect size which 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 outcome rate from 3 to 1.5 per 1000. In a study of 164 preterm infants [12], 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 approximately half a standard deviation below the standardised mean of 100. Assuming that a normal group of term infants would have a PARCA-R composite score half a standard deviation above this sample of preterm infants, we estimated a mean 2 year score of 96 (SD 33). Based on this estimate, a follow up sample of size 7,000 (3,500 per arm) had over 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 10 per 1,000 [18, 19, 20, 21]. A total of 46,000 women enabled us to detect a 28% relative risk reduction in this incidence with over 80% power, assuming a 5% level of significance, in those babies who have their cord artery pH measured.
During the early part of the trial, and with advice from the Data Monitoring Committee, the primary outcome definition was refined to ensure it captured babies who were likely to have experienced hypoxia during labour. The original definition of the component of the primary outcome “Admission to neonatal unit within 48 hours of birth for ≥ 48 hours with evidence of feeding difficulties, respiratory illness or encephalopathy” was initially capturing a range of conditions, many of which were unlikely to be related to hypoxia. Each case which fulfilled this component of the primary outcome was reviewed by an independent panel of neonatologists to ascribe each case (masked to allocation) as fulfilling the revised definition of the short-term primary outcome or not (appendix).

A statistical analysis plan was developed and approved by the Trial Steering Committee prior to analysis (appendix). For the main comparative 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 for women where a valid signed consent form could not be located, and for women who withdrew consent to use their data. The number (percentage) of babies with the composite primary outcome were presented for each group, and the risk ratio plus 95% confidence interval (CI) calculated. Risk ratios were estimated using generalised estimating equations (GEE), adjusting for the stratification factors used in the randomisation (centre and singleton/twin pregnancy) [9]. The mean (SD) PARCA-R Composite score was presented for each group, and the mean difference between groups plus 95% CI calculated and compared using GEE (Gaussian model with identity link). For secondary outcomes including the components of the primary outcome, a 1% level of statistical significance was employed.

To examine whether the effect of decision-support was consistent across specific subgroups of babies, the following prespecified subgroup analyses were undertaken, using the statistical test of interaction, for all neonatal outcomes; instrumental vaginal deliveries; caesarean section; centre; and the process outcomes: singletons versus twins; suspected fetal growth restriction at labour onset versus no growth restriction; BMI group (underweight <18.5; normal 18.5-24.9; overweight 25-29.9; obese >30; unrecorded); centre.
Major disability at two years was classified in the following domains: neuromotor function, seizures, auditory function, communication, visual function, cognitive function and other physical disability [22,23].

Stata/SE for Windows (version 13.1) was used 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 using three standard errors as the cut-off for consideration of early cessation, preserving the type-1 error across the trial [24].

The INFANT Trial is registered with Current Controlled Trials: ISRCTN98680152.

Role of the funding source
The trial was funded by the NIHR Health Technology Assessment programme as part of its commissioning stream. The funders of the study had no role in the trial design, data collection, data analysis, data 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 6 January 2010 and 31 August 2013, 47,062 women were randomised to the INFANT trial from 21 participating centres (appendix). A total of 1020 women (2.2%) were excluded from the analysis of the primary outcome (figure 1, appendix). The majority of these exclusions were 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 56% of those contacted, although data were sufficiently complete for the analysis for 6707 children.

Baseline characteristics were similar between the two groups of the trial (table 1). Mean maternal age was 29 years. The majority of women participating in the trial were of white
ethnic origin and the median BMI at booking was 25. Nearly 60% of women were having their first baby, and the majority of women in both groups had a gestational age between 38 and 41 completed weeks, although 11% were less than 38 weeks gestation and 6.3% were 42 weeks or above. Very few women had experienced a previous stillbirth (1%) and approximately 6% had undergone a caesarean section in a previous birth. Almost 60% of women had their labour induced.

There was no evidence of a difference in the incidence of the primary outcome of poor neonatal outcome between the groups, with 0.8% (172) babies having an poor outcome in the decision support group compared with 0.8% (171) babies in the no decision support group (aRR 1.01, 95% CI 0.82 to 1.25) (table 2). Similarly there was no evidence of a difference in any component of the composite primary outcome between the groups. In a pre-specified sensitivity analysis which used a different cut-off for defining compromise at birth (a score of 7 or greater indicating very severe compromise rather than a score of 3 or greater) this made no difference to the interpretation of the measure of effect for the primary outcome (aRR 0.97, 95%CI 0.58 to 1.63, appendix).

There was no evidence of any differences in any of the trials secondary outcomes for the baby (table 2), including Apgar scores, admission to the neonatal unit, metabolic acidosis of cord blood samples, the need for neonatal resuscitation, or duration of hospital stay.

Just over half of all births were spontaneous vaginal births and there was no statistically significant difference between the two groups (11823, 50.8%, of women in the decision support group versus 11959, 51.2%, of women in the no decision support group, aRR 0.99, 99%CI 0.97 to 1.01). Of the women who had an operative birth, half of these were by caesarean section and a half were instrumental (table 3). More women in the decision support group underwent fetal blood sampling, with 2366 (10.3%) of women in the decision support group versus 2187 (9.5%) in the no decision support group (aRR 1.08, 99%CI 1.01 to 1.16). No other statistically significant differences were found between the two groups from trial entry to birth in clinical outcomes (table 3).
Quality of care was assessed by an expert panel for all babies with an adverse outcome (trial primary outcome plus cord-artery pH less than 7.05 with base deficit 12 mmol/l or more) and for all neonatal deaths and intrapartum stillbirths (table 4, appendix). The addition of cord artery metabolic acidosis substantially reduced the number of babies with the primary outcome who were reviewed. There was no evidence of any difference in the proportion of babies judged to have received suboptimal care between the two groups (the overall observed incidence of suboptimal care likely to have affected the outcome was 38%, which is similar to that reported in previous studies).

The process outcomes collected in the trial are shown in table 5. For women in the no decision support group, the presence of alerts was calculated by the software during labour but not revealed to the woman or her care-givers. Using women with any level of concern as the denominator: blue levels of concern (the least severe alert) occurred frequently with women having a median of 9 such alerts during their labour (a rate of just below 1.4 per hour). The next, more severe alert, a yellow level of concern, occurred a median of twice per labour for women in the decision support group and twice for women in the no decision support group. There is 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) (table 5). For the most severe alert, the red level of concern, this occurred infrequently (a median of once per labour, with a rate of 0.14 per hour) and there was no evidence of a difference between the two groups, aRR 0.98, 99%CI 0.92 to 1.04..

Although there was concern that women in the decision support group in the trial may be left alone in labour more frequently given that the decision support software was running, there was no evidence to suggest that care-givers interacted with the Guardian system less frequently in this group. The rate of thumbprint entries on the Guardian system was 4.22 per hour in the decision support group versus 4.21 per hour in the no decision support group (adjusted rate ratio 0.99, 99%CI 0.95 to 1.03, table 5).

The time from the last red level of concern to birth was similar in both groups with a median of 58 minutes. Although this appears long, there were red levels of concern that did not prompt immediate delivery, for example when the CTG monitor was picking up the
maternal heart rate. In a sub-group of 500 traces containing at least one red level of concern, the last red level of concern was judged (by an expert co-investigator, PS) to be a valid fetal concern for 55%. For the remainder, the maternal heart rate triggered the red level of concern in 70%, and other reasons in 30%.

Families were contacted when the surviving child(ren) born in the INFANT trial reached two years of age. Nearly 7,000 families returned a questionnaire. There were statistically significant differences between the characteristics of the mothers who responded compared with the entire trial cohort, as well as between mothers who did and did not respond to an invitation to complete the questionnaire (appendix). Many of these differences are very small, but given the large numbers of participants in the trial, the differences are statistically significant. In general responders, when compared with the entire trial cohort were more likely to be slightly older, of white ethnic origin, to have given birth at a later gestational age, to be less likely to smoke and to have been having their first baby.

Of the 7066 infants for whom a questionnaire was returned, data could be analysed for 6707 of them (95%). There was no evidence of a difference between the two groups for any of the two year outcomes, including the long term primary outcome of the Parent Report of Children’s Ability (PARCA-R) with a mean (SD) score of 98.0 (33.8) in the decision support group and 97.2 (33.4) in the no decision support group (mean difference 0.63 95%CI -0.98 to 2.25) (table 6). Nearly 6% of children were classified as having major disability. The classification of disability used [21, 22] resulted in relatively large numbers of children being assigned a major disability as a consequence of poor growth (between 2.8% and 3% 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).

A number of subgroup analyses were pre-specified. These analyses are included in the appendix. There was no evidence that the decision support software performed significantly differently for any of the subgroups (multiple pregnancy; suspected fetal growth restriction; body mass index of the mother) for either the primary outcome or a
limited range of pre-specified secondary outcomes (appendix). There were also no differences found in the distribution of cord blood pH measurements (appendix). There were differences seen in the number of alerts in the analysis by centre (appendix). The reasons for this are unclear, particularly as there were no other significant differences for the other outcomes by centre (appendix).

**Discussion**

In this trial there is no evidence of a difference in the risk of a poor neonatal outcome using this decision support system in 46,000 women. A further randomised trial which evaluated the use of decision support in women monitored during labour using fetal ECG monitoring [25], similarly found no evidence that decision support improved their primary outcome of cord blood metabolic acidosis in 7730 women [26].

Using a composite primary outcome may not always be helpful if different components of the outcome respond differently to the intervention. We initially hypothesised that the components of the composite outcome (extended perinatal mortality, neonatal encephalopathy, and prolonged admission to a neonatal unit following birth in a poor condition) would be similar in incidence, with each component likely to contribute approximately a third to the composite. If one component of the composite dominates the others, then effectively the trial results reflect any differences being detected within this dominant component [27].

Estimates of the incidence of the components of the primary outcome for the eligible study population were difficult to find before the trial commenced [9]. In retrospect, the observed perinatal mortality in the study (stillbirth and neonatal death) was lower than anticipated (0.3 per 1000 versus 1.05 per 1000), and the incidence of neonatal encephalopathy requiring cooling was also lower than anticipated (0.8 per 1000 versus 1.3 per 1000). However, the incidence of prolonged neonatal unit admission with evidence of compromise at birth occurred much more frequently than anticipated given an overall primary event rate of 8 per 1000, compared with our estimated 3 per 1000. This allowed us to have power to detect more modest differences than we had originally planned.
The strength of this study lies in its contemporaneous data collection and its size, the latter being designed to detect differences in substantive perinatal outcomes, as well as in more frequent outcomes such as cord metabolic acidosis, and operative delivery. Potential weaknesses include whether the UK setting, where continuous EFM is not routine [10], may make generalisability of the findings to settings where EFM is routine difficult. One other concern is 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 potential weakness was identified when the trial was being planned. We acknowledged that passive learning from the decision support system was possible, and the only way to completely rule out this effect would be to conduct a cluster randomised trial. Such a design was unfeasible given the limited number of centres with the Guardian system in the UK and the very low incidence of the primary outcome measure. We therefore collected a range of process outcomes to measure the impact on clinician behaviour during the trial. We consider the impact of this potential learning effect to be small, partly because the overall incidence of poor perinatal outcome was higher than anticipated (8 per 1000 compared with 3 per 1000). In addition, there was evidence that clinical behaviour was changed in the decision support arm of the trial, with an increased incidence of fetal blood sampling, and a lower incidence of repeated yellow alerts. This suggests that different action was taken in response to the alerts in this arm of the trial. However, these changes in behaviour did not result in any evidence of a change in clinical outcomes.

Detecting abnormalities in the fetal heart rate can only improve outcome if care-givers 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 suboptimal care between the two groups.

The decision support system used in this trial clearly identifies fetal heart rate abnormalities [6, 7, 8]. However, it does not incorporate other information about the labour, such as the duration of labour, the rate of labour progress, presence of meconium, whether the woman has a temperature, and whether there is suspected fetal growth restriction, all of which may modify the way a clinician interprets the fetal heart rate and acts on this information.
Further development of decision support software may improve the quality of the information the system provides to clinicians to make a difference to outcomes. Given the importance for parents, clinicians and health services of the consequences of intrapartum hypoxia, there continues to be an urgent need to identify signs of early compromise during labour so that timely intervention can be used to decrease these poor outcomes.

Currently there is no evidence to support the use of computerised interpretation of the CTG in women who have continuous electronic fetal monitoring in labour to improve clinical outcomes for mothers or babies.

**Contributors**
All authors contributed equally to the development of the protocol and management and undertaking of the trial. LL did the analyses. PB wrote the report and revised it with input from all authors. All authors read and approved the final manuscript.

**Declaration of interests**
Keith Richard Greene. Founder shareholder of K2 medical systems. Clinical Director for development of the INFANT system.  
Dr Robert D F Keith. Director General, Chairman and Founder Shareholder K2 Medical Systems.  
Christopher Mabey. Under employ and shareholder of K2 Medical Systems the technology provider for study.

**Acknowledgements**
Funding source: This project was funded by the NIHT HTA programme (project number 06.38.01). This report presents independent research commissioned 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, MRC, CCF, NETSCC, the HTA programme or the Department of Health.
We thank the women who agreed to join the trial.

**Writing committee:** [alphabetical order]

**Investigator group:** [alphabetical order].

**Trial steering committee:** [alphabetical order]

**Data monitoring committee:** [alphabetical order]

**Clinical trials unit staff:** National Perinatal Epidemiology Unit, University of Oxford (date to date): [alphabetical order]

Comprehensive Clinical Trials Unit, UCL (date to date): [alphabetical order]

**References**


23. Disability and perinatal care: measurement of health status at two years. In A report of two working groups convened by the National Perinatal Epidemiology Unit and the former Oxford Regional Health Authority Oxford: NPEU; 1994