Authors: Banta, H. David; Thacke, Stephen B.

Title: Costs and Benefits of Electronic Fetal Monitoring: A Review of the Literature. NCHSR Research Report Series.

Institution: Center for Disease Control (DHEW/PHS), Atlanta, Ga.; National Center for Health Services Research (DHEW/PHS), Hyattsville, Md.

Report No: DHEW-PHS-79-3245; NCHSR-79-91

PUB DATE: Apr 79

NOTE: 40p.

AVAILABLE FROM: NCHSR Publications and Information Branch, 3700 East-West Highway, Room 7-44, Hyattsville, Maryland 20782 (no charge)

EDRS Price: ME01/PC02 Plus Postage.

Descriptors: *Birth; Clinical Diagnosis; *Cost Effectiveness; Infant Mortality; *Medical Care Evaluation; *Medical Treatment; *Quality Control; *Risk; Safety

Identifiers: Electronic Fetal Monitoring

Abstract: This report focuses on electronic fetal monitoring (EFM)--a technology that was developed during the 1960s and has rapidly spread into use in clinical obstetrics. The report includes a review of the extensive published literature on EFM and related subjects. It also contains original calculations concerning the technique's specificity and sensitivity, its predictive value as a diagnostic test, and the financial costs associated with its potential risks. (Author/MP)
Costs and Benefits of Electronic Fetal Monitoring: A Review of the Literature
The Research Report Series is published by the National Center for Health Services Research (NCHSR) to provide significant research reports in their entirety. Research Reports are developed by the principal investigators who conduct the research, and are directed to selected users of health services research as part of a continuing NCHSR effort to expedite the dissemination of new knowledge resulting from its project support.

Abstract

This report focuses on electronic fetal monitoring (EFM)—a technology that was developed during the 1960s and has rapidly spread into use in clinical obstetrics. The report includes a review of the extensive published literature on EFM and related subjects. It also contains original calculations concerning the technique's specificity and sensitivity, predictive value as a diagnostic test, and the financial costs associated with its potential risks.
Costs and Benefits of Electronic Fetal Monitoring: A Review of the Literature

H. David Banta, M.D., M.P.H.
National Center for Health Services Research
Currently: Office of Technology Assessment
Congress of the United States, Washington, D.C.

Stephen B. Thacker, M.D.
Bureau of Epidemiology
Center for Disease Control
Atlanta, Georgia

April, 1979
This NCHSR Research Report was written by H. David Banta, M.D., M.P.H., and Stephen B. Thacker, M.D. Although Dr. Banta started this study while employed by the Office of Technology Assessment (OTA), most of his work on the project was done during his tenure on the NCHSR staff in 1978. He has returned to OTA. Dr. Thacker was asked to participate in the literature analysis and the development of the cost figures because of his epidemiological expertise. He is employed by the Public Health Service's Center for Disease Control.

The authors wish to thank Ian Chalmers, M.D., Marshall Goldberg, M.D., Raymond Neutra, M.D., Judith Rooks, M.S., M.P.H., and Carl Tyler, M.D., for their criticisms and suggestions; Thomas Hodgson, Ph.D., for his help with the cost estimates; and Linda Burton, for clerical assistance.

Additional copies of this report can be obtained on request from the NCHSR Publications and Information Branch, 3700 East-West Highway, Room 7-44, Hyattsville, MD 20782 (Tel.: 301/436-8970). Other current NCHSR publications are announced on the last pages of this publication.

The views expressed in this publication are those of the authors, and no official endorsement by the National Center for Health Services Research is intended or should be inferred.
This report focuses on electronic fetal monitoring (EFM)—a technology that was developed during the 1960s and has rapidly spread into use in clinical obstetrics. The report includes a review of the extensive published literature on EFM and related subjects, as well as original calculations concerning the technique's specificity and sensitivity, predictive value as a diagnostic test, and financial costs.

Only recently have controlled clinical trials made it possible to try to answer questions about the benefits, risks, and costs of EFM, whose use is becoming increasingly controversial. EFM is an interesting case study in view of the current policy debate concerning the evaluation and control of medical technology. Although EFM is an important technique, it has even greater significance as an example of how little is known about many technologies used in medical practice. For instance, there is not much information available on the costs and benefits of EFM in a clinical setting, and research investments in that area have been pitifully small. Because there are few formal mechanisms for determining which medical technologies should be evaluated, almost any medical technology can be introduced into the health system for immediate purchase and use. With the lack of consumer and provider responsibility for the costs of medical technology, regulatory constraints have not been an effective mechanism for assuring that medical technology is well-evaluated before it is placed on the market.

Although the results of this study are controversial, I believe that they are well-grounded in the scientific literature, and that this report is a significant contribution to the literature on medical technology. Because of the professional associations of the authors while the study was in progress, this report represents the result of collaboration between three Federal health agencies: the National Center for Health Services Research, the Office of Technology Assessment, and the Center for Disease Control. We are happy to be associated with such an effort.
Electronic fetal monitoring (EFM) is an obstetric technique developed in the 1960s. Although it is available in virtually every delivery room in this country (1-4), its use as a medical procedure has become controversial. Questions about its efficacy, safety, and cost have been raised by women's groups, in recent newspaper articles (5-6), in U.S. Congressional hearings (7), and in the medical literature (8-9). The ethical implications of EFM also have been debated (10).

Today, EFM is done either externally, using ultrasound, internally by attaching electrodes directly to the fetus to monitor the electrocardiogram (ECG), or by sequential use of both techniques prior to and after rupture of the amniotic sac. In internal monitoring, a catheter is used to monitor uterine contractions. While EFM technology was evolving in this country, fetal scalp blood (FSB) sampling (to determine the pH of the fetal blood) was being developed simultaneously in Germany. It was not long before the two procedures were paired, and FSB sampling has gradually become an integral part of EFM (11-12).

When first used, EFM was applied largely to high-risk pregnant women, and some continue to advocate its use primarily for that group (13-14). However, an increasing number of obstetricians favor EFM of all patients in labor, and many institutions in this country (1,15-21) and elsewhere (22-24) do EFM routinely. As would be expected in light of such rapid dissemination, surveys of physicians in 1970 (25) and 1976 (26) have found a high degree of acceptance of the procedure, with the latter survey showing that 77 per cent of the physicians surveyed believe that all labors should be monitored electronically. In the USSR, Doppler ultrasound monitoring has aroused great interest (27), and in Germany 58 percent of the institutions surveyed in 1971 tried to electronically monitor all patients in labor (28).

However, the spread of EFM is part of the increasing dependence on technology in medical practice. Questions about the use of new technology in obstetrics have been raised both by women’s groups and some physicians (8,29-32) who point out that birth is a normal process.

This paper examines the evidence concerning the efficacy and safety of EFM by reviewing the English-language literature and makes some calculations on its cost.
### Contents

- Introduction
- History of fetal monitoring
- Quality of diagnostic information
  - FHR monitoring
- FSB pH monitoring
- EFM combined with FSB sampling
- Predictive value
- Impact on outcomes
  - Preventable stillbirths and neonatal deaths
  - High risk
  - Infant and perinatal mortality
- Brain damage
- Summary
  - Maternal outcomes
  - Fetal risk from EFM
  - Maternal risk from EFM
- Maternal reaction to EFM
- The future of EFM
  - Costs of EFM
- Conclusions
- Tables
- References
Biblical references may be found to fetal movements or signs of fetal distress in utero. In 1679, the passage of meconium was mentioned as a “certain sign of death.” Other books in the 1700s and 1800s merely identified it as a warning sign. By the 1700s, signs of life or death of the unborn child were clearly described in midwifery textbooks (38). In 1793, it was noted that the pulse of the fetus could be felt in the presenting part — head, foot, or cord.

De Kergaralec first described obstetric auscultation as a potentially important diagnostic tool in 1821 (33). He noted that by auscultation one could detect fetal life, lie, and distress during labor, and he described fetal bradycardia as a sign of distress (34). In 1838, Nagele associated head compression with bradycardia (35). Kennedy published his Observations on Obstetric Auscultation in 1843, which described the delay in recovery of the FHR after a contraction as a warning sign (34). In 1885, Preyer identified head compression, placental insufficiency, and anoxia as causes of fetal bradycardia (35). In 1893, Von Winckle established criteria for the diagnosis of fetal distress: a FHR of >160 or <120 (34).

Attempts to record the fetal heart tones were made by Pestalozza in 1891, Seitz in 1903 and by Hofbauer and Weiss in 1908 (36). Strassman and Mussey reported success in recording fetal heart tones in 87 per cent of 52 patients in 1938 (29). Cremer used a vaginal electrode to obtain a fetal ECG tracing in 1906 (29).

During the late 1970s and 1940s an increasing number of investigators began to take an interest in FHR monitoring, and a great deal of work was done with newborn rabbits (34) and with fetal lambs (34–35). In the post–World War II period, electronics developed rapidly, and an ultrasonic Doppler device was available by 1964.

In fact, in 1974 Hehre stated that abdominal recording had not advanced in 25 years (37).

In 1957, Hon first reported the successful recording of the fetal ECG for the abdomen (38). In 1959, Hon observed profound bradycardia on a dying fetus and also noted that manipulation of the prolapsed cord caused cardiac irregularities and bradycardia (39). In 1960, Hon described a clinical FHR monitor. However, a better method of recording was necessary. According to Hehre (37), Hon conceived the idea of an electrode to record the ECG, made a prototype in his home workshop, and appeared in the delivery suite at 3:00 a.m. for the initial test. The electrode was successful (40) and became the basis for presently used EFM devices. The major improvement was a spiral electrode developed by Hon in 1972 (41).

By 1969, EFM by ultrasound and by direct ECG monitoring had begun to diffuse rapidly (4). Hon and Caldeyro-Barcia were the most active investigators, describing a variety of changes in the FHR and correlating them with clinical states. Working independently, these two men devised different schemes for interpreting FHR patterns. Because of the resulting confusion, they agreed on a standard terminology in the 1970s (37).

Saling had found FHR monitoring to be nonspecific, and in 1961 he reported a technique for FSB sampling. In this technique, a scalpel is passed through the cervix, and a small wound is made in the baby’s scalp. A sample of blood is then collected in a capillary tube and analyzed for acid-base parameters. It is possible to do measurements such as glucose and electrolytes from scalp samples, but only pH, PO₂, and PCO₂ are routinely done. Saling has demonstrated that pH alone provides as much information as the three tests together. FSB sampling has become widely accepted throughout the developed world, and has remained essentially unchanged since.

Today FHR monitoring is done either externally, using ultrasound, internally with ECG, or these two techniques are used sequentially in individual pregnancies prior to and after rupture of the amniotic sac. There exist no national figures on the use of EFM equipment. In particular, there is no data on the relative use of these two modalities. Koh reported that 37.8 per cent of electronically monitored patients in his institution are followed only externally, 11.2 per cent are only internally monitored, while both techniques are used in 51.0 per cent of deliveries (43). Quilligan reported that internal monitoring is probably more generally used (4).

By 1972 an estimated 1000 EFM systems were in use in the United States. In a survey in 1976 of 360 programs with residencies in obstetrics, 278 of 279 respondents reported the use of EFM (2). We believe that the vast majority of obstetrics services now have EFM equipment, and that over half of the labors are monitored electronically.
Although the ultimate measure of efficacy is improved patient outcome, such improvement is often assumed for diagnostic procedures if the information obtained is reliable and valid. Therefore, evaluating the efficiency of EFM includes considering the quality of the information obtained. Two indices are used: changes in FHR and changes in the pH of the FSB.

The changes in the FHR with a uterine contraction, generally referred to as periodic FHR changes, have been extensively studied (17,34,43-46). Early bradycardia or "deceleration," which has been attributed to fetal head compression, is a normal finding that occurs more or less simultaneously with the development of the contraction. Late deceleration is slowing of the FHR that begins after the peak intensity of the uterine contraction and extends into the period between contractions. It is considered an abnormality caused by uteroplacental insufficiency.

Another abnormal FHR pattern, variable deceleration, occurs out of phase with the contractions, and is considered to be due to umbilical cord compression (46). Both late and variable decelerations often can be eliminated by changes in position, maternal oxygen therapy, or cessation of oxytocin, which by itself can produce abnormal FHR patterns (47-48). Either can be ominous if they persist. There is a normal fluctuation of FHR during labor and loss of these beat-to-beat (47) variations can signal danger to the fetus, as can cardiac arrhythmias. Putting these abnormalities together, Beard classified nine abnormal states. Extensive work with animals, with dying fetuses, and with fetuses with clinical fetal distress confirms that these patterns are often indicative of fetal distress (50). Gabert and Stenchever found that variable or late decelerations appeared in 38 to 42 per cent of high-risk labors (51).

Monitoring by auscultation provides only a small part of the information on fetal status that is available through other methods (1,16,19,34,37,46,52-54). For example, Kelly and Kulharni state that traditional monitoring samples only 1 to 2 per cent of the information on fetal status and uterine efficiency (55). Moreover, a large collaborative study showed that FHR alterations detected by auscultation were only clearly indicative of serious difficulty at the extremes (56). Strong agreement has been documented between auscultation and EFM for severe variable decelerations, the most ominous finding (57). However, agreement was not good for less serious patterns, and beat-to-beat variability and late decelerations could not be determined by auscultation. Because of these facts, EFM is often considered to be superior to auscultation. At the same time, there are those that feel that even late decelerations can be detected with the fetoscope (58), but not reliably (59). Yet, the
actual value of this added information must be evaluated.

The most reliable and valid method of following the FHR is by internal monitoring of the fetal ECG (53). The mother's ECG is sometimes picked up, especially if the fetus dies unexpectedly (60-61). External monitoring by Doppler ultrasound is not as reliable (17, 62-63). Women who are obese, uncomfortable, or experiencing a crisis are often difficult to monitor externally, and thus the technique often fails just when the data are presumed to be most needed. This difficulty with external monitoring has prompted the shift to the internal monitor. Likewise, phonocardiography cannot compare to internal monitoring of the fetal ECG in either reliability or validity (45).

The validity of EFM has been studied by comparing it to the infant's Apgar score at birth. The Apgar score is probably the most accurate routinely available measure of newborn status, but does have problems of validity. At the extreme, with scores of 0 or 1, it predicts death or morbidity quite well. However, in the middle ranges, from 5-7, it is not such a valid measure (64). Apgar scores have also been correlated with clinically diagnosed fetal asphyxia, but asphyxia does not necessarily lead to a low Apgar score. In one study only 40 per cent of the fetuses with moderate asphyxia had low Apgar scores at birth, although 80 per cent of those with severe asphyxia had a low Apgar score (65). At the same time, clinical decisions must be made on the best measures available, and the evaluative literature of EFM depends heavily on the Apgar score.

Using Apgar scores to predict the outcome, EFM is not a precise measure of fetal distress (66). In a study of 749 deliveries, Gabbert and Stenchever found 8.6 per cent false negatives (normal FHR tracing with Apgar <6) and 33.7 per cent false positives (abnormal FHR tracing with Apgar ≥6) (51). The former gives one a false sense of reassurance, whereas the latter could lead to inappropriate interference with the labor, especially cesarean section. Schiffin and Dame made similar observations, finding that a normal FHR pattern predicted a high Apgar score with 93 per cent accuracy, but that the abnormal pattern was much less specific, predicting with only 43 per cent accuracy with the one minute Apgar and only 20 per cent accuracy with the five minute Apgar (66). They concluded, "... it appears that the major value of FHR monitoring lies in the prediction of an apparently normal neonate regardless of the level of his heart rate." O'Gureck et al reported numbers of infants born depressed with different FHR patterns; their figures show a clear lack of specificity (67). Saldana and his colleagues found that less than 50 per cent of fetuses manifesting variable deceleration evidenced any clinical correlates, and that deceleration was also nonspecific (68). Shenker reported that 68 per cent of the babies with late deceleration were born without signs of depression (69). Thomas found that only half the fetuses with late deceleration had other signs of compromise (70). Finally, Goodlin found that deceleration patterns alone led to diagnoses of fetal distress in 4.5 times as many infants as were subsequently confirmed by overall clinical evaluation and Apgar score (71).
Fetal distress can also be defined by changes in the FSB pH, although defining the normal range for FSB pH has proven difficult. Approaches have included looking at the pH determinations of normal and abnormal babies as defined by Apgar scores, EFM patterns, and clinical status. Twelve studies were reviewed by Lunkey et al., who found that the lower limit of normal was usually a pH of 7.22 to 7.27, but that one study had a normal as low as 7.15. These investigators suggested 7.25 as the cut-off point (72). Saling carried out the largest studies, and found normal infants with FSB pH values as low as 7.21 (52). By using two standard deviations from the mean to define limits of normality, Saling's group suggested that 7.20 to 7.24 be considered as pre-pathologic, and below 7.20 as definitely pathologic (73).

Studies show that the FSB pH accurately reflects the acid-base balance of the fetus (52,74). Saling compared capillary and umbilical cord blood pH, and found a low incidence of error, which he stated could lead to unnecessary operative interference in 2.2 per cent of infants (52). He also concluded that capillary blood gave reliable information about the acid-base balance of the fetus. On the other hand, other investigators found FSB sampling to be less reliable, and that a fetus with a normal FSB pH had a 1 in 7 chance of being misdiagnosed as acidic (75). In addition, technical problems with FSB sampling due to maternal position, drugs such as oxytocin, peripheral vasoconstriction in the fetus, scalp edema and mechanical errors can produce pH alterations (65,76-77).

FSB pH has also been correlated with Apgar score. The rates of false negatives (normal pH with low Apgar) vary from 6 per cent (78) to nearly 50 per cent (54), but usually range from 10 per cent to 25 per cent (Table 1). At the same time, false positives (abnormal pH with normal Apgars), are found in much greater numbers, ranging as high as 84.6 per cent (81), and usually found in 20 per cent to 50 per cent of FSB pH readings (Table 1). Similar FSB pH readings are found in fetuses with and without distress with significant overlap (81, 83). Roux et al observed five fetuses who had a normal FSB pH shortly before birth but were depressed (84). Caldeyro-Barcia found a lower oxygen saturation with type II dips in the FHR record (late decelerations), but the ranges overlapped considerably (85).
EFM combined with FSB sampling

The lack of sensitivity and specificity* of both tests has led to their more frequent use together. Beard examined the results of using both EFM and FSB pH in this manner in 270 high-risk patients. Of the 68 of these patients who had low Apgar scores (<7) at birth, 46 had perfectly normal FHR tracings and normal FSB pH sampling, and only 22 of the 68 were abnormal on both tests. Seventeen patients were abnormal on both tests but had normal Apgar scores at birth. Therefore, even when used together the tests are imprecise with 44 percent false positives and 19 percent false negatives (50).

In summary, results reported largely using Apgar score as the validation for test results show that both FHR recording and FSB sampling, used separately and together, have alarming rates of false positives and false negatives even in the most skilled hands. Many authors have reported correlations of abnormal FHR patterns (such as late decelerations) and low Apgar scores (89), clinical status at birth with pH (90), and so on. But it should be noted that few of the authors whose figures are quoted discuss the low sensitivity and specificity of this procedure to question the value of these techniques. However, other investigators explicitly recognize the diagnostic inaccuracy of FHR monitoring (71,91) and the high false-positive rate (92). The physiologic basis for FHR changes is not known (93).

---

* Sensitivity is the extent to which fetuses that are abnormal are correctly classified. Specificity is the extent to which normals are correctly classified.
The predictive value (PV) of a diagnostic procedure such as EFM relates sensitivity and specificity to the occurrence of the disease or condition in the population studied (94). The PV of a positive test is the percentage of the time that a positive test denotes a true abnormality. Similarly, the PV of a negative test is the percentage of the time that a negative test will detect a person who is actually unaffected by the disease or condition.

Table 2 displays a variety of predictive values over a range of sensitivities and specificities. The three prevalence rates listed for demonstration represent current estimates of neonatal mortality (10.20/1000 live births) and 5 minute Apgar scores less than 7 (about 5 per cent) (21,51,95).

Assuming that the sensitivity of EFM is 80 per cent and the specificity is 90 percent (Table 2), the PV of a negative or normal test is 99.5 per cent. This suggests that a normal tracing very accurately predicts a favorable outcome in the infant. However, such a conclusion reflects the low prevalence of the disease rather than the quality of the test as a diagnostic tool. The PV of the normal test would remain over 99 per cent even if the specificity were as low as 50 per cent.

Assuming relatively high rates of sensitivity of 80 per cent and specificity of 90 per cent, the PV of a positive or abnormal test is only 14.0 per cent. This indicates that an abnormal EFM pattern incorrectly predicts outcome 86 percent of the time. Finally, if the neonatal mortality were to decrease to 10/1000, the PV of an abnormal tracing and the utility of EFM would also decline.

Even if neonatal asphyxia were to cause a preventable disability in 5 per cent of infants, and all of these could be detected by EFM, the PV of an abnormal test would still be only 50 percent, meaning that half of abnormal tests would provoke inappropriate anxiety or unnecessary intervention. When trying to predict an uncommon event like neonatal mortality, the large number of false positives found with EFM limits its clinical application. A diagnostic tool must be very specific (>99 more (7)).

Even if neonatal asphyxia were to cause a preventable disability in 5 per cent of infants, and all of these could be detected by EFM, the PV of an abnormal test would still be only 50 percent, meaning that half of abnormal tests would provoke inappropriate anxiety or unnecessary intervention. When trying to predict an uncommon event like neonatal mortality, the large number of false positives found with EFM limits its clinical application. A diagnostic tool must be very specific (>99 more (7)). The impact on EFM on CSR is difficult to evaluate.

One type of institutional report found the EFM group to have a consistently higher CSR than the auscultated group (20-21,55,109-111). For example, Paul reported 7 per cent of deliveries ending in cesarean sections in the auscultated group, with 16 percent ending in cesarean sections in the EFM group (20). Another type of report has shown doubling of the CSR after the introduction of EFM (15,26,43,47,112-116). A typical result is reported by Koh and his colleagues, who found a CSR of 6.4 per cent in 1971 and 12.5 per cent in 1973 (43). These rises are steady and do not generally level off by the time of the report.

Two institutions in Britain reported decreases in the CSR after the introduction of EFM (116-117). Beard and his coworkers, for example, found a decrease from 9 percent in 1972 to 5.8 per cent in 1974, and reported that this was due to the addition of FSB sampling to EFM, a technique that the author felt avoided incorrect diagnosis of fetal distress, thereby avoiding unnecessary cesarean sections (116). In Germany, Safling reported no change in CSR after the introduction of FSB (52).

However, before-and-after studies such as those cited above fail to account for secular trends that might explain the rise of CSR independently of EFM. In addition, these results are complicated by the fact that high-risk pregnancies are more likely to end in cesarean section and are also more likely to result in increased perinatal morbidity and mortality.

Perhaps the most impressive result was obtained in two randomized, controlled, clinical trials (RCTs) from Colorado. In the first study, one high-risk group was monitored by auscultation and had a CSR of 6.6 per cent, and the other was monitored electronically and had a rate of 16.5 per cent (see below) (118). A second RCT assigned 690 high-risk women to one of three study groups: auscultated, EFM, and EFM with FSB sampling (102). Once again, the route of delivery was significantly different in the three groups: 6 per cent of the auscultated group was delivered by cesarean section, whereas 18 percent of the EFM and 12 per cent of the FSB sampled had a cesarean section. Thus, FSB sampling did apparently prevent some unnecessary cesarean sections, but the combined procedure was still associated with a doubling of the CSR. Two other RCTs also showed more than a doubling of the CSR in the EFM group compared to the auscultated group (119-120).

Despite these results, EFM is not entirely responsible for the increase in cesarean sections. Indications for
cesarean section such as breech presentation have been liberalized markedly in the past decade. Hughey attempted to analyze one institution's experience and found that the rise from 5.5 per cent in 1969 to 10.8 per cent in 1975 was largely due to breech deliveries and to the diagnosis of dystocia, which increased from 10 per 1000 deliveries to 37 per 1000. Only 18 per cent of primary cesarean sections at most were found attributable to the diagnosis of fetal distress (121). Paul and Hon found a decrease in cesarean sections for fetal distress after introduction of EFM, although the overall CSR increased (25). Haddad and Lundy, on the other hand, recorded a striking increase in primary cesarean section for cephalopelvic disproportion (CPD) (2-fold), breech (50-fold), as well as fetal distress (70-fold). In that hospital, CPD was the indication for nearly 50 per cent of the cesarean sections, while fetal distress was the second most frequent indication (17 per cent) (99). Similarly, other investigators have found fetal distress and breech presentations to be the primary causes of increased numbers of cesarean section in their institution (21,55,122-126). An RCT determined that 60 per cent of the increase in cesarean sections with EFM was due to the diagnosis of fetal distress (127).

The complex causes for the rise in CSR have been discussed in a recent report (127). These causes include CPD, breech, fetal distress, repeat cesarean section, alterations in residency training programs, the more interventionistic stance of the obstetric community, and the financial incentives which may encourage cesarean sections. The reported increase in the diagnosis of CPD as an indication for cesarean section is difficult to explain. Using specific criteria, O'Driscoll et al found that CPD occurred in about 2 per cent of 1000 primigravid labors (128). Yet, in many reports CPD is the primary cause of cesarean section and is the primary indication for ½ to ¾ of the operative procedures. Fetal distress, on the other hand, while increasing markedly as an indication, accounts for only about 1/6 of the primary indications for cesarean section. Nevertheless, data from the RCTs (102,118-120) suggest that half of the rise in the CSR could be attributed to EFM and, thereby, one would assume fetal distress. This inconsistency could be explained by a change in attitude of the obstetrician towards the delivery with the increased mechanization associated with EFM (129-130).
Preventable stillbirths and neonatal deaths. The use of EFM is based on the supposition that diagnosing fetal distress and intervening aggressively can make a significant difference in perinatal mortality and morbidity. This section examines causes of stillbirth and neonatal death to determine if some deaths are preventable.

High risk. Since most pregnancies have a favorable outcome, high-risk pregnancies are routinely singled out for special care. Goodwin developed a high-risk scoring system which included obstetric history, present pregnancy, and gestational age as well as known risk factors such as breech presentation and meconium staining (131). Goodwin's score correlates with Apgar scores, but unexpected events often do occur (132). Many other investigators have developed their own systems (11,17,96,120,133-136), and report from 16 to 50 per cent of their patients as high-risk. Unfortunately, a significant proportion of perinatal deaths cannot be predicted by a high-risk score.

One study showed that 30 percent of patients accounted for 60 per cent of abnormal outcomes (137). This is consistent with findings that 18 per cent of patients classified as high-risk prenatally, were reclassified as low-risk based on intrapartum factors, while 20 per cent initially classified as low-risk were reclassified as high-risk based on intrapartum factors. The latter group suffered a 24.3 per cent neonatal morbidity and a perinatal mortality rate of 35 per 1000 (134). Fetal distress and abnormal FHR tracings during labor are more common among fetuses with intrauterine growth retardation, the small-for-gestational-age fetus (138-140). EFM is suggested as routine for this group (141), which suffers increased rates of death and illness; in particular, neurologic sequelae such as cerebral palsy, abnormal EEG, and learning difficulties (152). Management of this group is difficult, and a high CSR is to be expected (140). At the same time, early delivery by cesarean section frequently leads to respiratory distress syndrome (RDS) (140) (see below).

Baird and his co-workers examined the clinical records of more than 1000 stillbirths to determine the cause of death (143) and found, as others have (21), that prematurity is the major contributor to stillbirth; they concluded that skillful intervention could lessen risk in up to 46 per cent of the patients studied. Prematurity is also the major contributor to neonatal mortality (23,144-145).

Other investigators have examined specific causes of perinatal mortality. A large collaborative study examined 83 full-term stillbirths and found definite causes in 43, few avoidable with EFM. The remaining 40 deaths could not be explained, although 10 of these 40 had fetal bradycardia and 15 had fetal tachycardia (146).

Alberman, in a study of causes of perinatal mortality among over 16,000 deaths, found the most common causes for stillbirth were intrapartum factors such as labor, cord complications, and placental complications (29.9 per cent), and antepartum factors such as placental infarction (24.2 per cent), congenital malformation (21 per cent), and toxemia (9.6 per cent). Intrapartum factors included both acute fetal distress, where prompt and effective intervention can save the baby, and chronic fetal distress, where the outcome is generally not expected to be favorable. Intrapartum complications were especially associated with multiparity, thus reducing parity through birth control has perhaps had the greatest effect. In the neonatal group, over 40 per cent of deaths were related to immaturity; 28.2 per cent to intrapartum factors; 18 per cent to congenital malformations (147). The contribution of multiple causes to perinatal mortality found here and in other studies (148) underscores the fact that there is a limited number of cases where intervention during labor and delivery could be expected to be helpful.

Infant and perinatal mortality. The most frequently described outcome from EFM is a decrease in infant and perinatal mortality rates. Rates for the entire United States dropped from 29.2 per 1000 in 1950, to 20.0 in 1970, and to an estimated 14.0 in 1977. Similarly, the death rate of infants ≤ 28 days of age has fallen from 13.0 per 1000 in 1973 to an estimated 9.8 in 1977 (149-150). Perinatal mortality fell from 32.5 in 1950 to 20.1 in 1973 (151). The neonatal mortality rate in New York City fell from 20 in 1962 to 15 in 1971 (107). Reported deaths due to asphyxia of the newborn fell 25 per cent from 1976 to 1977 (149).
A large proportion of this fall in infant and perinatal mortality has been attributed to the introduction of EFM (15,20,23,26,47,51,87,108,114-116,121,152-153). Perinatal mortality rates typically have been reported as around 50 per 1000 until 1969, when the use of EFM became widespread, and then progressively falling to 21 per 1000 in 1974 (20). Paul, and Hon compared mortality rates among infants over 1500g and found that the high-risk, EFM patients did slightly better than the low-risk, non-EMF patients, but that the difference was not statistically significant (154). However, these results are not completely consistent with those of other studies. Some investigators, for example, have found either no difference or a slight rise in death rate during the period after introduction of EFM (43,113).

These impressive results are subject, however, to a variety of questions. Increased attention at delivery by itself may improve neonatal morbidity and mortality. The use of contraception and abortion, better nutrition, better prenatal and obstetrics care, and patient education (many of these made available because of increasing public funds) have led to fewer births, lower average parity, fewer very young and very old pregnant women, and fewer infants under 2500g (8,18,84,107,121,155). Data from one study showed a marked drop in neonatal and perinatal mortality prior to the introduction of EFM with no significant change subsequent to the widespread use of EFM in that institution (99). Within obstetric practice itself, amniocentesis, broadened indications for cesarean section, better management of complications, improved anesthesia and resuscitation, use of the lecithin-sphingomyelin (L/S) ratio to determine fetal maturity, analysis of urinary estriols, and improved delivery technique have probably contributed to decreased mortality (15,20,84,121).

In addition, the comparability of populations and the relative quality of obstetric care are not explicitly described in these before-and-after studies, and one must be concerned that an unequal distribution of possible confounding factors such as precipitous deliveries might be reflected in neonatal morbidity and mortality figures. Interpretation of these results is complicated further by the fact that EFM often only involves high-risk patients and the differential impact of EFM on high- and low-risk patients is not clearly delineated.

A few reports break down the results by birth weight, to examine the common assertion that EFM is most useful in low birth weight babies. In one study low birth weight accounted for 60 percent of neonatal mortality, and the EFM fetuses in this group did better (154). Two other groups reported substantial falls in intrapartum stillbirth associated with EFM, but these findings were inconclusive because of small sample size and lack of controls (22,156). Observations in Vermont supported a conclusion of possible benefit in the low birth weight group. Perinatal death rates from 1965 to 1975 were unchanged in that state in the over 2500g group, while falling significantly in that period in the low birth weight group (157).

It is recognized in the obstetric community that these results do not constitute scientific proof of benefit of EFM in reducing death (84,158-159), and that RCTs are necessary to resolve the questions (1,18,20,24,44). Only four RCTs have been carried out to date. The first RCT was performed in Colorado by Haverkamp and his coworkers among 483 high-risk pregnant women (118). The patients were randomly assigned either to an ausculated group or an EFM group (FB sampling was not done). Infants <1500g were excluded from the study. Both groups were cared for by the same obstetrical staff. The ausculated group had internal EFM done, but the results were not accessible to those caring for the patients. The pediatric house staff determined Apgar scores. Neonatal mortality was noted, and infant morbidity during the first 72 hours was recorded. The records from the pediatric clinic were reviewed at six weeks. The two groups were very similar in outcome, but as mentioned above, the EFM group had a CSR of 16.5 per cent, compared with the CSR of 6.6 per cent in the ausculated group.

The study conclusions have been criticized for several reasons. First, in early labor 16 late decelerations or severe variable decelerations were observed in the EFM group, while only five were seen in the ausculated group. This raised the question of comparability of the two groups. Secondly, the study design required rather close nurse-patient contact in the ausculated group. For this reason critics have felt that the outcomes may reflect more the intensity of nurse-patient relationship than the relative impact of auscultation versus EFM.
Furthermore, the study was not a blind study, so that investigator bias and the Hawthorne effect* could have influenced the results. Finally, and perhaps most important, the findings were consistent with a small degree of benefit which might not be found in a clinical trial of this size (see below).

Haverkamp and his coworkers subsequently undertook a second trial among 690 women randomly placed into three groups: ausculated, EFM, and EFM with FSB sampling (see above) (102). In response to the criticism about the intensity of nursing care in the ausculated group during the previous trial, all women were attended by individual study nurses. Fetal distress was diagnosed by standard techniques, and in the FSB-sampled group, a pH value of ≤7 was considered as an indication for immediate delivery. Again, the study was not blind. The three groups were quite similar in make-up by demographic characteristics and pregnancy status. The outcomes of the three groups, measured by a variety of techniques, were similar, except for an increased CSR in both EFM groups as noted previously.

A third clinical trial was done in Australia on 350 high-risk patients randomly allocated to a control or “intensive care” group. The control group was managed by auscultation, the intensive care group by EFM and FSB sampling. The two groups were similar by several measures of risk (120). The number of neurologic sequelae was significantly higher in the control group (13 vs. 2). In addition, the pH, PO2 and the PCO2 done at delivery on cord blood were significantly better in the intensive care group. However, perinatal mortality and Apgar scores were not significantly different between the two groups, and there was a significant increase in the maternal infection rate in the study group. The CSR was 22.3 per cent in the study group and 13.7 per cent in the control group. However, six patients in the study group had had previous cesarean section, and when they were removed, the difference in CSR was not significantly different.

This study, done in a center where EFM has been extensively investigated, was not a blind study, and was subject to investigator bias as well as Hawthorne effect. Indeed, during the study one of the eight participating obstetricians withdrew his patients because he felt it was unethical not to provide EFM. Moreover, the fact that the study group had six patients who had had a previous cesarean section while the control group apparently had none raises the question of comparability of the two groups. Finally, there was a potential ascertainment bias in that FSB sampling was available to the control group only during regular laboratory hours (I. Chalmers, personal communication).

Finally, an RCT done in England studied EFM in 504 low-risk patients who were randomized into an auscultated group and an EFM with FSB sampling group (119). The two groups were treated by the same obstetric staff. There were no significant difference in the two groups in outcome on a variety of measures. However, the CSR in the EFM group was more than double that in the auscultated group. The authors concluded that there were no beneficial or harmful effects as a direct result of the use of EFM. Like the other three RCTs, this study was not done in a double blind fashion, and the sample size was such that small differences in rates could not be detected.

Consequently, the four controlled clinical trials have methodologic problems and possibly conflicting conclusions as to the benefits and complications of EFM. It is noteworthy that all four RCT's showed no benefit attributable to EFM in terms of perinatal mortality. Neutra et al took another approach, retrospectively studying the data from a group of 16,529 live-born infants delivered at a large hospital over a seven-year period. They found that failure to use EFM increased the risk of neonatal death 1.4 times, but that use of EFM in the 24 per cent of labors with demonstrable risk factors would avert 83 per cent of the potentially preventable neonatal deaths (160). This finding of a small decrease in risk could be consistent with the RCT findings of no decrease, since a result this small would require a study of more than 100,000 women to find a statistically significant difference. In addition, the non-EMF patients in the low-risk group had somewhat lower neonatal death rates than did the EFM patients, so the small decrease in mortality (5
Evidence of some benefit in terms of mortality from EFM could be an artifact of the model. The study does give evidence of some benefit in terms of mortality from EFM in selected patients.

In response to these recent findings of little or no decrease in mortality from EFM, proponents have argued that the benefit is hard to evaluate because perinatal mortality is low (43,50). Supervision of active labor alone might reduce perinatal mortality rates by only 2.8 per 1000 live births, whereas the management of maternal disease, fetal malnutrition, RDS and infections could diminish the rate by 6.6 per 1000 live births (161). The neonatal period may be the critical risk period, and the neonatologist may prove to be most responsible for improved neonatal mortality rates (7).

Brain damage. The association between complications of pregnancy and subsequent motor and intellectual impairment of the child has been made in numerous investigations, and it has been estimated that 10 per cent of mental retardation is attributable to avoidable antenatal and early postnatal complications (18,162). Hypoxia is a major cause of antepartum and intrapartum death (29), and an estimated 60 per cent of cerebrally-impaired children have undergone a phase of hypoxia and acidosis (23,50,84,163). Experimental brain damage in monkeys has been attributed to hypoxia (164). Investigators recognize that a study with long-term follow-up of infants is necessary to demonstrate this benefit (20,35,62), but many seem to agree with Quilligan (1). He argued that such a study would be very difficult to do, would take at least 1500 births and 7 to 10 years follow-up to achieve 95 percent confidence limits, and then the results would be questionable because of lack of good outcome measures for the child. Quilligan and Paul estimated that 44.00% mentally retarded infants are born each year, with about half of the severely retarded individuals having causative factors associated with delivery. As a result, they estimate that universal EFM could save $2 billion a year in care for the handicapped (164).

However, these assumptions have been seriously questioned (29,44,142-165). Cerebral palsy or brain damage is usually related to prepartum or other events, not intrapartum asphyxia (8,166). Several British at-risk registers have been unable to demonstrate sequelae of hypoxia at birth (29). In addition, the lesions in monkeys produced by intrapartum hypoxia are not similar to those of cerebral palsy (29). Primate work demonstrates that the threshold severity of asphyxia required to produce brain injury is so close to that which causes fetal death that fetal asphyxia usually either leads to no apparent neurologic sequelae or causes death of the fetus (167-168). This "all-or-nothing" effect is suggested by several authors (8,29,168) and is supported by monkey studies showing no pathologic brain damage 3 to 9 months after controlled asphyxia (169). Studies of infants who had had hypoxia around the time of birth found no differences from control groups in physical or mental score at 8 months to 4 years of age (170-171).

A collaborative study followed 14,115 children from birth and examined them at one year of age. Only 1.9 per cent had a definite neurologic abnormality. Of those with a 5-minute Apgar score of 3 or less, a finding considered to indicate severe depression of the infant, only 7.4 per cent had a definite neurologic abnormality at one year (64). In another study, 15 children who had had 1-minute Apgar scores of ≤3 at birth were examined on the average of 22.4 months after birth. Neurologic examinations were normal in 10, two were borderline abnormal, and three were abnormal (172). The three abnormal children also had had intrauterine growth retardation. The authors concluded that the prognosis for asphyxiated infants is good. Thus, neurologic abnormality is an uncommon sequela to labor and delivery, even when the baby is depressed at birth. These studies also point to low birth weight as the important indicator of low scores and a possible confounding variable in the search for causes of cerebral palsy.

Thomson et al compared 31 children with severe birth asphyxia (Apgar at 1 minute equal to zero, or at 5 minutes a score of <4) with 31 normal controls matched for sex, birth weight, gestational age, and social class, and completed neurologic and psychologic follow-up at 5 to 10 years of age (173). Definite neurologic abnormalities were found in three cases, although one of these had satisfactory school performance despite bilateral deafness. There were no significant differences between cases and controls in numbers of children with borderline neurologic abnormalities, psycholinguistic quotient and
abilities, and Bender-Gestalt (although cases showed better than average performance more frequently). The presence of a neurologic abnormality seemed to be related only to time required for the infant to obtain spontaneous respiration and to abnormal behavior in the postnatal period. There was no association with birth weight, gestational age, social class, duration of fetal distress, Apgar score, and cardiac arrest.

Two other controlled studies, following mature infants for a year in one and 2-5 years in the other, measured several physical, psychological, neurologic, and behavioral characteristics. Neither showed significant long-term impact of intrapartum fetal asphyxia (65,174). Such results suggest that if the asphyxia of a mature fetus is not severe enough to cause noticeable cerebral injury and residual disability, then more subtle changes will not be manifested. Indeed, EFM may lead to rescue of the seriously ill infant with cerebral palsy or severe anomalies (8), and result in prolonged hospitalization before an inevitable early death.
The evidence for benefit from EFM is contradictory and confined to a small decrease in mortality among high-risk patients, particularly low birth weight patients. However, EFM with or without FSB sampling has been compared to auscultation in only four RCTs. With the likelihood that auscultation is of some benefit, it must be considered a strong possibility that EFM is no better than auscultation. Indeed, Goodlin and Haesslein have concluded that in a well-screened, healthy population, EFM adds little except to increase the CSR (8).

Maternal outcomes. In physical terms, EFM is of no benefit to the mother. Maternal mortality is very low, and instead of reducing risk, EFM increases risk to the mother (see below). EFM can, however, be of psychological benefit to some patients (see below).

Fetal risk from EFM. The most immediate risk to the infant from EFM is laceration by either the electrode (21,84,175-176) or by the knife that punctures the scalp (177). If amniotomy is done for internal monitoring, a prolapsed cord can result (17). Hemorrhage occurs in about 0.3% of cases (4, 13,71,84,178-179), and has been reported as arterial on one occasion (180). Scalp abscesses are fairly common, occurring in about 0.1 to 4.5 per cent (4,97,178,181-187). Severe sequelae include rectal-vaginal laceration (188), fatal hemorrhage (189), subgaleal abscess (190), osteomyelitis of the skull (191-192), gonococcal abscess (193-194), persistent bacteremia (195) and fatal bacterial and herpetic sepsis (196-197).

It is unknown whether or not ultrasound has long term effects, although clinical opinion and limited experimental evidence (198-199) suggest that it is of no risk to the mother or infant at currently used levels. Indeed, experiments have shown no evidence of chromosomal damage (200-201), or of changes in the EEG of newborns (202) after ultrasound exposure. The Food and Drug Administration is currently unwilling to say that ultrasound is without risk, and is initiating long-term follow-up studies of children as a first step to determine whether ultrasound is safe for the fetus (203).

Several clinical (95,163,204-205) and basic research (206-208) studies indicate that there are physiologic differences between babies delivered by cesarean section and those delivered vaginally. The short- and long-term implications of such differences are less evident, although delivered by cesarean section have a greater risk of morbidity and mortality (205).

The risks to the infant as a result of cesarean section center on RDS and hyaline membrane disease (HMD) (138,209-211). RDS is most common before 36 weeks of gestation, but occurs more frequently among babies delivered abdominally, even when controlling for gestational age (211). Some investigators feel that the bulk of RDS related to cesarean section is composed of the benign syndrome, transient tachypnea of the newborn (212). Others point out that if the fetal lung is mature, as determined by tests such as the L/S ratio, there will be no RDS (14,213-212), although other factors may be involved (215). Neither of these issues has been carefully studied by RCT, so the association of RDS and cesarean section must continue to be a concern. Certainly, in the emergent setting of fetal distress, L/S ratios become less useful in the decision-making equation.

Separation of mother and infant may interfere with bonding and later maternal behavior (216). EFM does not directly interfere with bonding, although the mechanized approach to birth and the separation engendered by the increased maternal morbidity following surgery may interfere. There has been at least one study suggesting an increase in child abuse among children following early postpartum separation (217).

Maternal risk from EFM. Mortality and morbidity are higher among women undergoing EFM. Lacerations of the mother or the placenta from the electrodes can occur (84,218), and uterine perforations from the catheter have been reported (21,219-220). Cesarean section is associated with a maternal mortality rate 3 to 30 times that found among vaginally delivered mothers (221-224). Of course, women-who are delivered abdominally often have independent risk factors that contribute to or cause death. The individual contribution of mode of delivery has not been specified, but a study done in Rhode Island indicated that four of the nine deaths reported in association with cesarean section could be attributed to the method of delivery (224).

Cesarean section also leads to morbidity and mortality (223,225-231) associated with anesthesia (232-233),
anesthetic psychosis (234-235), urinary tract infection secondary to foley catheters, operative trauma to other organs (236), supine hypotension (237-238), pulmonary embolus (239), sepsis (240), wound healing, hernia, bowel obstruction, hemorrhage, respiratory infection, and pneumonia. In addition, in the United States the operation usually assures a second operation if another pregnancy occurs (128). Finally, there is an increase in low birth weight babies subsequent to cesarean section (157).

Both EFM and cesarean section increase the risk of maternal infection (55,109,118, 219, 242-248), although the independent effect of EFM has not been adequately described (243,247). For example, Gibbs found that 87 of 132 women monitored internally with ruptured membranes had intrauterine infection (243). Bacteremia has been associated with both EFM and cesarean section (249). In another study women with EFM and cesarean section had a rate of uterine infection of 40.4 per cent; those with only a cesarean section had a rate of 20.4 per cent; those with EFM vaginal deliveries had a rate of 2.7 per cent; and women without EFM who delivered vaginally had a rate of 1.4 per cent (109). Hagen reported a 33.5 per cent overall rate of febrile morbidity with cesarean section, varying from 17 per cent in those with no labor and no EFM to 32.9 per cent in those with labor but no EFM to 54.4 per cent in those who had EFM labor prior to cesarean section (246). The rate of infection increases the longer EFM continues (241,245).

The increase in febrile morbidity has led to another situation which complicates the issue of maternal and child care during and following delivery. There is indication that rates of febrile morbidity are lowered by antibiotic prophylaxis (250-258). This places the mother at risk of antibiotic side effects, fosters the growth of resistant organisms, and complicates care of the newborn.

Maternal reactions to EFM. In the medical literature maternal reactions are usually expressed through the obstetrician (51). Some physicians and nurses state that mothers should be taught to accept EFM (161,259). EFM can cause anxiety (91), but many women suffer anxiety during labor, and EFM properly explained to the patient may reduce maternal concern (120). Some women complain of discomfort with the electrode, but overall acceptance has been good (260). External EFM is more acceptable than internal EFM, especially for prepared childbirth patients (15), although even external EFM interferes with Lamaze techniques (261).

Replacing nursing care with electronic devices could be expected to produce negative reactions. Haverkamp et al noted: “Very close physical contact with the patient was necessary for the nurse to auscultate fetal heart tones adequately. This was not true to the same degree with the monitored group. Nursing attention to the gravida with respect to maternal comfort, emotional support, and ‘laying on of hands’ could have a significant impact on the fetus. . . . The authors have the impression that the measuring psychological atmosphere created by personal nurse interaction and the absence of the recording machine in auscultated patients contributed to the excellent infant outcome in auscultated patients” (118). Indeed, a growing minority of women are demanding delivery without interference and are turning to midwives and home delivery to find such human support (9).

Only two groups have systematically approached the question of patient acceptance. In 1970, Roux et al reported that 23 per cent of patients experienced major discomfort, 24 per cent were frightened of the equipment and objected to the excessive pain, and there were difficulties in 50 per cent because of technical problems or lack of cooperation. On the other hand, 92 per cent liked the medical and paramedical support associated with EFM. Fourteen per cent said that they would not be electrically monitored again (84).

The second study which consisted of structured interviews with 25 women found that 14 have overall positive responses to EFM monitoring, and 10 had mixed or negative responses (202). Women with a problem in a previous pregnancy were more positive. The positive group stressed the comfort and protection, relieved anxiety, and the electronic monitor as an aid in cen-
munication, husband-wife interactions, and "mastery." The negative group developed competitive feelings toward the electronic monitor, found it a source of discomfort, annoyance, anxiety, and fear. Women who had had prior uncomplicated labors and healthy children especially tended to resent the mechanization (263).

A negative maternal reaction could have an adverse effect on the fetus. Monkey studies suggest catecholamine release provoked by maternal stress leads to bradycardia and hypotension in the fetus, subsequently causing episodic aggravation of underlying fetal asphyxia (264). The physiologic response to maternal anxiety could, for example, explain the higher rate of abnormal FHR patterns during early labor in the EFM patients in the first Colorado RCT.
Improvements in EFM, both in technique and interpretation, are possible. One may soon be able to use telemetry to reliably record the fetal ECG from the mother’s abdomen, obviating the need for direct electrodes (5,265). If not, fetal scalp electrodes can be improved (24). Improved Doppler ultrasound may make external monitoring more reliable (265). Better indirect methods for monitoring uterine contractions are being developed (267).

A great deal of work is being done to improve interpretation of FHR tracings. Improving display and alarm systems will be helpful (24,198,268-269), and curves generated by the EFM are being studied to gain a better understanding of them (270-272). These studies can be used to develop computer interpretation of the tracings, which is already approximately as accurate as a trained expert (273-277). Continuous pH monitoring with glass electrodes may become routine (28,45,278-279). Remote monitoring by telemetry is becoming more and more reliable and would allow more free movement by the woman in labor (24,27,267,280-282).

Finally, a variety of other methods of monitoring have been suggested as having promise. These include monitoring fetal arousal (158,283), fetal movements (272, 284-285), fetal respiration (35,286-288), and fetal EEG (274,289-290). It has also been suggested that monitoring all of pregnancy is necessary (28), or that fetal blood testing before labor might be useful (267). On the maternal side, monitoring decidual or endometrial blood flow has been suggested (291). Uterine hypoperfusion is another potential indicator of fetal distress (292).

The costs of EFM may be estimated by examining both direct and indirect costs (Table 3). The direct costs are the medical care costs of delivering EFM services themselves. The indirect costs are those of complications affecting either mother or child from EFM, including the additional medical care costs from associated cesarean section. As discussed previously, some newborns and mothers die after EFM and delivery. In addition, morbidity results from such conditions as scalp abscess and RDS in the infant and pelvic infection in the mother.

In estimating the direct costs of EFM, it is assumed that one half of the 3.2 million deliveries per year are electronically monitored, and that the use of EFM adds $50 to the cost of each delivery. Direct costs independent of FSBI sampling had previously been estimated at $33.50 (164).

The number of cesarean sections done in the United States has risen from 160,000 in 1965 to 353,000 in 1975, an increase of 193,000 (105). Based on the RCTs, half of these additional cesarean sections, or 96,500, can be attributed to EFM. The additional cost if a cesarean section is done instead of a normal delivery has been estimated to be $2300.

If 1.6 million deliveries were electrically monitored, an incidence of scalp abscess of 0.5 per cent would result in 8,000 abscesses. A scalp abscess requires 10 days of hospitalization for the infant (186), or an additional 7 days. It is estimated that the hospitalization of a newborn costs $100 a day; the cost of physician services, antibiotics, and so forth is disregarded.

The incidence of RDS in cesarean sections is about 5 per cent, so 4825 cases would result from 96,500 operations. However, about 1255 cases would have resulted from vaginal delivery (211), so 3570 cases can be attributed to EFM. It costs between $2700 and $3400 to treat a case of RDS, whether the infant dies or not (209-210). A figure of $3000 per case is used in this analysis.

The incidence of death from RDS is 1.3 per cent (129), so 96,500 cesarean sections would result in 1255 deaths. If one assumes that 20 per cent of the 1255 cases calculated above would have died in any case following vaginal delivery (211), 1000 deaths can be attributed to EFM. A cost can be put on these deaths by using the present value of lifetime earnings approach, discounting
at 10 per cent. The value of lifetime earnings for a male under the age of one year is $37,659 and for a female $29,802 (292). It is assumed that the deaths are half male and half female.

A significant amount of morbidity results from cesarean section. Relatively uncommon complications such as herniation, thromboiphlebitis, and so forth have not been projected. However, the rate of maternal infection is quite high (40 per cent) and results in an average of six additional days of hospitalization (294). Using $190 a day as the cost of the hospital beds and ignoring physician and drug costs (295), this morbid event costs $44 million a year.

Three per cent of vaginally delivered women who are electronically monitored become infected (109) about double the rate in vaginal deliveries without EFM. An additional 19,500 infections result from 1.5 million EFM deliveries. This group stays an estimated three additional days in the hospital at a cost of $11.1 million, again disregarding the cost of physician services and drugs.

The maternal mortality attributable to cesarean section is 3.1 per 10,000 (224). Thus, 30 deaths result from 96,500 cesarean sections. Using the present value of lifetime earnings approach, assuming that these women are age 30-34, and using a 10 per cent discount rate, the cost of these deaths is $114,057 each (293).

The estimate of total cost of $411 million per year compares to $80 million spent annually for all public and private childhood immunization programs (A. Hinman, personal communication).

Although evidence that EFM prevents mental retardation is lacking, this contention cannot be completely ignored. At the same time, one cannot dismiss the possibility that EFM keeps alive infants destined to live with intellectual handicaps who might otherwise have died. Because the data are not available to allow an accurate estimate of the cost or benefit to society of EFM with regard to mental retardation, such estimates have not been included in the cost analysis of EFM. Nonetheless, increases or decreases in the number of mentally retarded children born could have a profound effect on the calculations presented in this section.

Severe mental retardation has been estimated to cost society approximately $250,000 during the lifetime of each affected individual; milder handicaps cost an estimated $30,000 per patient (164). If 1 per cent of live-born children suffer measurable mental retardation, 30,000 such infants will be born in this country each year, half of them with severe impairments. The cost to society of these infants is then a little more than $4 billion over their lifetimes. (These calculations were done in 1974, so inflation would make them considerably higher now.) It has been estimated that up to one-half of mental retardation can be eliminated if serious intrapartum asphyxia is detected, immediate delivery is successfully undertaken, an appropriate post-partum resuscitation and care is employed (164). This would save society about $2 billion over the lifetime of those handicapped individuals. On the other hand, an increase of 50 per cent of those with mental retardation would cost about $2 billion for those individuals. Without further study, the effect of EFM on the costs of mental retardation to the society cannot be estimated.
EFM was developed to prevent damage to the fetus, especially from asphyxia, during labor and delivery. Empirical work has demonstrated that it is possible to record accurately the FHR either by ultrasound or by direct ECG-monitoring. In recent years, FSB pH sampling by puncture of the fetal scalp has become part of the monitoring procedure. Auscultation of the fetal heart has been replaced by interpretation of patterns in the electronically-recorded FHR. The obstetric literature reflects the commonly held belief in medicine that more information will lead to a better outcome. The technical advances required and the demonstration that reliable recording could be done seems to have blinded most observers to the fact that this additional information will not necessarily produce better outcomes. Nonetheless, many obstetricians justify the procedure because they believe it to be a reliable indicator of normality.

Careful review of the literature indicates little increased benefit from EFM compared to auscultation. This is not surprising, given the lack of precision of EFM for the diagnosis of fetal distress, and the general difficulty in separating normal fetal stress during labor from fetal distress. If EFM has benefit, it appears to be for low birth weight infants, but no RCT of its use in this group has been carried out.

The risk from EFM is substantial, especially but not wholly through the increased CSR that its use apparently engenders. We estimate the addition to the annual cost of childbirth to be $411 million if 50 per cent of deliveries are monitored electronically.

Because of these risks and costs, a few signs of dissent have appeared (8,14,29,296). Hohe, for example, takes the cautious view expressed by his own patients in complications, and it is unfair to subject all patients to the routine use of EFM (9). In light of the increasing concern with the costs of medical care, all of society, especially the medical profession, must be concerned with the widespread use of an expensive technology such as EFM in the absence of scientific evidence as to its benefits.
Table 1. Diagnostic precision of fetal scalp blood sampling using Apgar score as the measure of outcome.*

<table>
<thead>
<tr>
<th>Source</th>
<th>n</th>
<th>False negatives (%)</th>
<th>False positives (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>d, Filshie, et al. (1971) (50)</td>
<td>279</td>
<td>19.1</td>
<td>42.1</td>
<td>32.4</td>
<td>92.4</td>
</tr>
<tr>
<td>e, et al. (1970) (79)</td>
<td>355</td>
<td>13.9</td>
<td>30.3</td>
<td>62.6</td>
<td>69.4</td>
</tr>
<tr>
<td>Art (1969) (80)</td>
<td>295</td>
<td>23.6</td>
<td>42.2</td>
<td>30.6</td>
<td>90.1</td>
</tr>
<tr>
<td>a, Rama and Merkat (1970) (78)</td>
<td>208</td>
<td>9.0</td>
<td>56.7</td>
<td>44.8</td>
<td>90.5</td>
</tr>
<tr>
<td>Khazin, and Paul (1969) (81)</td>
<td>214</td>
<td>15.4</td>
<td>69.2</td>
<td>39.0</td>
<td>79.2</td>
</tr>
<tr>
<td>minute Apgar &lt;7, pH &lt;7.20</td>
<td>214</td>
<td>5.6</td>
<td>84.6</td>
<td>47.1</td>
<td>77.7</td>
</tr>
<tr>
<td>Khazin, and Paul (1969) (81)</td>
<td>78</td>
<td>10.2</td>
<td>55.6</td>
<td>57.1</td>
<td>84.1</td>
</tr>
<tr>
<td>od, et al. (1967) (54)</td>
<td>118</td>
<td>47.4</td>
<td>21.7</td>
<td>28.6</td>
<td>90.9</td>
</tr>
</tbody>
</table>

*In many studies have been recalculated because of inaccurate estimates of false positives and false negatives.

Table 2. Predictive values of a diagnostic test over a range of sensitivities and specificities when actual prevalence of a condition varies from 1 to 5 per cent.

<table>
<thead>
<tr>
<th>Prevalence Rate</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Predictive Value Negative (%)</th>
<th>Predictive Value Positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal Mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) 10/1000</td>
<td>80</td>
<td>90</td>
<td>99.8</td>
<td>7.5</td>
</tr>
<tr>
<td>2) 20/1000</td>
<td>50</td>
<td>90</td>
<td>99.6</td>
<td>3.9</td>
</tr>
<tr>
<td>3) 20/1000</td>
<td>80</td>
<td>80</td>
<td>99.5</td>
<td>7.6</td>
</tr>
<tr>
<td>4) 20/1000</td>
<td>80</td>
<td>90</td>
<td>99.5</td>
<td>14.0</td>
</tr>
<tr>
<td>Abnormal 5” Apgar Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5) 50/1000</td>
<td>95</td>
<td>95</td>
<td>99.7</td>
<td>50.0</td>
</tr>
</tbody>
</table>

Table 3. Estimated financial cost of fetal monitoring, 1977-1978, United States.

<table>
<thead>
<tr>
<th>Cost Description</th>
<th>Cost (in millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Cost of Monitoring</td>
<td>$80 million</td>
</tr>
<tr>
<td>Direct Cost of Cesarean Sections</td>
<td>$222 million</td>
</tr>
<tr>
<td>Indirect Costs of Fetal Monitoring—Neonatal Mortality</td>
<td>$34.2 million</td>
</tr>
<tr>
<td>Neonatal Morbidity Scap Abscess</td>
<td>$5.6 million</td>
</tr>
<tr>
<td>Respiratory Distress Syndrome</td>
<td>$10.7 million</td>
</tr>
<tr>
<td>Neonatal Mortality</td>
<td>$34.2 million</td>
</tr>
<tr>
<td>Indirect Costs of Fetal Monitoring—Maternal Mortality</td>
<td>$411 million</td>
</tr>
<tr>
<td>General Morbidity from Cesarean Section</td>
<td>$44 million</td>
</tr>
<tr>
<td>Maternal Infection from Cesarean Section</td>
<td>$44 million</td>
</tr>
<tr>
<td>Maternal Infection from Fetal Monitoring</td>
<td>$11.1 million</td>
</tr>
<tr>
<td>Maternal Mortality</td>
<td>$3.4 million</td>
</tr>
</tbody>
</table>
References


31


National Center for Health Services Research publications of interest to the health community are available on request to NCHSR, Publications and Information Branch, 3700 East-West Highway, Room 7-44, Hyattsville, MD 20782 (telephone: 301/436-8970). Mail requests will be facilitated by enclosure of a self-addressed, adhesive-backed mailing label. These publications also are available for sale through the National Technical Information Service (NTIS), Springfield, VA 22161 (telephone: 703/557-4650). PB and HRP numbers in parentheses are NTIS order numbers. Publications which are out of stock in NCHSR are indicated as available only from NTIS. Prices may be obtained from the NTIS order desk on request.

Research Digests

The Research Digest Series provides overviews of significant research supported by NCHSR. The series describes either ongoing or completed projects directed toward high priority health services problems. Issues are prepared by the principal investigators performing the research, in collaboration with NCHSR staff. Digests are intended for an interdisciplinary audience on health services planners, administrators, legislators, and others who make decisions on research applications.

(HRA) 77-3144 Evaluation of a Medical Information System in a Community Hospital (PB 264 353)
(HRA) 77-3145 Computer-Stored Ambulatory Record (COSTAR) (PB 268 342)
(HRA) 77-3160 Program Analysis of Physician Extender Algorithm Projects (PB 264 610, available NTIS only)
(HRA) 77-3161 Changes in the Costs of Treatment of Selected Illnesses, 1951-1954-1971 (HRP 0014596)
(HRA) 77-3163 Impact of State Certificate-of-Need Laws on Health Care Costs and Utilization (PB 264 392)
(HRA) 77-3164 An Evaluation of Physician Assistants in Diagnostic Radiology (PB 266 507, available NTIS only)
(HRA) 77-3166 Foreign Medical Graduates: A Comparative Study of State Licensure Policies (PB 265 233)
(HRA) 77-3171 Analysis of Physician Price and Output Decisions (PB 273 312)
(HRA) 77-3173 Nurse Practitioner and Physician Assistant Training and Deployment (PB 271 000, available NTIS only)
(HRA) 77-3177 Automation of the Problem-Oriented Medical Record (PB 266 881, available NTIS only)
(PHS) 78-3190 Uncertainties of Federal Child Health Policies: Impact In Two States (PB 283 202)

Research Summaries

The Research Summary Series provides rapid access to significant results of NCHSR-supported research projects. The series presents executive summaries prepared by the investigators at the completion of the project. Specific findings are highlighted in a more concise form than in the final report. The Research Summary Series is intended for health services administrators, planners, and other research users who require recent findings relevant to immediate programs in health services.

(HRA) 77-3162 Recent Studies In Health Services Research, Vol. I (July 1974 through December 1976) (PB 266 460)
(HRA) 77-3183 Recent Studies In Health Services Research, Vol. II (CY 1976) (PB 279 198)
(HRA) 77-3176 Quality of Medical Care Assessment Using Outcome Measures (PB 272 455)
(PHS) 78-3193 Optimal Electrocardiography (PB 281 558)
(PHS) 78-3201 A National Profile of Catastrophic Illness (PB 287 291)
(PHS) 78-3187 Criterion Measures of Nursing Care Quality (PB 287 449)
(PHS) 78-3192 Assessing the Quality of Long-Term Care

Policy Research

The Policy Research Series describes findings from the research program that have major significance for policy issues of the moment. These papers are prepared by members of the staff of NCHSR or by independent in-

38
vestigators. The series is intended specifically to inform those in the public and private sectors who must consider, design, and implement policies affecting the delivery of health services.

(HRA) 77-3182 Controlling the Cost of Health Care (PB 266 885)

Research Reports

The Research Report Series provides significant research reports in their entirety upon the completion of the project. Research Reports are developed by the principal investigators who conducted the research, and are directed to selected users of health services research as part of a continuing NCHSR effort to expedite the dissemination of new knowledge resulting from its project support.

(HRA) 76-3143 Computer-Based Patient Monitoring Systems (PB 266 508)
(HRA) 77-3152 How Lawyers Handle Medical Malpractice Cases (HHR 0014319)
(HRA) 77-3159 An Analysis of the Southern California Arbitration Project, January 1956 through June 1975 (HHR 0012466)
(HRA) 77-3165 Statutory Provisions for Binding Arbitration of Medical Malpractice Cases (PB 264 409, available NTIS only)
(HRA) 77-3184 1960 and 1970 Spanish Heritage Population of the Southwest by county (PB 260 658, available NTIS only)
(HRA) 77-3188 Demonstration and Evaluation of a Total Hospital Information System (PB 271 079)
(HRA) 77-3189 Drug Coverage under National Health Insurance: The Policy Options (PB 272 074)

Research Management

The Research Management Series describes programmatic rather than technical aspects of the NCHSR research effort. Information is presented on the NCHSR goals, research objectives, and priorities; in addition, this series contains lists of grants and contracts, and administrative information on funding. Publications in this series are intended to bring basic information on NCHSR and its programs to research planners, administrators, and others who are involved with the allocation of research resources.

(PHS) 79-3220 Emergency Medical Systems Research Projects, 1978
(PHS) 79-3241 NCHSR Research Priorities

Research Proceedings

The Research Proceedings Series extends the availability of new research announced at key conferences, symposia and seminars sponsored or supported by NCHSR. In addition to papers presented, publications in this series include discussions and responses whenever possible. The series is intended to help meet the information needs of health services providers and others who require direct access to concepts and ideas evolving from the exchange of research results.

(HRA) 76-3138 Women and Their Health: Research Implications for a New Era (PB 264 359, available NTIS only)
(HRA) 76-3150 Intermountain Medical Malpractice (PB 266 344, available NTIS only)
(HRA) 77-3154 Advances in Health Survey Research Methods (PB 262 230)
(HRA) 77-3181 NCHSR Research Conference Report on Consumer Self-Care in Health (PB 273 811)
(HRA) 77-3186 International Conference on Drug and Pharmaceutical Services Reimbursement (PB 271 386)
(PHS) 78-3195 Emergency Medical Services: Research Methodology (PB 279 096)
(PHS) 78-3227 Effects of the Payment Mechanism on the Health Care Delivery System
(PHS) 78-3207 Health Survey Research Methods, Second Biennial Conference
(PHS) 78-3208 Drug Coverage Under National Health Insurance
(PHS) 79-3209 Health Services Research in Puerto Rico
(PHS) 79-3228 A National Conference on Health Policy, Planning, and Financing the Future of Health Care for Blacks in America

Program Solicitations

(HRA) 77-3196 Conference Grant Information
(PHS) 78-3224 Grants for Dissertation Research Support
## Abstract

This report focuses on electronic fetal monitoring (EFM) -- a technology that was developed during the 1960s and has rapidly spread into use in clinical obstetrics. The report includes a review of the extensive published literature on EFM and related subjects, as well as original calculations concerning the technique's specificity and sensitivity, predictive value as a diagnostic test, and financial costs associated with its potential risks.

## Key Words and Document Analysis

<table>
<thead>
<tr>
<th>Health services research</th>
<th>Cost/benefit analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic fetal monitoring</td>
<td>Literature on medical technology</td>
</tr>
<tr>
<td>Medical technology</td>
<td>Technology assessment</td>
</tr>
<tr>
<td>Health care costs</td>
<td></td>
</tr>
</tbody>
</table>

## Availability Statement

Releasable to the public. Available from the National Technical Information Service, Springfield, VA 22161. (Tel.: AC 703/557-4650)