



Published in final edited form as:

Semin Perinatol. 2016 August ; 40(5): 307–317. doi:10.1053/j.semperi.2016.03.008.

What We Have Learned About Intrapartum Fetal Monitoring Trials in the MFMU Network

Steven L. Bloom, MD^a, Michael Belfort, MD, PhD^b, George Saade, MD^c, and Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network

^aDepartment of Obstetrics and Gynecology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390

^bDepartment of Obstetrics and Gynecology, Baylor College of Medicine, Texas Children's Hospital Pavilion for Women, 6651 Main Street, Houston, TX 77030

^cDepartment of Obstetrics and Gynecology, The University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555

Abstract

The vast majority of pregnant women in the United States are subjected to electronic fetal heart monitoring during labor. There is limited evidence to support its benefit compared with intermittent auscultation. In addition, there is significant variability in interpretation and its false positive rate is high. The latter may have contributed to the rise in operative deliveries. In order to address the critical need for better approaches to intrapartum monitoring, the MFMU Network has completed 2 large multi-site randomized trials, one to evaluate fetal pulse oximetry and the other to evaluate fetal ECG ST segment analysis (STAN). Both of these technologies had been approved for clinical use in the U.S. based on prior smaller trials. These technologies were evaluated in laboring women near term and their primary outcomes were overall cesarean delivery for the oximetry trial and a composite adverse neonatal outcome for STAN. Both trials failed to show a benefit of the technology, neither in the rates of operative deliveries nor in the rates of adverse neonatal outcomes. The experience with these trials, summarized in this report, highlights the need for rigorous evidence before introduction of new technology into clinical practice and provide a blueprint for future such trials to address the need for better intrapartum monitoring approaches.

When first introduced, electronic fetal heart rate monitoring was used primarily in complicated pregnancies, but gradually it came to be used during most labors. In 1978 it was estimated that nearly two-thirds of American women were being monitored electronically during labor.¹ By 1998, nearly 3.3 million American women, comprising 84 percent of all live births, underwent electronic fetal heart rate monitoring.²

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

By the end of the 1970s, however, questions about the efficacy, safety, and costs of electronic monitoring were being voiced by the Office of Technology Assessment, the United States Congress, and the Centers for Disease Control and Prevention. Banta and Thacker¹ analyzed 158 reports and concluded that “the technical advances required in 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”. They attributed the apparent lack of benefit to the imprecision of electronic monitoring to identify fetal distress. Moreover, increased usage was linked to more frequent cesarean delivery. They estimated that additional costs of childbirth in the United States, if half of labors had electronic monitoring, were approximately \$400 million per year in 1979 dollars.

The *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) appointed a task force to study these concerns, and a consensus report was published in 1979.³ After an exhaustive review of the electronic fetal heart rate monitoring literature, the group concluded that the evidence only suggested a trend toward improved infant outcome in complicated pregnancies. They emphasized that few scientifically rigorous investigations had been done to address perinatal benefits. A subsequent NICHD consensus panel,⁴ convened to address the dramatic increase in cesarean births in the United States, concluded that use of electronic fetal heart rate monitoring was a contributing factor.

Almost 20 years later, the NICHD Fetal Monitoring Workshop⁵ formulated research recommendations intended to assess the reliability and validity of fetal heart rate patterns in the prevention of asphyxial brain damage. The workshop participants concluded that the effectiveness of fetal heart rate monitoring still remains to be established despite widespread use in the United States. Another reason for the failure of fetal heart rate monitoring technology to be proven beneficial is the now well-accepted non-specificity of fetal heart rate patterns to predict fetal compromise. This poor specificity of fetal heart rate pattern interpretation has resulted in a continuing search for adjunctive tests that could be used to distinguish false-positive fetal heart rate patterns.

A number of adjunctive measures have been proposed, including fetal scalp sampling for pH, fetal scalp stimulation, fetal lactate measurement, determination of fetal oxygen saturation, and monitoring of fetal ECG. The *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Maternal-Fetal Medicine Units (MFMU) Network identified intrapartum monitoring as one of the areas in need for more research, especially given that one of the major aims of the Network is to “evaluate maternal and fetal interventions for efficacy, safety, and cost-effectiveness”.⁶ In this review, we attempt to present the rationale, the findings, and our experience with two large randomized trials conducted by the Network designed to measure the efficacy and safety of two promising adjuvants to electronic fetal monitoring – fetal pulse oximetry and fetal ECG. Although neither of these trials showed benefit, we believe that these ambitious efforts helped ensure that interventions were not introduced prior to their efficacy and safety being validated and hope that this information will be valuable for future research designed to improve intrapartum fetal assessment.

Fetal Pulse Oximetry

In May 2000, the United States Food and Drug Administration (FDA) granted conditional approval of the Nellcor OxiFirst Fetal Pulse Oximetry System for use as an adjunct to electronic fetal monitoring.⁷ This technology was designed to improve knowledge of the intrapartum condition of the fetus in the presence of a non-reassuring fetal heart-rate pattern by continuously measuring fetal oxygen saturation. With this technology, a specialized sensor is inserted through the dilated cervix after ruptured membranes and positioned against the fetal face. Once in contact with the fetal skin, the device permits measurement of fetal oxygen saturation during labor.⁸

The fetal pulse oximetry system was designed based upon principles of spectrophotometry and plethysmography.⁸ The sensor contains two low-voltage, light-emitting diodes as light sources and one photo detector. One light-emitting diode emits red light (735 nm), and the other emits infrared light (890 nm). When light from each light-emitting diode passes through fetal tissue at the sensor application site, a fraction is absorbed. The photodetector measures the light that was not absorbed – that is, the light that is reflected. Because oxyhemoglobin and deoxyhemoglobin have different light-absorption characteristics – relatively less red light is absorbed by oxyhemoglobin compared with deoxyhemoglobin and relatively more infrared light is absorbed by oxyhemoglobin compared with deoxyhemoglobin – pulse oximetry employs the ratio of these differences to calculate fetal oxygen saturation during each arterial pulse.^{8,9}

A number of published observational studies in both animals and humans demonstrated a correlation between fetal metabolic acidosis and increasing duration of fetal pulse oximetry saturations below 30 percent in the setting of a non-reassuring fetal heart rate pattern.^{10–12} These studies were followed by the first randomized controlled trial of fetal pulse oximetry, which was published by Garite and colleagues in 2000.¹³ In this trial, a total of 1010 women with term pregnancies in active labor with an abnormal fetal heart rate pattern were randomly assigned to electronic fetal monitoring alone (the control group) or to electronic fetal monitoring plus continuous fetal pulse oximetry (the study group). The primary outcome was a reduction in cesarean deliveries for the indication of non-reassuring fetal status.

As shown in Table 1, the frequency of the primary outcome was significantly lower in the study group compared with the control group (5% vs. 10%, $P<0.001$).

However, as also shown in Table 1, there were no significant differences in the overall cesarean rate, the overall operative vaginal delivery rate, or the rate of operative vaginal delivery for non-reassuring fetal status. Looking at the results from a different angle, although the rate of cesarean delivery for non-reassuring fetal status was halved in the fetal pulse oximetry arm, the rate of cesarean delivery for dystocia more than doubled in this same group (9% vs. 19%, $P<0.001$). This increase in the rate of cesareans for dystocia was unexpected, and the results of the study raised several key questions. For example, were the discrepancies in the effect of the oximeter according to the indication for cesarean delivery reproducible? Was the sample size sufficient to permit assessment of infant safety in

circumstances when an obstetrician withholds cesarean delivery in the presence of an abnormal fetal heart rate because fetal oxygenation is deemed to be normal? Such questions prompted the American College of Obstetricians and Gynecologists to withhold endorsement of the oximeter for use in clinical practice until additional studies were completed.¹⁴ Similarly, final FDA approval was contingent upon the results of post-approval studies.

In 2001, the investigators of the NICHD MFMU Network began an ambitious trial to test the efficacy and safety of fetal pulse oximetry.¹⁵ The study protocol required that 10,000 women with term pregnancies all have their labors monitored with both an electronic fetal monitor and a fetal pulse oximeter. Women were randomly assigned to a group in which the fetal oxygen saturation data were made available to the managing clinicians or to a group in which the fetal oxygen saturation data were masked. This design and the large sample size strengthened the study in several ways. First, it was hypothesized that like electronic fetal monitoring, fetal pulse oximetry, if widely adopted, would inevitably be used in labors not only complicated by non-reassuring fetal heart rate patterns but also those with either reassuring or “less non-reassuring” fetal heart rate patterns. The sample size ensured that both a large number of labors complicated by non-reassuring fetal heart rate patterns and a large number of less complicated fetal heart rate patterns would be available for study. Second, the masked design permitted that the actual experiment tested the clinical utility of fetal oxygen saturation data and helped eliminate other potential confounding variables. For example, some patients with non-reassuring fetal heart rate patterns which developed in the second stage of labor could have gone on to deliver vaginally during the occasionally time-consuming minutes that it took to insert the device. Such vaginal deliveries were less a result of the utility of the fetal oxygen saturation data and more a result of the additional time required attempting to obtain such data. Third, the masked design permitted blinded ascertainment and recording of fetal oxygen saturation data in approximately 5000 laboring patients. This archived data would serve to help better elucidate the natural history of fetal oxygen saturation during labor, as well as to refine the parameters used to define abnormal. Fourth, the sample size permitted a statistically valid assessment of the safety of fetal pulse oximetry for both mother and fetus.

Following a comprehensive training phase – including centralized and local on-site training of Network physicians and nurses by educators from the manufacturer, as well as passing of both written and practical examinations by the study personnel – recruitment began in May 2002. At the third, planned interim analysis, the data and safety monitoring committee halted the trial early after concluding that, given the data collected up to that point, an adequate number of subjects had already been enrolled to address the major study outcomes.

At the conclusion of the study, a total of 27,571 nulliparous women had been screened at 13 participating academic medical centers. Of the eligible women, 5553 consented to the study and sensor insertion was attempted. Of these, 212 did not undergo randomization because in 170 the device could not be successfully placed and in 42 attempted insertion of the device was associated with a previously unreported finding of prolonged fetal heart rate decelerations. The remaining 5341 women were randomly assigned to either the open or masked group.

As shown in Table 2, the availability of intrapartum fetal oxygen saturation data had no impact on the route of delivery.

Specifically, the overall cesarean rate was similar in the open and masked groups (26.3% vs. 27.5%, $P = 0.31$) as were the rates of cesarean for non-reassuring fetal heart rate (7.1% vs. 7.9%, $P = 0.30$) and dystocia (18.6% vs. 19.2%, $P = 0.59$). The findings were the same for the 2168 women with labors complicated by a non-reassuring fetal heart rate pattern prior to randomization (Table 2). There were also no differences in neonatal outcomes between the two groups including a composite measure comprised of one or more of the following: 5-minute Apgar score of 3 or less, umbilical–artery blood pH value of less than 7.0, seizures, intubation in the delivery room, stillbirth, neonatal death, or admission to the neonatal intensive care unit for more than 48 hours.

As reviewed by Dildy,¹⁶ there have been at least four other smaller randomized studies of fetal pulse oximetry published around the same time as or after the Network trial.^{17–20} At the University of Mississippi Medical Center, Klauser and colleagues randomly assigned 360 laboring women with a non-reassuring fetal heart rate pattern to electronic fetal monitoring alone or electronic fetal monitoring plus fetal pulse oximetry.¹⁷ Like the Network trial, the overall rate of cesarean delivery, as well as the rates of cesarean for non-reassuring fetal heart-rate tracing and for dystocia were similar between the two groups. And although the decision-to-incision time was approximately 10 minutes shorter in the fetal pulse oximetry arm, there were no statistically significant differences in neonatal outcome.

Conducted at four maternity hospitals in Australia, the FOREMOST trial included 600 laboring women with non-reassuring fetal status who were randomly assigned to electronic fetal monitoring alone or electronic fetal monitoring plus fetal pulse oximetry.¹⁸ Similar to the Garite trial,¹³ these investigators also found that fetal pulse oximetry was associated with a lower rate of cesarean delivery for non-reassuring fetal status compared with electronic fetal monitoring alone (13.8% vs. 20.2%, $P = 0.042$). However, the overall cesarean rate was comparable between the two groups (45.9% vs. 48.1%, $P = 0.584$), and there were no differences in neonatal outcomes.

Two smaller randomized studies – one from Turkey with 230 patients¹⁹ and one from Germany with 146 patients²⁰ – demonstrated significant reductions in both cesarean for non-reassuring fetal status and overall cesareans or overall operative delivery in those women randomly assigned to fetal pulse oximetry. These two trials were later included in a synthesis of the available data in 2014 by the Cochrane Collaboration.²¹ The Cochrane authors concluded that “the addition of fetal pulse oximetry does not reduce overall cesarean section rates” and that “a better method than pulse oximetry is required to enhance the overall evaluation of fetal well-being in labour”.

ST Segment Analysis

Fetal electrocardiogram (ECG) ST segment analysis (STAN) was approved by the FDA in 2005 as an adjunct to electronic fetal heart rate (FHR) monitoring to determine whether obstetrical intervention is warranted when there is an increased risk for developing

metabolic acidosis.²² Clinical application of STAN to fetal monitoring began in 1979 and further refinements between 1979 and 1989 to improve signal processing capability then made it possible for clinical trials to be conducted.²³ STAN has now been available in Europe for more than two decades; and there have been several randomized controlled trials (RCTs)^{24–28} that have suggested its utility in decreasing (i) cord blood acidosis, (ii) need for fetal scalp blood sampling during labor, and (iii) need for operative vaginal delivery and emergency cesarean delivery for fetal indications. In 2006 the Cochrane Library²⁹ endorsed the use of STAN “when a decision has been made to undertake continuous electronic FHR monitoring during labour”. However, despite these endorsements, enthusiasm for this technology in the United States was muted, and there were concerns that needed to be addressed before widespread adoption of another electronic fetal monitoring modality could be considered.³⁰ None of the randomized trials published prior to 2006 had been performed in the United States, and given the different patient case mix, different health care delivery models, and different obstetrical practices in the United States and Europe, direct extrapolation of the European data to the U.S. population was not thought to be appropriate. As a result the largest randomized controlled trial of ST analysis ever undertaken was performed by the NICHD MFMU Network and the results were published in 2015.³¹

The fetal STAN system uses established principles of ECG analysis and select components of the fetal ECG signal are specifically isolated to determine the presence of myocardial ischemia.^{32–35} Fetal ECG does not provide the same perspective as an adult ECG recording because only one lead is used (scalp lead) to provide a global electrical picture of what is happening within the fetal heart as a whole, relying on the principle that the fetal ECG is a surface representation of changes in action potentials within the myocardium.

While a detailed description of myocardial electrophysiology is beyond the scope of this article, some description of the ST analysis methodology used in the STAN system may be helpful. A normal fetal ST interval is made up of an ST segment characterized by a horizontal or upward-sloping ST segment and a T wave with a constant and stable amplitude. A normal ST interval usually indicates a positive energy balance and aerobic myocardial function. Dawes and colleagues³⁶ showed that during hypoxia, fetal myocardial function and survival depends on myocardial glycogenolysis. As glycogenolysis increases, so does the amplitude of the T wave in the fetal ECG³³ and this relationship has been demonstrated to be linear.³⁴ The amplitude of the QRS complex remains relatively stable until quite late in the hypoxia/acidosis process, providing a metric against which the change in the height of the T wave is standardized—the T/QRS ratio.

In fetal guinea pigs, progressive hypoxia results in changes in the ST segment and T-wave that are clearly seen and occur within a few minutes of the initiation of hypoxia.³⁷ The fetus responds to moderate hypoxemia with a catecholamine surge, beta-adrenergic activation and myocardial glycogenolysis, all of which stimulate an increase in the T-wave amplitude. Repolarization of the myocardium (as reflected by the ST segment and T wave) is an energy-consuming process. When there is hypoxia, the energy balance within the myocytes becomes negative and the cells resort to beta-adrenoceptor-mediated anaerobic glycolysis. This pathway produces both lactate and potassium ions, and these ions affect the myocyte cell membrane potential and cause an elevation in the T-wave amplitude.³⁸ When energy

balance cannot be maintained by the compensatory mechanisms (vasodilatation and anaerobic metabolism), the endocardium becomes ischemic, altering the sequence of repolarization and direction of electrical flow. This imbalance between the endocardium and epicardium causes depression of the ST segment, with or without inversion of the T wave.³⁹ In some fetuses, certain conditions prevent the usual myocardial response of ST segment elevation and T-wave amplitude increase, and there is the development of a biphasic shape to the ST segment, with progressive depression of this biphasic ST segment as the hypoxia worsens.

Because the perfusion pressure in the endocardium is always lowest when the mechanical strain is highest and because the response of the myocardium (beta-receptor activation and enhanced Frank-Starling curve) to an increased volume load is not instantaneous, there can be delays in the repolarization, which is manifested as ST changes. Thus, any stimulus that substantially alters the balance and performance characteristics of the myocardium may result in ST depression. Conditions that may be associated with ST depression and biphasic fetal ECG waveforms include prematurity, infection, maternal fever, myocardial dystrophy, cardiac malformations,⁴⁰ chronic hypoxia, and the initial phase of acute hypoxia (when the fetus has not had enough time to develop the classic response). In addition, Yli and colleagues⁴¹ showed that ST depression occurs more frequently in fetuses of mothers with diabetes mellitus (possibly related to the higher prevalence of myocardial dystrophy in such babies).

An important concept that must be taken into account is that the STAN system is based on the ability to detect changes in the ST interval when the fetus mobilizes its compensatory mechanisms against hypoxia. A fetus that is already hypoxic, or one that has a significantly decreased capability to mount a response to hypoxia, may not show a change in the T-wave amplitude with further hypoxia. In this case, even progressive metabolic acidosis does not elicit a T-wave response. This fact highlights the importance of using STAN only in a fetus that is initially deemed to be capable of mounting a hypoxic response and underlines the adjunct nature of this system. In other words, the system can only be relied upon to act as designed when there has been accurate interpretation of the standard electronic fetal heart pattern. Therein lies a very real source of bias – the STAN technology depends upon another technology (electronic fetal heart rate monitoring and its interpretation) – a technology that in and of itself is subjective and open to interpretation.

The STAN System

The STAN system includes a fetal ECG electrode, a maternal skin reference electrode, and a microprocessor-based monitor that identifies the fetal ST segment and T-wave changes and compares them with the normal baseline values that have been individually established for that patient using the fetal heart rate monitoring strip as the basis for the normal reference indicator. This is an important distinction between this technology and others that have been used to evaluate fetal condition in labor. The STAN monitor requires that all fetuses monitored have had a normal heart rate pattern for at least 500 consecutive heartbeats, which equates to approximately 4 to 5 minutes of monitoring. This initial period of signal acquisition is used to generate the baseline parameters against which the algorithms will

compare all subsequent changes. This stipulation means that, despite its capability of discerning ST segment and T-wave changes, the STAN technology depends completely on the initial subjective decision by the clinical team as to whether a fetal heart rate tracing is reassuring or not. This requirement has probably created a significant source of confounding and bias since a number of different classification systems have been used to categorize the tracings in the published studies.⁴² In addition, different providers in different medical systems have differing opinions as to what is reassuring or normal. In fact, most of the published trials have used different classification systems and have not used a standardized objective measure to classify the fetal heart rate tracing. Most published randomized, controlled trials,^{24–28} used inconsistent interpretation of the 1987 FIGO-CTG categories⁴³ and were not constant in their definition of, or response to, ST-segment events. In addition, there was inconsistent use of fetal scalp sampling and intermittent monitoring of the fetal heart rate. These inconsistencies in study design, equipment, and clinical practice patterns between Europe and the U.S., underlined the need for the NICHD MFMU network trial. There was significant resistance by U.S. clinicians and medical systems to the introduction of a new technology that was not proven in an appropriately designed study using a U.S. population and practice patterns. In addition, the manufacturer had modified the color-coded fetal heart rate classification system from the four color classification used in Europe (Normal [green], Intermediary [yellow], Abnormal [orange], and Preterminal [red]), to a three color classification (green, yellow and red, based on the NICHD three category system)^{44,45} for their FDA application and for use in the U.S.⁴⁶ To add to the confusion, most recently (after completion of the U.S. trial), the latest (2015) FIGO Classification system⁴⁷ has introduced yet another classification that is different from that used in the STAN guidelines on which the prior European studies were based (Table 3).

The U.S. study by Belfort et al³¹ used the STAN guidelines that were approved by the FDA⁴⁶ and which were promoted by and used in the manufacturer's training program. It was felt that using the four color classification coding system promoted by the manufacturer in Europe would not have made any sense in the U.S. study given that the marketed product in the U.S. used a three color coding system. The subjective nature of the eligibility criteria for use of ST analysis technology was considered by many to be a major drawback, and the concern was raised that if the interpretation of STAN could be so heavily swayed by the fetal-heart interpretation, was the technology clinically reliable and efficacious in the U.S. setting? This skepticism was the basis for the U.S. trial which is discussed later in this article.

Clinical Use of the STAN System

During labor, fetuses essentially fall into three categories:^{43,44} those who are tolerating labor without any issue and the monitoring strip is strongly predictive of normal acid-base balance (NICHD category I); those who are clearly in trouble, with a strip predictive of abnormal fetal acid-base balance with a need for urgent intervention of some sort, if not urgent delivery (category III); and those who are neither a category I nor category III (category II). The category II monitoring strip is regarded as indeterminate, not predictive of an acidotic fetus but without adequate evidence to classify the strip as category I or category III. The recommendation for the management of a category II strip is to evaluate, continue

surveillance, and reevaluate, taking into account the entire clinical circumstances. Category II tracings are clearly the arena in which most of the more difficult decision making occurs. This was also the area where STAN has been reported to be potentially helpful in Europe, and where it required evaluation in the U.S. context. As mentioned above, the STAN system is based on a combination of FHR interpretation and adjunctive ST analysis. Because the entire premise of the system is based on an assurance that the fetus is not acidotic at the baseline, the ability to identify an abnormal FHR tracing is the rate limiting step of this technology. Competence in the assessment and management of FHR monitoring is an important component of any usage of the STAN system and this underscored one of the important reasons why STAN had to be studied in the U.S. context before any widespread use could be promulgated or endorsed. The system had never been tested using the NICHD FHR monitoring guidelines⁴⁴ (as opposed to International Federation of Gynecology and Obstetrics (FIGO) and local European guidelines) and had not been extensively studied using the FDA-approved STAN guidelines.⁴⁸

FDA approved classification system

In the U.S., STAN uses a FHR classification that separates the FHR tracing into 1 of 3 zones: green, yellow, or red. Much of the European work done on STAN has used a four color code classification that includes an intermediary [orange] zone between the yellow and red zones. Heart rate tracings classified as green zone do not require any intervention and may be watched expectantly regardless of any ST change. Tracings classified as red zone need expeditious delivery regardless of ST changes. Yellow-zone tracings in the U.S. (and in Europe yellow and orange zone tracings) have a more complex management schema that relies on the presence or absence of ST changes, and the degree of baseline T/QRS increase. The color zones in the FDA approved STAN system are roughly analogous to the NICHD categories but with two significant differences. (i) The green zone allows the presence of variable decelerations as long as they are less than 60 seconds in duration and less than 60 beats per minute in depth. NICHD category I does not allow any variable decelerations, and as such, the FDA approved STAN system may be slightly more lenient than the NICHD classification in terms of what is regarded as acceptable to continue to monitor without any concern for acidosis. (ii) In contrast, NICHD category II is less stringent than the FDA approved STAN system in what it regards as indeterminate in that NICHD category II tracings allow absent variability without recurrent decelerations, whereas the STAN system classifies such a strip as red zone needing expeditious delivery. The four tiers system used in Europe required a greater degree of abnormality in the intermediary [European yellow] and abnormal [European orange] zones than what was required in the FDA approved guidelines for intervention to be initiated. For example:

- i. In the FDA approved guidelines, a physician is required to evaluate the patient after there has been a yellow zone tracing for > 60 minutes versus in Europe no intervention is recommended for yellow or orange zone tracings regardless of duration.
- ii. In the FDA approved guidelines, when an episodic TQRS rise of > 0.05 generates an ST event in a yellow zone tracing intervention is required. In

Europe, intervention is required only when the episodic T/QRS rise is greater than 0.15 (3 times higher than in the USA) in the yellow zone, and 0.10 (2 × higher than in the USA) in the orange zone.

- iii. In the FDA approved guidelines, when a baseline TQRS rise of > 0.05 generates an ST event in a yellow zone tracing intervention is required. In Europe, intervention is required only when the baseline T/QRS rise is greater than 0.10 (2 times higher than in the USA) in the yellow zone.
- iv. In the FDA approved guidelines, when 2 biphasic log messages have generated an ST event in a yellow zone tracing intervention is required. In Europe, intervention is required only when there have been 3 biphasic log messages in the yellow zone.

These differences could result in later intervention in the European context and may explain some of the differences seen in the European studies and the U.S. study.

The promise of the STAN system lies in helping to manage those patients who are classified in the yellow zone (in the U.S.) or in the yellow and orange zones elsewhere - ST changes in this zone are believed to indicate progressive hypoxia, and the initiation of anaerobic metabolism, with its potential for metabolic acidosis.

As mentioned above, the recent changes to the FIGO classification system⁴⁷ have introduced yet another potential confounder to use of the STAN system. There are some important differences to the interpretation of what is accepted as normal between the two (FIGO and STAN) classifications. The main differences between the 2015 FIGO CTG classification and the STAN CTG classification are: (i) A baseline fetal heart rate of 150–160 bpm is classified as normal by FIGO but as intermediary (yellow zone) by STAN. (ii) Variability 5–25 bpm: accelerations are not needed for classifying a pattern as a normal CTG by FIGO, but are required by STAN, and (iii) Absent variability (silent pattern) and pre-terminal patterns are not classified by FIGO, but constitute a fourth CTG class (pre-terminal CTG) in the STAN CTG classification system.

Approved training, certification, and credentialing (as mandated by the FDA) are needed before use of the STAN system. In addition, as has been noted in most publications, there is a definite learning curve to the use of this system and indiscriminate use without adequate supervision is not advised. The system is FDA approved only for use in singleton pregnancies in which the fetus is more than 36–0/7 weeks, the membranes are ruptured, and the mother is in the first stage of labor without active or involuntary pushing. There should also be no contraindication to a fetal scalp electrode analysis or STAN.

The Randomized Trials and their Analysis

The Plymouth Trial²⁴ was the first large randomized controlled trial in which intervention rates and neonatal outcomes in fetuses monitored with cardiotocography (CTG) alone were compared with those in fetuses monitored using the combination of ST waveform analysis plus CTG. In this study of 2434 patients, there was a 43% reduction in operative interventions for fetal distress in the ST + CTG group. There was also a trend toward fewer

cases of cord artery metabolic acidosis and low Apgar score. The second, large-scale randomized controlled trial to compare outcomes following use of CTG alone and CTG + ST waveform analysis was the rather controversial Swedish Randomized Controlled trial.²⁵ The study included 4966 term fetuses in three large labor and delivery wards in Sweden. After exclusion of inadequate recordings and fetuses with malformations (n = 5574), the findings showed a 61% decrease in the number of fetuses born with umbilical cord arterial metabolic acidosis in the CTG + ST group. There was also a 28% decrease in operative interventions because of fetal distress in the STAN-monitored group. These findings were consistent with the results of the Plymouth trial.²⁴ It was concluded that intrapartum monitoring with CTG combined with automatic ST waveform analysis increased the ability to identify fetal hypoxia and to intervene more appropriately, resulting in an improved perinatal outcome. There were a number of problems with this trial and concerns were raised regarding patient selection, exclusion and statistical analysis.⁴⁹ Errors in the analysis were subsequently confirmed by a committee of the Swedish Research Council.⁵⁰ A revised modified intention-to-treat analysis confirmed that after correction for errors at data collection there was still a 52% reduction of metabolic acidosis, but the p-value of the difference was now 0.038.⁵¹

The Finnish randomized controlled study aimed to examine whether STAN could reduce the rate of neonatal acidemia and the rate of operative intervention during labor compared with fetal heart rate (FHR) monitoring alone.²⁶ A total of 1483 women in active labor with a singleton term fetus in cephalic presentation were randomly assigned to STAN + FHR monitoring or to FHR monitoring alone. Fetal scalp sampling was optional in both groups. The main outcome measures were neonatal acidemia (umbilical artery pH <7.10), neonatal metabolic acidosis (umbilical artery pH <7.05 and base excess <-12 mmol/L) and operative interventions: cesarean delivery rate, vacuum delivery rate, and fetal scalp sampling rate. There were no statistically significant differences between the STAN group and FHR monitoring group in the incidence of neonatal acidemia (5.8% vs 4.7%) or metabolic acidosis (1.7% vs 0.7%). The cesarean delivery rate (6.4% vs 4.7%) and the vacuum delivery rate (9.5% vs 10.7%) were also similar in the STAN and FHR monitoring groups. The incidence of fetal scalp sampling was lower (P<.001) in the STAN group (7.0%) than in the FHR monitoring group (15.6%). The investigators concluded that intrapartum fetal monitoring with STAN did not improve the neonatal outcome or decrease the cesarean rate. However, the need for fetal blood sampling (FBS) during labor was lower in the STAN group.

The French randomized controlled trial was conducted in two French maternity centers.²⁷ Its objective was to assess whether knowledge of ST-segment analysis was associated with a reduction in operative deliveries for non-reassuring fetal status (NRFS) or with a need for at least one scalp pH during labor. A total of 799 women in labor at 36 weeks or more, with a single fetus with cephalic presentation, and either abnormal cardiotocographic tracing or thick meconium-stained amniotic fluid were randomized to either monitoring with CTG alone (n = 400) or CTG + STAN (n = 399). Scalp pH sampling was an option in both groups. There was no difference between the groups in the primary outcome of operative delivery for NRFS. The proportion of patients who had at least one scalp pH measurement during labor was substantially lower in the CTG + STAN group as compared with CTG

alone: 27% compared with 62% (relative risk, 0.44; 95% confidence interval [CI], 0.36–0.52). There was no significant difference in a composite abnormal neonatal outcome between the groups (pH 7.05 or base deficit in extracellular fluid [BDecf] 12 mmol/L or 5-minute Apgar score < 7 or neonatal intensive care unit admission or convulsions or neonatal death).

The Dutch randomized controlled trial was a multicenter randomized pragmatic trial in three academic and six nonacademic teaching hospitals in The Netherlands.²⁸ The objective of the trial was to estimate the effectiveness of intrapartum fetal monitoring by CTG plus ST analysis using a strict protocol for performance of fetal blood sampling. Performance of fetal blood sampling was restricted to three situations: (1) start of STAN registration with an intermediary or abnormal CTG tracing; (2) abnormal CTG tracing for more than 60 minutes during the first stage without ST events; and (3) poor ECG signal quality in the presence of an intermediary or abnormal CTG tracing. The reason for this protocol was to control performance of scalp pH, a factor that may have affected the results of prior trials in which scalp pH use was not left up to the discretion of the provider. The primary outcome measure was the incidence of metabolic acidosis, defined as an umbilical cord artery blood pH below 7.05 and a base deficit calculated in the extracellular fluid compartment (BDecf) above 12 mmol/L according to the Siggaard-Andersen acid-base chart algorithm. Although base deficit in blood is the value most commonly reported by umbilical cord blood analyzers, and the one often used in clinical practice, the investigators in this trial thought that BDecf better reflects the true metabolic component of acidosis. There were 2832 and 2849 women randomly assigned to monitoring by CTG with ST analysis (index) or CTG only (control). The fetal blood sampling rate was 10.6% in the index compared with 20.4% in the control group (relative risk 0.52; 95% CI 0.46–0.59). The primary outcome occurred 0.7% in the index compared with 1.1% in the control group (relative risk 0.70; 95% CI 0.38–1.28; number needed to treat 252). Using metabolic acidosis calculated in blood, these rates were 1.6% and 2.6%, respectively (relative risk 0.63; 95% CI 0.42–0.94; number needed to treat 100). The number of operative deliveries, low Apgar scores, neonatal admissions, and newborns with hypoxic-ischemic encephalopathy was comparable in both groups. Although this study did not show a difference in the primary outcome, it did show a lower rate of metabolic acidosis defined using base deficit in blood, the most common method used in the United States. And this benefit was achieved despite performing 48% less fetal blood sampling when ST segment analysis was used. It should also be noted that the actual rates of metabolic acidosis using base deficit in extracellular fluid (1.1% in the control and 0.7% in the index group) were much lower than the 3.5% assumed in the sample size calculations.

All of the trials described above begged the question as to how would ST segment analysis perform when fetal blood sampling was not performed. It became clear to clinicians in the U.S. that whereas data seemed encouraging in the European system, the utility of this new method of fetal monitoring had yet to be confirmed in a U.S. population. There were two main reasons why another randomized controlled trial (RCT) was needed before widespread general use of STAN could be endorsed in the United States. First, STAN had never been evaluated with an RCT under the specific conditions of obstetric health care delivery in this country. STAN is not simply a stand-alone technology but rather a system of management. How the U.S. obstetric community would adapt to this technology and its associated

protocol was not something that could be assumed to mirror the experience in Europe. The U.S. model for obstetric care delivery is different from that in European countries and the direct extrapolation of data from these countries to a U.S. population is not appropriate. Second, although metabolic acidosis and some aspects of neonatal outcome had been studied, no RCT had specifically addressed a composite neonatal outcome as the primary outcome of interest. Although a decrease in operative delivery is a very important outcome in obstetric practice, the reduction in adverse neonatal outcomes is the ultimate goal for a monitoring system designed to reduce fetal and neonatal acidosis.

This then was the rationale for the U.S. RCT that was published in 2015.³¹ The study, which consisted of a pilot phase and a randomized trial, was conducted at 16 university-based clinical centers – each comprised of one to five delivery hospitals (26 total) – of the NICHD MFMU Network. The level of attention paid to ensuring that participating providers were adequately trained and experienced prior to enrolling patients in the RCT was unprecedented. Participating care providers and research personnel were trained and certified in the correct use of the fetal ECG ST analysis system to a level exceeding FDA requirements in a program in which the manufacturer participated and had oversight to ensure that correct training and usage of their equipment was in place. Each hospital participated in a pilot phase consisting of enrollment and management of at least 50 patients monitored with fetal ECG ST analysis, with central review of labor management decisions and re-training as needed, before being approved to start the trial. Once approved to begin the RCT, each site enrolled women with a singleton gestation attempting vaginal delivery at more than 36 weeks' gestation, and who had cervical dilation between 2 to 7 cm (inclusive). Patients were randomly assigned to “open” or “masked” fetal ST analysis monitoring. The masked system functioned as a normal fetal heart rate monitor. The open system displayed additional information for use with uncertain (yellow zone) fetal heart rate patterns. The primary outcome was a composite of fetal death, neonatal death, Apgar score ≤ 3 at 5 minutes, neonatal seizure, cord artery pH ≤ 7.05 with base deficit ≥ 12 mmol/L, intubation for ventilation at delivery, or neonatal encephalopathy.

There were 11,108 patients randomized (5,532 open; 5,576 masked). The primary outcome occurred in 52 patients (0.9%) in the open arm and 40 (0.7%) in the masked arm (relative risk 1.31; 95% CI 0.87 to 1.98; $p=0.20$). Among the individual components of the primary outcome, only the frequency of 5-minute Apgar score ≤ 3 differed significantly between the open and masked arms (0.31% versus 0.11%; respectively $P=0.02$). There were no significant differences between groups in cesarean delivery (16.9% versus 16.2%; $P=0.31$) or any operative delivery (22.8% versus 22.0%; $P=0.31$). Adverse events were rare, but similar in the two groups.

The U.S. trial finding of no improvement in neonatal outcomes or reduction in cesarean delivery rates is consistent with the results of an individual patient data meta-analysis of ST analysis trials.⁵² However, that meta-analysis showed a reduction in the frequency of fetal blood sampling (not routine in the U.S.) and operative vaginal delivery in women who had electronic fetal monitoring and adjunctive ST analysis, compared with those who had electronic fetal monitoring alone.⁴¹

Concerns about the failure of STAN to predict cases of intrapartum metabolic acidosis have been raised over the years.⁵³ Three cases of intrapartum metabolic acidosis were documented by the Dutch group⁵⁴ and a paper from St. George's Hospital in London reported on 14 cases of neonatal encephalopathy in the first 1052 patients who were monitored using STAN and stated that in only 7 of these 14 had there been a significant ST event.⁵⁵ It was subsequently established that STAN monitoring was being started when the fetal heart rate tracing was already abnormal, a finding that reinforces the importance of only starting STAN monitoring *after* a reassuring pattern has been established (and maintained until the baseline ST analysis is completed). In the same study STAN also failed to identify seven (30%) of 23 cases of metabolic acidosis. Another paper published in 2008 reported on approximately 5000 patients monitored using STAN, and showed that in only two of three cases (66%) with severe, and in 20/48 (42%) cases with moderate metabolic acidemia, were there ST events coinciding with CTG abnormalities.⁵⁶

The Cochrane Review

Seven randomized trials were included in the most recent Cochrane review (September 2015) on the topic.⁵⁷ The objective of this review was to compare the effects of analysis of fetal electrocardiogram (ECG) waveforms during labor with alternative methods of fetal monitoring. Seven trials (27,403 women) were included: six trials of ST waveform analysis (26,446 women) and one trial of PR interval analysis (957 women).⁵⁸ The reviewers felt that the trials were generally at low risk of bias for most domains and the quality of evidence for ST waveform analysis trials was graded moderate to high. The major finding was that in comparison with continuous electronic fetal heart rate monitoring alone, the use of adjunctive ST waveform analysis made no obvious difference to primary outcomes: births by cesarean (risk ratio (RR) 1.02, 95% confidence interval (CI) 0.96 to 1.08; six trials, 26,446 women; high quality evidence); the number of babies with severe metabolic acidosis at birth (cord arterial pH less than 7.05 and base deficit greater than 12 mmol/L) (average RR 0.72, 95% CI 0.43 to 1.20; six trials, 25,682 babies; moderate quality evidence); or babies with neonatal encephalopathy (RR 0.61, 95% CI 0.30 to 1.22; six trials, 26,410 babies; high quality evidence). There were, however, on average fewer fetal scalp samples taken during labor (average RR 0.61, 95% CI 0.41 to 0.91; four trials, 9671 babies; high quality evidence) although the findings were heterogeneous and there were no data from the largest trial (from the U.S.). There were marginally fewer operative vaginal births (RR 0.92, 95% CI 0.86 to 0.99; six trials, 26,446 women); but no obvious difference in the number of babies with low Apgar scores at five minutes or babies requiring neonatal intubation, or babies requiring admission to the special care unit (RR 0.96, 95% CI 0.89 to 1.04, six trials, 26,410 babies; high quality evidence). The authors concluded that the modest benefits of fewer fetal scalp sampling procedures during labor (in settings in which this procedure is performed) and fewer instrumental vaginal births have to be considered against the disadvantages of needing to use an internal scalp electrode, after membrane rupture, for ECG waveform recordings. They found little strong evidence that ST waveform analysis had an effect on the primary outcome measures in their systematic review.

Conclusion

As the rate of cesarean delivery remains high in the U.S., there is need for better methods to evaluate fetal well-being intrapartum. The MFMU Network has evaluated two of the most promising methods, without finding a benefit. Although the results were negative, the studies actually prevented adoption of these methods without appropriate evidence, as happened with the original introduction of fetal heart rate monitoring.

Without the unique setup of the MFMU Network, these studies would likely never have been performed in the U.S. Given that both of these technologies were already approved for use in the U.S., without the MFMU Network trials, the outcome would have likely gone the way of the fetal heart rate monitoring technology. Our experience in these two situations also highlights the importance of performing the trials that are adequately powered to address the appropriate outcomes and that are generalizable to the U.S. population and management. These two case studies clearly illustrate that hurried adoption of new technologies based on surrogate outcomes or on studies in clinical settings other than the U.S. may not be a good approach to health care. They also highlight the importance of having infrastructures such as the MFMU Network which allow such large, complex and rigorous trials to be conducted, and its impact on maternal and child health to be rigorously measured.

Acknowledgments

The project described was supported by grants from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) [HD21410, HD27860, HD27869, HD27915, HD27917, HD34116, HD34136, HD34208, HD53097, HD40545, HD40560, HD27869, HD40485, HD40500, HD40512, HD40544, M01 RR00080 (NCRR); HD68282, HD68268, HD27917, HD36801]. Comments and views of the authors do not necessarily represent views of the NICHD.

References

1. Banta HD, Thacker SB. Assessing the costs and benefits of electronic fetal monitoring. *Obstet Gynecol Surv.* 1979; 34:627–42. [PubMed: 113713]
2. Ventura, SJ.; Martin, JA.; Curtin, SC.; Mathews, TJ.; Park, MM. National vital statistics report. Vol. 48. Hyattsville, Maryland: National Center for Health Statistics; 2000. Births: Final data for 1998.
3. National Institute of Child Health and Human Development. Antenatal Diagnosis. Report of a Consensus Development Conference, Assisted by the Office for Medical Applications of Research and the Fogarty international Center; March 5–7, 1979; National Institutes of Health, Bethesda, Maryland. U.S. Department of Health, Education and Welfare. Public Health Service, National Institutes of Health; Apr. 1979 NIH Publication No. 79-1973
4. National Institute of Child Health and Human Development. Cesarean Childbirth. Report of a Consensus Development Conference in Conjunction with the National Center for Health Care Technology and Assisted by the Office for Medical Applications of Research; September 22–24, 1980; National Institutes of Health, Bethesda, Maryland. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health; Oct. 1981 NIH Publication No. 82-2067
5. National Institute of Child Health and Human Development Research Planning Workshop. Electronic fetal heart rate monitoring: research guidelines for interpretation. *Am J Obstet Gynecol.* 1997; 177:1385. [PubMed: 9423739]
6. [Accessed January 5, 2016] Maternal Fetal Medicine Network Home Page. <https://mfmu.bsc.gwu.edu>

7. [Accessed October 30, 2006] Summary minutes: meeting of the Obstetrics and Gynecology Devices Advisory Panel, open session. Jun 9. 2003 at http://www.fda.gov/ohrms/dockets/ac/03/minutes/3963m1_summary%20minutes.pdf
8. Yam J, Chua S, Arulkumaran S. Intrapartum fetal pulse oximetry. Part I: principles and technical issues. *Obstet Gynecol Surv.* 2000; 55:163–72. [PubMed: 10713982]
9. Bloom SL, Swindle RG, McIntire DD, Leveno KJ. Fetal pulse oximetry: duration of desaturation and intrapartum outcome. *Obstet Gynecol.* 1999; 93:1036–40. [PubMed: 10362177]
10. Nijland R, Jongsma HW, Nijhuis JG, van den Berg PP, Oeseburg B. Arterial oxygen saturation in relation to metabolic acidosis in fetal lambs. *Am J Obstet Gynecol.* 1995; 172:810–9. [PubMed: 7892869]
11. Dildy GA, Thorp JA, Yeast JD, Clark SL. The relationship between oxygen saturation and pH in umbilical blood: Implications for intrapartum fetal oxygen saturation monitoring. *Am J Obstet Gynecol.* 1996; 175:682–7. [PubMed: 8828434]
12. Goffinet F, Langer B, Carbonne B, Berkane N, Tardif D, Le Goueff F, et al. Multicenter study on the clinical value of fetal pulse oximetry. The French Study Group on Fetal Pulse Oximetry. *Am J Obstet Gynecol.* 1997; 177:1238–46. [PubMed: 9396924]
13. Garite TJ, Dildy GA, McNamara H, et al. A multicenter controlled trial of fetal pulse oximetry in the management of non-reassuring fetal heart rate patterns. *Am J Obstet Gynecol.* 2000; 183:1049–58. [PubMed: 11084540]
14. American College of Obstetricians and Gynecologists. Fetal pulse oximetry: ACOG committee opinion no. 258, September 2001. *Obstet Gynecol.* 2001; 98:523–4. [PubMed: 11547792]
15. Bloom SL, Spong CY, Thom E, Varner MW, Rouse DJ, Weinger S, Ramin SM, et al. Fetal pulse oximetry and cesarean delivery. *N Engl J Med.* 2006; 355:2195–202. [PubMed: 17124017]
16. Dildy GA. Fetal pulse oximetry. *Clin Obstet Gynecol.* 2011; 54:66–73. [PubMed: 21278503]
17. Klausner CK, Christensen EE, Chauhan SP, Bufkin L, Magann EF, Bofil JA, Morrison JC. Use of fetal pulse oximetry among high-risk women in labor: A randomized clinical trial. *Am J Obstet Gynecol.* 2005; 192:1810–9. [PubMed: 15970816]
18. East CE, Brennecke SP, King JF, Chan FY, Colditz PB. on behalf of The FOREMOST Study Group. The effect of intrapartum fetal pulse oximetry, in the presence of a non-reassuring fetal heart rate pattern, on operative delivery rates: A multicenter, randomized, controlled trial (the FOREMOST trail). *Am J Obstet Gynecol.* 2006; 194:606.e1–e16. [PubMed: 16522387]
19. Caliskan E, Cakiroglu Y, Corakci A, Ozeren S. Reduction in cesarean delivery with fetal heart rate monitoring and intermittent pulse oximetry after induction of labour with misoprostol. *J Matern Fetal Neonatal Med.* 2009; 22:445–51. [PubMed: 19530004]
20. Kuhnert M, Schmidt S. Intrapartum management of nonreassuring fetal heart rate patterns: A randomized controlled trial of fetal pulse oximetry. *Am J Obstet Gynecol.* 2004; 191:1989–95. [PubMed: 15592281]
21. East CE, Begg L, Colditz PB, Lau R. Fetal pulse oximetry for fetal assessment in labour. *Cochrane Database Syst Rev.* 2014; (10) Art No.: CD004075. doi: 10.1002/14651858.CD004075.pub4
22. [Accessed 7 February, 2011] FDA document. Available at: http://www.accessdata.fda.gov/cdrh_docs/pdf2/p020001a.pdf
23. Rosen KG, Lindcrantz K. STAN—the Gothenburg model for fetal surveillance during labour by ST analysis of the fetal electrocardiogram. *Clin Phys Physiol Meas.* 1989; 10(Suppl B):51–6. [PubMed: 2698308]
24. Westgate J, Harris M, Curnow JS, et al. Plymouth randomized trial of cardiotocogram only versus ST waveform plus cardiotocogram for intrapartum monitoring in 2400 cases. *Am J Obstet Gynecol.* 1993; 169:1151–60. [PubMed: 8238177]
25. Amer-Wahlin I, Hellsten C, Norén H, et al. Cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram for intrapartum fetal monitoring: a Swedish randomised controlled trial. *Lancet.* 2001; 358:534–8. [PubMed: 11520523]
26. Ojala K, Vaarasmaki M, Makikallio K, et al. A comparison of intrapartum automated fetal electrocardiography and conventional cardiotocography—a randomized controlled study. *BJOG.* 2006; 113:419–23. [PubMed: 16553653]

27. Vayssiere C, David E, Meyer N, et al. A French randomized controlled trial of STsegment analysis in a population with abnormal cardiocograms during labor. *Am J Obstet Gynecol.* 2007; 197:299, e1–6. [PubMed: 17826428]
28. Westerhuis ME, Visser GH, Moons KG, et al. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. *Obstet Gynecol.* 2010; 115:1173–80. [PubMed: 20502287]
29. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. *Cochrane Database Syst Rev.* 2006; 3:CD000116. [PubMed: 16855950]
30. Belfort MA, Saade GR. ST segment analysis (STAN) as an adjunct to electronic fetal monitoring, Part II: clinical studies and future directions. *Clin Perinatol.* 2011; 38:159–67. [PubMed: 21353096]
31. Belfort MA, Saade GR, Thom E, et al. A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. *N Engl J Med.* 2015 Aug 13; 373(7):632–41. [PubMed: 26267623]
32. Rosen KG, Kjellmer I. Changes in the fetal heart rate and ECG during hypoxia. *Acta Physiol Scand.* 1975; 93:59–66. [PubMed: 1155132]
33. Rosen KG, Isaksson O. Alterations in the fetal heart rate and ECG correlated to glycogen, creatine phosphate and ATP levels during graded hypoxia. *Biol Neonate.* 1976; 30:17–24.
34. Hokegard KH, Eriksson BO, Kjellmer I, et al. Myocardial metabolism in relation to electrocardiographic changes and cardiac function during graded hypoxia in the fetal lamb. *Acta Physiol Scand.* 1981; 113:1–7. [PubMed: 7315431]
35. Greene KR, Dawes GS, Lilja H, et al. Changes in the ST waveform of the fetal lamb electrocardiogram with hypoxemia. *Am J Obstet Gynecol.* 1982; 144:950–7. [PubMed: 7148927]
36. Dawes GS, Mott JC, Shelley HJ. The importance of cardiac glycogen for the maintenance of life in foetal lambs and newborn animals during anoxia. *J Physiol.* 1959; 146:516–38. [PubMed: 13665675]
37. Towell ME. Catecholamine depletion and the response to asphyxia in the fetal guinea pig. *Biol Neonate.* 1971; 18:212–24. [PubMed: 5155591]
38. Fenn WO, Cobb DM. Electrolyte changes in muscle during activity. *Am J Physiol.* 1936; 115:345–56.
39. Wohlfart B. A simple model for demonstration of STT-changes in ECG. *Eur Heart J.* 1987; 8:409–16. [PubMed: 3609036]
40. Rosen KG. Fetal electrocardiogram waveform analysis in labour. *Curr Opin Obstet Gynecol.* 2001; 13:137–40. [PubMed: 11315867]
41. Yli BM, Kallen K, Stray-Pedersen B, et al. Intrapartum fetal ECG and diabetes. *J Matern Fetal Neonatal Med.* 2008; 21:231–8. [PubMed: 18330818]
42. Belfort MA, Saade GR, Thom EA. Intrapartum Fetal ECG ST-Segment Analysis. *N Engl J Med.* 2015 Dec 17; 373(25):2480–1. [PubMed: 26672861]
43. FIGO Subcommittee on Standards in Perinatal Medicine. Guidelines for the use of fetal monitoring. *Int J Gynaecol Obstet.* 1987; 25:159–67.
44. American College of Obstetricians and Gynecologists. Intrapartum Fetal Heart Rate Monitoring: Nomenclature, Interpretation, and General Management Principles. *ACOG Practice Bulletin No. 106.* *Obstet Gynecol.* 2009; 114:192–202. [PubMed: 19546798]
45. Macones GA, Hankins GD, Spong CY, et al. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretations, and research guidelines. *Obstet Gynecol.* 2008; 112(3):661–6. [PubMed: 18757666]
46. [Accessed 7 February, 2011] FDA-approved STAN guidelines. Available at: <http://www.fda.gov/cdrh/pdf2/p020001a.pdf>
47. Ayres-de-Campos D, Spong CY, Chandrharan E. for the FIGO Intrapartum Fetal Monitoring Expert Consensus Panel. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. *Int J Gynecol Obstet.* 2015; 131:13–24.
48. Devoe LD, Ross M, Wilde C, et al. United States multicenter clinical U.S. ge study of the STAN 21 electronic fetal monitoring system. *Am J Obstet Gynecol.* 2006; 195:729–34. [PubMed: 16949404]

49. Lund University. [accessed April 4, 2014] Press release regarding alleged research fraud 22nd October 2008. Available online at: http://www.med.lu.se/english/research/press_release_22_october_2008
50. Lund University. [accessed April 4, 2014] Press release: a decision relating to suspected misconduct in research has been made. Available online at: http://www.med.lu.se/english/research/press_release_16_september_2010
51. Amer-Wahlin I, Kjellmer I, Marsal K, Olofsson P, Rosen KG. Swedish randomized controlled trial of cardiotocography only versus cardiotocography plus ST analysis of fetal electrocardiogram revisited: analysis of data according to standard versus modified intention-to-treat principle. *Acta Obstet Gynecol Scand*. 2011; 90:990–6. [PubMed: 21623743]
52. Schuit E, Amer-Wahlin I, Ojala K, et al. Effectiveness of electronic fetal monitoring with additional ST analysis in vertex singleton pregnancies at >36 weeks of gestation: an individual participant data metaanalysis. *Am J Obstet Gynecol*. 2013; 208:187, e1–187. [PubMed: 23333546]
53. Steer PJ, Hvidman LE. Scientific and clinical evidence for the use of fetal ECG ST segment analysis (STAN). *Acta Obstet Gynecol Scand*. 2014; 93:533–8. [PubMed: 24597897]
54. Westerhuis M, Kwee A, van Ginkel A, Drogtróp A, Gyselaers W, Visser G. Limitations of ST analysis in clinical practice: three cases of intrapartum metabolic acidosis. *BJOG*. 2007; 114:1194–201. [PubMed: 17501963]
55. Doria V, Papageorgiou AT, Gustafsson A, Ugwumadu A, Farrer K, Arulkumaran S. Review of the first 1502 cases of ECG-ST waveform analysis during labour in a teaching hospital. *BJOG*. 2007; 114:1202–7. [PubMed: 17877672]
56. Melin M, Bonnevier A, Cardell M, Hogan L, Herbst A. Changes in the ST-interval segment of the fetal electrocardiogram in relation to acid–base status at birth. *BJOG*. 2008; 115:1669–75. [PubMed: 19035941]
57. Neilson JP. Fetal electrocardiogram (ECG) for fetal monitoring during labour. *Cochrane Database Syst Rev*. 2015 Dec 21.12:CD000116. [PubMed: 26690497]
58. Strachan BK, van Wijngaarden WJ, Sahota D, Chang A, James DK. for the FECG Study Group. Cardiotocography only versus cardiotocography plus PR-interval analysis in intrapartum surveillance: a randomised, multicentre trial. *Lancet*. 2000; 355:456–9. [PubMed: 10841126]

Table 1

Routes of and indications for delivery in the randomized trial of fetal pulse oximetry published by Garite and colleagues.¹³

Route of Delivery and Indication	EFM Alone n = 502	EFM + Fetal Pulse Oximetry n = 508	P value
Spontaneous vaginal delivery	255 (51)	241 (47)	NS
Operative vaginal delivery:			
Overall	117 (23)	120 (24)	NS
For non-reassuring fetal status	57 (11)	55 (11)	NS
For all other indications	60 (12)	65 (13)	NS
Cesarean delivery:			
Overall	130 (26)	147 (29)	NS
For non-reassuring fetal status	51 (10)	23 (5)	< 0.001
For dystocia	43 (9)	94 (19)	< 0.001
For non-reassuring fetal status + dystocia	35 (7)	27 (5)	NS
For other indications	1 (0)	3 (1)	NS

Data are shown as n (%). EFM = electronic fetal monitoring. NS = non-significant

Table 2

Routes of and indications for delivery in the randomized trial of fetal pulse oximetry published by the MFMU Network.¹⁵

Route of Delivery and Indication for the Entire Cohort	Open Group N = 2629	Masked Group N = 2712	Relative Risk (95% CI)
Spontaneous vaginal delivery	1557 (59.2)	1565 (57.7)	1.03 (0.98–1.07)
Operative vaginal delivery	380 (14.5)	400 (14.7)	0.98 (0.86–1.12)
Cesarean delivery:			
Overall	692 (26.3)	747 (27.5)	0.96 (0.87–1.04)
For non-reassuring FHR	187 (7.1)	213 (7.9)	0.91 (0.75–1.09)
For dystocia	490 (18.6)	521 (19.2)	0.97 (0.87–1.08)
Cesarean Delivery for Women with non- reassuring FHR before randomization	Open Group N = 1055	Masked Group N = 1113	
Cesarean delivery:			
Overall	327 (31.0)	339 (30.5)	1.02 (0.90–1.15)
For non-reassuring FHR	104 (9.9)	123 (11.1)	0.89 (0.70–1.14)
For dystocia	216 (20.5)	210 (18.9)	1.09 (0.92–1.29)

Data are shown as n (%). FHR = fetal heart rate

Table 3

Summary of the 2015 FIGO intrapartum cardiotocography (CTG) classification system.⁴⁷

	Normal CTG ^a	Suspicious CTG	Pathological CTG
Baseline ^b	110–160 bpm	Lacking at least one of normal characteristics, but with no pathological features	<100bpm
Variability ^{c,d,j}	5–25 bpm		Reduced/increased variability ^{c,d} ; sinusoidal pattern ^j
Decelerations ^{e,f,g,h,i}	No repetitive * decelerations		Repetitive * late or prolonged decelerations for >30 min (or >20 min if reduced variability); one deceleration >5 min
Interpretation	No hypoxia/acidosis	Low probability of hypoxia/acidosis	High probability of hypoxia/acidosis

* Decelerations are repetitive if they occur at >50% of uterine contractions.

^aAccelerations (amplitude >15 bpm, lasting >15 s but <10 min) are not included in the intrapartum CTG classification system because their absence have uncertain significance, but presence of accelerations indicates a neurologically responsive fetus without hypoxia/acidosis.

^bTachycardia defined as >160 bpm for >10 min, bradycardia defined as <110 bpm for >10 min. A baseline 100–110 bpm may occur in normal fetuses, especially if postdate.

^cReduced variability is bandwidth <5 bpm for >50 min in baseline segments, or for >3 min during decelerations.

^dIncreased variability (saltatory pattern) is bandwidth >25 bpm for >30 min.

^eA deceleration is a loss of heart rate of >15 bpm for >15 s.

^fEarly decelerations are shallow, short-lasting, have normal variability, and are coincident with uterine contractions.

^gVariable decelerations have a rapid drop to nadir within 30 s after onset, good variability, rapid return to baseline, are V-shaped but vary in shape, size and relation to uterine contractions.

^hLate decelerations are U-shaped with reduced variability, start more than 20 s after the onset of uterine contraction with a gradual onset with >30 s to nadir, have a nadir after the acme of contraction, and return to baseline gradually with >30 s from nadir to baseline.

ⁱProlonged decelerations are lasting >3 min. Decelerations remaining <80 bpm for >5 min and with reduced variability are serious.

^jSinusoidal pattern is a regular, smooth, undulating pattern resembling a sinus wave, with amplitude 5–15 bpm at frequency 3–5 cycles/min, lasting for >30 min and coinciding with absent accelerations.