Continuous fetal heart rate monitoring: is there a conflict between confidential enquiry findings and results of randomized trials?

Mary C M Macintosh MD MRCOG

J R Soc Med 2001;94:14-16

SECTIONS OF EPIDEMIOLOGY & PUBLIC HEALTH AND OBSTETRICS & GYNAECOLOGY, 9 MARCH 2000

Continuous electronic fetal heart rate monitoring (EFM) is used in all but 9 of 250 maternity hospitals in the UK (Kelly T, RCOG survey). The technology became available commercially in the 1960s. Previously the fetal heart rate had been measured by intermittent auscultation between contractions. Initially there were high expectations that continuous monitoring would allow early recognition of hypoxia, with consequent timely intervention and avoidance of death or brain damage in the newborn. However, in the mid-1990s a Cochrane review of evidence from randomized control trials (RCTs) concluded that the only benefit from EFM was a reduction in neonatal seizures¹. Unless EFM was used in conjunction with fetal blood sampling it was associated with an increase in caesarean section rates. There was no reduction in perinatal mortality.

Paradoxically, the fourth report of Confidential Enquiries into Stillbirths and Deaths in Infancy (CESDI) concludes that the major correctable error for avoidance of intra-partum related deaths was the interpretation of EFM traces. Since more than half such deaths were judged from the anonymized case notes to have been avoidable, the report indicates that EFM does have a part to play in avoidance of deaths². Here I review the relative strengths and weaknesses of trial, meta-analysis of trials, observational and enquiry evidence, and discuss whether this conflict can be resolved.

TRIAL EVIDENCE

The first two RCTs of continuous electronic fetal heart monitoring versus intermittent auscultation were published in 1976^{3,4}. Neither found an improvement in newborn outcome with continuous monitoring but both demonstrated an approximately twofold increase in caesarean delivery. By 1994 at least 12 RCTs had been published with a total of 58 855 labours. All showed higher caesarean section rates with EFM and in half this was statistically significant. There was no reduction in perinatal deaths but one trial found a reduction in perinatal deaths due to fetal hypoxia⁶.

Most of the trials were small, ranging from 250 to 950. The two largest were the Dublin trial⁷ involving just under 13 000 cases, and the Athens trial⁶ with nearly 1500. But the weakness of all RCTs was the lack of statistical power to detect differences in rare outcomes, in particular perinatal death due to fetal hypoxia. In the UK in the 1990s, one such death occurred in 1500 deliveries²; a trial size of 85 000 would be required to demonstrate a halving from 2 in 1500 to 1 in 1500. All trials recorded perinatal deaths as an outcome but most of the deaths in this category are from causes other than hypoxia in labour. Neonatal seizures occur in 2 per 1000 deliveries⁸. The Dublin trial⁷, the only RCT of adequate power, showed a significant reduction in this outcome.

The effect of continuous EFM on more frequent outcomes such as caesarean section or instrumental delivery is easier to demonstrate. An increase in both interventions was observed: caesarean section rates increased in all trials; instrumental delivery increased in all but two trials.

In summary, the trials show a consistent increase in operative delivery with the use of EFM but all trials were too small to address the benefit or otherwise of the technology in terms of rare serious outcomes for the baby, such as death or cerebral palsy.

META-ANALYSIS

Systematic reviews are crucial to public health policy-making. Meta-analysis, in which the data from individual trials are combined, increases the power to assess rare outcomes. By the mid-1990s 12 RCTs, and 2 systematic reviews^{1,5} comprising 9 RCTs and 18 561 labours, had been published. This is still well short of the 85 000 needed to address death from perinatal asphyxia. Even if a meta-analysis could achieve adequate power it would still have the drawback of combining studies from substantially different clinical settings. The wide variation in outcome measures is illustrated by caesarean section rates in the 'intermittent-auscultation' arms of the trials which ranged from 2% to 15.3% and perinatal deaths which ranged from 1 to 35⁹.

The size of the individual studies weights the meta-analysis. The Dublin study⁷ dominates all systematic reviews, being ten times larger than the next largest study

and up to thirty times the size of other studies. This trial dealt with low-risk labours, half the women delivering within two hours of admission to the labour ward; the caesarean section rate was 2.4% in the EFM group and 2.2% in the intermittent-auscultation group. The heterogeneity of the RCTs dilutes the effects observed in high-risk populations; for example, the Athens study⁶ (the second largest) involved high-risk women and the EFM group showed a 60% reduction in perinatal deaths from hypoxia.

Conclusions may depend on which trials reviewers choose to include or exclude¹⁰. One meta-analysis showed a significant reduction in perinatal deaths due to fetal hypoxia⁹ whereas another¹ did not: the reason lay in inclusion or exclusion of a single large trial⁶.

Not all trials identified intrapartum hypoxia deaths at the time of publication, and subsequent subanalysis on this is subject to variation in interpretation; thus, with cumulative meta-analysis of the existing trials we can expect continued debate on the merits or otherwise of EFM.

OBSERVATIONAL EVIDENCE

Randomized control trials are usually the best way to evaluate an intervention. However, when the outcome is rare then trials of appropriate size are likely to be impracticable. Despite their inherent biases, observational studies must contribute to any review of the evidence. Examination of historical trends might be of value.

In England and Wales, perinatal mortality fell from 21 per 1000 total births in 1973 to 8.2 per 1000 in 1998¹¹. This was the period after introduction of EFM, but there are many confounding factors. Intrapartum death rates were not routinely collected until 1986, when death certificates were amended to indicate whether the stillbirth occurred before or during labour. In the years before this change, mortality coded as 'anoxia of unspecified cause' had been stable at 2 per 1000 in England and Wales until 1973, when there was a decline, again coinciding with the introduction of EFM. Since 1986 the Office for National Statistics (ONS) has reported a category termed 'intrapartum stillbirths with associated cause'. This accounts for 10% of all stillbirths and seems to be falling slightly¹².

A review by Hornbuckle *et al.*¹³ cites 13 observational studies comparing monitored with unmonitored pregnancies, and in 12 of them intrapartum deaths declined with the introduction of monitoring. However, before-and-after studies of this type are subject to secular trends. Furthermore, there may be a trend to publication bias. Only one of these studies adjusted for background risk: probably the women being monitored are a higher risk group, so the effect on mortality could be an underestimate. There have been two case-control studies^{14,15} demonstrating a significant relationship between suboptimal intra-

partum care and fetal death during labour, but both are subject to confounding factors—namely antenatal risk, the use of continuous EFM and 'presumed' intrapartum distress.

ENQUIRY EVIDENCE

The Confidential Enquiry into Stillbirths and Deaths in Infancy (CESDI) dates from 1992. Its remit is to examine the causes of stillbirths, neonatal deaths and infant deaths and in particular to identify aspects of care that might have contributed to the unsatisfactory outcome. These functions are served by 16 regional coordinators (a clinically integrated network covering England, Wales and Northern Ireland) and a central secretariat.

CESDI is notified of 99% of all deaths occurring between 20 weeks' gestation and 1 year of life, approximately 10 000 deaths annually. A subset of these is chosen for enquiry. Clinicians enter the cause of death on the notification form after a review of the clinical and pathological findings. A case-by-case ascertainment check with the Office for National Statistics ensures identification of all relevant deaths. The first enquiry topic was deaths of normally formed babies weighing above 1.5 kg due to an event in labour, 1993–1996; the findings were published in 1997².

The enquiries are held on a regional basis by a multidisciplinary panel which reviews the anonymized case records. Panel members are from within the region but are not from the unit associated with the case. There is no direct feedback to the units or individuals involved and the results are summarized regionally and nationally to form the basis of the annual report.

The findings suggest that over half the intrapartum-related deaths are avoidable. Panels were asked to comment on areas of suboptimal care in the antenatal, intrapartum or postpartum period. Of the 1708 comments made on intrapartum care the commonest area of concern related to the interpretation of fetal heart traces (791 comments)². The robustness of the panel assessments in 1995 was assessed in a second-pass panel exercise¹⁶ and the topic most likely to be identified consistently by both panels was difficulty in interpreting EFM. However, the unstructured approach used in 1995 led to the recommendation of providing predefined standards for guidance at a panel enquiry¹⁶. These findings related to deaths alone; they did not include observations in women with normal outcomes, and the remit of the panels was to identify suboptimal care. Probably, therefore, the enquiries overemphasized the importance of obstetric care in causing death.

RESOLVING THE CONFLICT

Future studies, be they trial, observational or enquiry, need to address the lack of standardization in interpretation of the FHR pattern¹⁷. Interobserver and intraobserver reliability is low^{18,19}. One-third of labours have some form of variant, and

standard definitions of FHR abnormalities are a priority¹⁷. This was recognized by the US National Institute of Child Health and Human Development in 1997²⁰, but the US group felt unable to offer guidelines on management of specific patterns because of incomplete knowledge of the underlying pathophysiology of FHR variation in hypoxia. Within England and Wales a national evidence-based guideline funded by the Department of Health is in preparation but standard definitions of FHR abnormalities have yet to be agreed.

Meta-analysis will always suffer the disadvantage of joining studies which differ substantially in their setting and entry criteria. Retrospective subdivision of perinatal deaths due to hypoxia is subjective and small changes in these numbers are highly influential in the interpretation of the meta-analysis. Many of the past studies are not representative of current practice and meta-analysis will not provide an agreed and definitive answer.

The ideal remains the RCT, but can we realistically expect a randomized trial involving 85 000 labours, comparing continuous fetal heart monitoring with intermittent auscultation? This number represents just over one-eighth of all deliveries in the UK annually. If the question is restricted to low-risk women and if we assume two-thirds of all pregnancies are in this category, then at least one in five of all suitable maternities would need to be recruited to achieve adequate numbers in a year and this takes no account of the individual preferences of health professionals and mothers. Strongly held opinions in either direction are likely to result in non-participation. Even if the issues of numbers and preference could be dealt with, we would have to address the practicalities of the one-to-one midwifery required for intermittent auscultation; unfortunately, human companions are more resource-intensive than their electronic equivalents. Thus a trial of adequate size seems unlikely. This leaves observational and enquiry studies—the case-control approach where the intrapartum care in deaths is compared with that in well babies. In case-control studies the choice of appropriate controls is particularly problematic. The major confounding factors are antenatal risk factors and intrapartum fetal distress. Probably, more complex decisions will be necessary in the cases than the controls. Agreed criteria are needed for management of suboptimal traces and there is the additional difficulty of assessing care in the intermittent-auscultation group where a continuous record will not exist (so presumed fetal distress is less likely to occur). Ultimately the findings of the case-control approach can only be supportive or otherwise, not definitive.

With the founding of the National Institute for Clinical Excellence, central guidance on participation in trials becomes a possibility. Some people may argue that the possible risks and benefits of EFM technology need clear answers. But if the evidence went against EFM, would the

technology then be removed? Before this could happen there would need to be widespread debate, including not just health professionals but mothers as well. Meanwhile, this is a key area for the National Institute for Clinical Excellence. The conflict between the enquiries and the trial evidence is destined to continue rather than be resolved.

REFERENCES

- 1 Neilson JP. EFM+scalp blood sampling vs intermittent auscultation in labour. In *Cochrane Pregnancy and Childbirth Database; Issue 2*. Oxford: Update Software, 1995
- 2 *Confidential Enquiries into Stillbirths and Deaths in Infancy, 4th Report*. London: Maternal & Child Health Research Consortium, 1997
- 3 Haverkamp AD, Thompson HE, McFee JG, Cetrulo C. The evaluation of continuous fetal heart rate monitoring in high risk pregnancy. *Am J Obstet Gynecol* 1976;**125**:310–17
- 4 Renou P, Chang A, Anderson I, Wood C. Controlled trial of fetal intensive care. *Am J Obstet Gynecol* 1976;**126**:470–6
- 5 Thacker SB, Stroup DF. Continuous electronic heart rate monitoring for fetal assessment during labor. *Cochrane Library; Issue 2*. Oxford: Update Software, 2000
- 6 Vintzileos AM, Antsaklis A, Varvarigos I, Papas C, Sofatzis I, Montgomery JT. A randomized trial of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation. *Obstet Gynecol* 1993;**81**:899–907
- 7 Grant A, O'Brien N, Joy MT, Hennessy E, MacDonald D. Cerebral palsy among children born during the Dublin randomized trial of intrapartum monitoring. *Lancet* 1989;**ii**:1233–6
- 8 Mongelli M, Chung TK, Chang AMZ. Obstetric intervention and benefit in conditions of prevalence. *Br J Obstet Gynaecol* 1998;**105**:246–7
- 9 Vintzileos AM, Nochimson DJ, Guzman MD, Knuppel RA, Lake M, Schiffrin BS. Intrapartum electronic fetal heart rate monitoring versus intermittent auscultation: a meta-analysis. *Obstet Gynecol* 1995;**85**:149–55
- 10 Gardosi J. Systematic reviews: insufficient evidence on which to base medicine. *Br J Obstet Gynaecol* 1998;**105**:1–5
- 11 Office for National Statistics. *Health Statistics Quarterly* 3, London: ONS, 1999
- 12 Alberman E, Blatchley N, Botting B, Dunn A. Medical causes on stillbirth certificates in England and Wales; distribution and results of hierarchical classifications. *Br J Obstet Gynaecol* 1997;**104**:1043–9
- 13 Hornbuckle J, Vail A, Abrams KR, Thornton JG. Bayesian interpretation of trials: the example of intrapartum electronic fetal heart monitoring. *Br J Obstet Gynaecol* 2000;**107**:3–10
- 14 Niswander K, Henson G, Elbourne D, et al. Adverse outcome of pregnancy and the quality of obstetric care. *Lancet* 1984;**ii**:827–31
- 15 Gaffney G, Sellars S, Flavell V, Squier M, Johnson A. Case-control study of intrapartum care, cerebral palsy, and perinatal death. *BMJ* 1994;**308**:743–50
- 16 *Confidential Enquiries into Stillbirths and Deaths in Infancy, 5th Report*. London: Maternal and Child Health Research Consortium, 1998
- 17 Parer JT, King T. Fetal heart rate monitoring: is it salvageable? *Am J Obstet Gynecol* 2000;**182**:982–7
- 18 Beaulieu MD, Fabia J, Leduc B, et al. The reproducibility of the intrapartum cardiotocogram assessments. *Can Med Assoc J* 1982;**127**:214–16
- 19 Barrett JFR, Jarvis GJ, MacDonald HN, Puchan PC, Tyrell SN, Lilford RJ. Inconsistencies in clinical decisions in obstetrics. *Lancet* 1990;**336**:549–51
- 20 National Institute of Child Health and Human Development Planning Workshop. Electronic fetal heart rate monitoring: research guidelines for interpretation. *Am J Obstet Gynecol* 1997;**177**:1385–90