New perspectives in electronic fetal surveillance*

Emily F. Hamilton1,2,** and Philip A. Warrick1
1PeriGen, Montreal, QC, Canada
2McGill University, Montreal, QC, Canada

Abstract

Despite its recognized limitations, fetal heart rate monitoring is a mainstay of intrapartum care. Although the basic technology in standard electronic fetal monitors has changed little in recent decades, clinical behavior in response to heart rate monitoring has changed considerably. In addition to clearly defined nomenclature and clinical guidelines, there is an increased awareness that environmental and human factors can impair clinical judgment, resulting in delayed intervention and, consequently, birth-related injury. This review examines three essential steps that affect clinical outcome: (1) signal acquisition, (2) associations with physiological outcome, and (3) clinical intervention. Only the third step is directly responsible for changing clinical outcome. However, timely initiation of interventions is dependent upon the second step, which is dependent upon the first step. Thus, deficiencies at each step tend to accumulate and contribute to the worsening of overall clinical outcome. This review article summarizes advances occurring at each step. The synergy and convergence of innovations in engineering, mathematics, and behavioral science shows considerable promise in intrapartum fetal surveillance.

Keywords: Electronic fetal monitoring; fetal heart rate; metabolic acidemia.

Introduction

Controversy is no stranger to electronic fetal monitoring (EFM). Few technologies in medicine can claim association with such an extensive and debated literature. Despite disagreements, it remains a mainstay of intrapartum care, suggesting that clinicians find its benefits outweigh its dis-advantages. The merits of this clinical behavior are supported by encouraging reports of falling rates of hypoxic-ischemic encephalopathy (HIE). Several large jurisdictions with population-wide statistics have shown steady declines, often reaching reductions in the range of 50% per decade. Preventable intrapartum stillbirths are almost eliminated [4, 7, 17, 32, 54, 67].

A myriad of factors influence outcomes, such as HIE, and it would be incorrect to attribute the improvement to any single one. The basic sensor technology of EFM has changed little over the past few decades. Thus, improvement in outcome is not likely to be related to simply better recording of the fetal heart rate (FHR). In contrast, clinical practices today are very different from those in the mid-1980s, when the largest randomized clinical trial (RCT) on EFM was reported. During this randomized study of 12,964 mothers, perinatal hypoxic death or newborn seizures occurred at a rate of approximately 1 in every 285 births, and primary cesarean rates were around 2.4% [43]. US cesarean rates now exceed 30% [47]. The rise is attributed to a number of factors in addition to an EFM effect. A recent review of over 1.7 million US births in 2004 found no difference in the primary cesarean rate among women with EFM (17.7%) compared with those without EFM (17.8%) [10].

Another change has been a growing recognition that human factors and system failures play a substantial role in adverse outcomes across all of branches of medicine [59]. Human actions, such as delayed recognition of tracing abnormality and delayed intervention are reported to have occurred in approximately half of asphyxial injuries or death [11, 23, 35, 57, 63]. The patient safety movement has ascertained many causes of medical error and, drawing from aviation and military experience, provides guidance on how to build less error-prone health-care systems. Consequently, policies and procedures to redress the reasons underlying human error are now widespread. Some actions are directed at system vulnerabilities, such as legislation to limit working hours, recommendations on staff-to-patient ratios, setting standards for the availability of obstetricians and operating room facilities, simulation training for emergency procedures, and formal feedback on performance to clinicians [9, 12, 37, 56]. Other actions are specific to EFM, such as standardizing nomenclature, defining graded classifications of abnormality, and establishing formal guidelines for clinical management [1, 24, 41, 44, 51, 62]. Finally, improved understanding about the clinical significance between some FHR patterns and outcome helps clinicians respond better. For example, defining the correlation between rates of fetal death, or HIE, and the interval between persistent bradycardia and delivery is crucial to establishing desirable response times [39].
Three basic steps in clinical monitoring

This brief review of the evolution of fetal monitoring underscores three essential steps that affect clinical outcome: (1) signal acquisition, (2) associations with physiological outcome, and (3) effective clinical intervention. This article will demonstrate how they interact, discuss evaluation strategies, and present advances that may improve the overall process.

Understanding the hierarchy and dependencies of these steps is essential. Better signal acquisition increases the potential to recognize when intervention is needed. For example, consistent and reliable detection of the FHR improves its clinical utility. However, failure to intervene when a reliable test is abnormal would negate any benefit gained through testing. Thus, each step is dependent upon the preceding step and the ultimate clinical outcome is dependent upon all three. Moreover, each step should be measured separately to identify and rectify weak links in the process.

The most basic step in fetal surveillance is signal acquisition from the monitoring device and subsequent display. Evaluation of this step should include performance measures, such as the frequency of missing signals, accuracy of the readings, and clarity of displays. Practical questions, such as “how easy is it to use the device?” and “are there complications associated with its application?” are also relevant to clinical practice. All fetal surveillance devices are subject to health-care regulatory agencies that scrutinize aspects related to safety and effectiveness carefully in their review process. It is important to note that effectiveness in this context means, “does the device do what its manufacturer states it does?” For example, if the manufacturer claims that the device can measure the FHR, “effectiveness” relates to how well it does exactly that. It does not refer to effectiveness in reducing a particular birth complication. It is up to clinical professional to decide if the device has clinical value. Nevertheless, how often the device actually measures what it is intended to measure is a fundamental determinant of clinical utility. For instance, an imaging device that produces satisfactory images only 50% of the time is unlikely to ever achieve high diagnostic sensitivity.

The second step in the sequence is the fundamental physiological association between the measured parameter measured and the pathology of interest. Exactly what patterns in the FHR recordings are associated with metabolic acidemia? How often are these patterns found in uncomplicated births and how discriminating are they for adverse outcome? The task of correctly inferring the state of the baby and its internal environment from its FHR patterns is daunting because the relationships between the FHR and metabolic acidemia are very indirect and non-specific. For example, low heart rate variability may indicate a depressed fetal central nervous system in the presence of metabolic acidemia, but it can also be seen in a range of normal conditions from, for example, fetal sleep or with certain medications. Decelerations provide some information about cardiac output, which is related to the development of metabolic acidemia, but decelerations with something relatively innocuous, such as intermittent cord compression may look identical to those with an impending catastrophic event like uterine rupture [29].

The usual measures of the relationship between a diagnostic test and its target disease are sensitivity (rate that the test is positive among diseased individuals) and false-positive rates (rate that the test is positive among healthy individuals). These measures are not entirely appropriate for fetal surveillance where the goal is prevention not detection. For example, when a baby is born after an intervention for an abnormal tracing and has a normal outcome, it is not possible to be certain if this was a “successful prevention” or an “unnecessary intervention” unless there is a good marker of impending peril. In animal studies, one can create hypoxemia and measure sensitivity and false-positive rates associated with the heart rate patterns without the interference of therapeutic intervention. This is not possible with EFM in humans.

One approach to estimating sensitivity and false-positive rates in humans uses historical cases and measures the incidence of certain patterns in the groups with or without metabolic acidosis at birth. Although this underestimates the properties of the device, because monitoring can instigate intervention that prevents adverse outcome, it does provide some insight on performance.

The third step is clinical intervention: What should be done? How urgently should it be accomplished? How effective is the clinical intervention? Only the third step is directly responsible for changing clinical outcome. However, timely initiation of interventions is dependent upon the second step, which is dependent upon the first step. Thus, deficiencies at each level tend to accumulate and contribute to worsening the overall clinical outcome. Randomized clinical trials where a clinical outcome is the endpoint measure the combined effect of all three steps.

If we could optimize each of the steps, the task for clinicians would be considerably easier than it is today. The next section will focus mostly on exciting developments in engineering and mathematics pertaining to steps 1 and 2. We turn to advances in the social sciences of human behavior and health-care management for the third step.

Step 1: signal acquisition

Standard fetal monitoring devices have changed little in recent decades. FHRs are still calculated based on the identification of cardiac pulsation by either an external Doppler ultrasound sensor or a fetal ECG sensor on the fetal scalp. Uterine pressure is measured directly with intrauterine pressure sensors or indirectly using external sensors that measure abdominal wall tension.

There are two broad groups of technologies with respect to innovation in monitoring devices. One group includes external sensors to measure the FHR or uterine contractions non-invasively. With concerns about infection, reliable non-invasive monitoring is highly desirable. Wireless transmission allows the patient to be mobile, provided the sensors can still perform well with ambulation.

External sensors Electrodes placed on the surface of the maternal abdomen can detect changes in the electromagnetic field that arise from uterine contractions and from the fetal...
heart. The changes induced by the beating fetal heart are relatively weak, given its size and are mixed with changes from maternal heart beats and other contracting muscles. It is therefore a challenging engineering problem to isolate the fetal cardiac signal and calculate heart rate and even more challenging to construct the actual ECG complex with sufficient fidelity to measure fine details, such as the R-R interval, ST segment, or other aspects related to altered myocardial function [13, 30].

R-R interval measurements or beat-to-beat variability calculations are not available with conventional fetal monitors with their standard data exporting methods. Heart rates from the Doppler sensors are averaged over several seconds (i.e., several heart beats) to improve signal-to-noise ratios, thus eliminating the potential to estimate individual R-R intervals. Although scalp electrodes acquire cleaner signals and permit heart rate estimation with each heart beat, the standard protocol for exporting the heart rate does not preserve this level of detail, which is important for characterization of the short-term FHR variability. Thus, both scalp ECG and surface ECG methods could give beat to beat intervals. However, this is not readily available with most current commercial applications [15, 20].

Surface detection of the fetal ECG complex with sufficient precision to analyze the P-QRS-T waves would open another dimension of possibilities. Based on the premise that myocardial changes induced by hypoxemia are reflected in changes in the ECG waveform, ECG analysis could provide additional information that helps to determine which fetus is experiencing clinically significant hypoxemia [5, 6, 40, 61, 70].

Uterine contraction detection is also possible with these sensors. Using surface electrodes, Euliano et al. [26] have measured the characteristics of the electrohistogram (EHG) of uterine contractions. In this small study, uterine contractions propagated with less fundal dominance in mothers with active phase arrest of dilation compared with mothers with vaginal delivery. It is not yet clear if the altered contraction characteristics caused non-progressive labor or resulted from it. Lucovník et al. [42] have used a similar technology with preterm contractions to determine what characteristics of contractions discriminate preterm labor from false labor.

Nevertheless, these experimental data indicate the encouraging potential of wearable sensors to provide clinical advantage, particularly if they could provide additional information about the nature or strength of contractions and/or myocardial response to hypoxemia.

Internal sensors A second group of technologies includes specialized internal sensors. Examples include devices that measure fetal oxygen saturation levels or scalp electrodes that can measure fetal oxygen saturation aspects of the ECG waveform, such as the ST segment of the fetal electrocardiogram [5, 6, 21]. Both of these technologies are used in combination with EFM. Logically, we would expect that additional information about the oxygen-carrying capacity of the baby or myocardial reactions that indicate hypoxia would add to our understanding of the fetal cerebral condition. However, fetal O₂ levels in peripheral blood or fetal ST segment characteristics are still indirectly related to fetal cerebral oxygenation. Compensatory mechanisms can maintain oxygen delivery to the fetal brain cell despite lowered oxygen saturation levels and cerebral perfusion is preserved preferentially over cardiac perfusion. Moreover, some ST segment changes are not specific to hypoxemia and can be induced by normal fetal movement. Thus, these measures do not necessarily reflect actual fetal brain oxygenation [5, 21]. This underscores the critical importance of animal studies to fully understand their associations with metabolic acidemia before launching a clinical trial. A clinical trial using poorly selected thresholds may conclude with negative findings whereas the effect with a different threshold could be positive. A simple analogy might be an RCT based on a subtherapeutic dosage of an anticoagulant.

Step 2: association with metabolic acidemia

Traditional EFM The relationship between EFM and metabolic acidemia has a large literature, which will not be reviewed in detail here, as it is well summarized by several professional societies in the USA, UK, Canada, and Japan [1, 24, 41, 44, 51, 62]. This animal and human literature was the basis for defining what combinations of EFM patterns are most closely associated with metabolic acidemia and to describe a graded classification of increasing abnormality, which in turn determines the nature and urgency of clinical management. None of these classifications were published with their sensitivity or false-positive measures. This information is critical to understanding their potential clinical effect. High sensitivity of the classification is required to prevent a substantial portion of metabolic acidemia and good specificity is needed to avoid unnecessary interventions.

We have measured the sensitivity and specificity of one of the most methodical and detailed classification methods that defined five color-coded levels based on baseline variability, baseline level and the severity of recurrent decelerations [25, 51]. Red was the most abnormal, with orange, yellow, and blue defining progressively less concerning patterns. Green included only tracings with completely normal features.

We computerized these definitions and measured how often the various levels appeared in the last 3-h tracings of babies with metabolic acidemia defined by an umbilical artery base deficit over 8 mM [25, 60]. The digital tracings were analyzed using PeriCALM Patterns™ version 1.05 (PeriGen, Princeton, NJ, USA), an EFM pattern recognition software that identified and measured baseline, baseline variability, FHR decelerations, and contractions. Table 1 shows the sensitivity and specificity for each of the levels.

This performance is not stellar at first glance and bears discussion. These simple classification definitions do not consider the duration of tracing abnormality or trends over time. Another key problem is the assessment of baseline variability, which was a defining feature for each level of abnormality. Minute variations in baseline variability can change the classification level. A recent review of levels of clinician agreement on tracings classification using the ACOG Category definitions showed that disagreements were often due to how
clinician rated the baseline variability [3]. In addition, baseline variability measured electronically can be different from visual estimations. An example of this disparity is provided in Figure 1.

The tracing in Figure 1 was obtained about 20 min before the birth of a baby with Apgar scores of 0 and 4 and an umbilical artery pH of 7.03 with a base excess of -10 mM. Severely reduced baseline variability is clearly evident. Most clinicians would describe this as absent variability, but precise automated software measurement of variation around the baseline produces a value of 3.2 beats per minute (bpm).

The stated clinical significance of moderate variability is also problematic. For convenience, the relevant text about baseline variability from the updated NICHD definitions is reproduced in Table 2 [44].

The four tracings in Figure 2 were taken a few minutes before birth. Electronically measured baseline variability exceeds 5 bpm in each panel. All babies showed abnormal base excess levels in the umbilical artery at birth. These cases demonstrate clear exceptions to the statement that “moderate variability reliably predicts the absence of fetal metabolic acidemia at the time it is observed.”

To determine if these examples are typical or merely rare exceptions, it is necessary to examine a large number of cases systematically. The following graphs show summarize the analysis of baseline and baseline variability (measured electronically as a single complex) using PeriCALM Patterns™ from a series of term 3320 babies born with normal umbilical artery gases (base excess over -8 mM) and a group of 316 with base excess under -12 mM. We divided each tracing into 10-min segments, with 0 representing the last 10 min before birth. The mean and its 95% confidence interval for baseline and baseline variability in each time segment are shown in Figures 3 and 4.

On average, the baseline heart rate increased slightly and gradually as labor advanced especially in the babies with an elevated base deficit. Within these groups, some tracings developed low variability and others developed excessive variability. However, on average, baseline variability increased terminally in both groups, especially those with metabolic acidemia. These observations are consistent with experimental sheep and monkey literature, where heart rate

<table>
<thead>
<tr>
<th>Percentages of cases that reached</th>
<th>Sensitivity in the group with metabolic acidemia group (n=280)</th>
<th>Specificity in the group with normal umbilical artery gases (n=2132)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>2.9</td>
<td>98.3</td>
</tr>
<tr>
<td>Orange or above</td>
<td>30.0</td>
<td>81.9</td>
</tr>
<tr>
<td>Yellow or above</td>
<td>73.2</td>
<td>40.1</td>
</tr>
<tr>
<td>Blue or above</td>
<td>93.6</td>
<td>13.1</td>
</tr>
</tbody>
</table>

Table 1 Sensitivity and specificity of a modern feature- and rule-based classification method.

<table>
<thead>
<tr>
<th>Definition</th>
<th>Baseline FHR variability is defined as fluctuations in the baseline FHR that are irregular in amplitude and frequency. The fluctuations are visually quantitated as the amplitude of the peak-to-trough in beats per minute. No distinction is made between short-term variability (or beat-to-beat variability or R-R wave period differences in the electrocardiogram) and long-term variability, because in actual practice, they are visually determined as a unit. Hence, the definition of variability is based visually on the amplitude of the complexes, with exclusion of the sinusoidal pattern.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical significance</td>
<td>Moderate FHR variability reliably predicts the absence of fetal metabolic acidemia at the time it is observed. Minimal or absent FHR variability alone does not reliably predict the presence of fetal hypoxemia or metabolic acidemia.</td>
</tr>
</tbody>
</table>

Table 2 Current ACOG definition, classification, and clinical significance of baseline variability.

Figure 1 EFM tracing approximately 20 min before birth.
variability indices also increased with induced hypoxemia [8, 18, 28, 33, 46]. It also mirrors the known human cardiac response to increasing levels of catecholamine [45]. In sheep, reduced variability occurred only with very severe metabolic acidemia [8, 18, 46]. Thus, it appears that increased baseline variability can be associated with metabolic acidemia and may precede the development of the reduced variability that is often seen with very severe metabolic acidemia.

The purpose of presenting these examples was to demonstrate how better information could improve clinical guidelines. With computerized analysis of tracings now available, we can rigorously test hypotheses about EFM patterns, such as the optimal definition of reduced variability. Furthermore, it is also possible to separate the different frequency components of heart rate variability corresponding to the clinical concepts of short- and long-term variability [22, 38, 50, 64, 69]. Considerable debate remains regarding the best method to measure the various components of variability and what components or ratio of components are the most discriminating of metabolic acidemia in general [22, 50, 52, 64, 68].

**ST segment analysis (STAN)** Several small prospective human studies provide indications on the relationship between ST segment from the fetal ECG measurements and fetal acidemia. All of the studies described below relied upon ST segment and QRS complex measurements obtained with direct scalp electrodes. The STAN algorithm defines an ST event as an elevation in the relative amplitudes of the T-wave and QRS complex (the T/QRS ratio) or as an ST-segment morphology that is biphasic [61].

![Figure 2](image)

**Figure 2** EFM tracings taken within 4 min of birth in four term babies.
Metabolic acidemia was defined by an umbilical artery pH of ≤7.15 and a base deficit of ≥12 mM. Tracings were classified visually by two authors using the FIGO scoring system. Sensitivity for ST segment parameters with the FIGO classification was 43% (3/7) and specificity was 75% (101/136).

A later study by Costa et al. [16] examined the last hour of tracing in 148 term patients. Tracings were analyzed and classified using a computerized EFM pattern recognition program, Omniview Sis Porto 3.5 (Speculum SA, Alfragide, Amadora, Portugal), before applying the STAN algorithm. Metabolic acidemia was defined as an umbilical artery pH of ≤7.05. Sensitivity was 100% (7/7) and specificity was 94% (133/141) in this study. Vaysserie et al. [72] retrospectively examined 411 term tracings. When the outcome of interest was an umbilical artery pH of <7.15, sensitivity was 38% (41/108) and specificity was 83% (252/303). When a more stringent outcome (pH <7.05) was chosen, sensitivity rose to 62.5% (10/16) and specificity fell to 79% (313/395).

A larger retrospective review of tracings from 787 term tracings underlines the importance of interpreting ST events simultaneously with the EFM tracing [48]. ST events alone were associated with sensitivity for severe acidosis of 79% (19/24). However, specificity was only 50% in the 177 normal controls. Requiring an abnormal EFM tracing in addition to the ST event finding reduced sensitivity to 69% and improved specificity to 88%.

Even this small review quickly demonstrates that it is not easy to see the incremental benefit of a new technology. To draw some reasonable conclusions, the reader needs to see the performance for the same outcome and to be able to match specificity and then see the changes in specificity or vice versa.

To date there have been five prospective RCTs using ST segment technology in five different countries and a number of observational studies. Results have not been consistent across the five RCTs with respect to rates of operative delivery and reduction in measures of metabolic acidemia. In some studies (UK and Swedish), the operative interventions decreased in the STAN group but not in others (Finnish, French, and Dutch) [2, 49, 71, 75, 76]. Outcomes related to metabolic acidosis were defined differently across the studies. Again, an inconsistent pattern was seen, with reductions in metabolic acidosis related outcomes is reported in some studies (Swedish and Dutch) but unchanged in others (Finnish and French) and borderline in another (UK). Fetal blood sampling is the most definitive test of intrapartum acid base status, and its use would strongly affect outcome irrespective of the method of fetal surveillance. The utilization of fetal blood sampling varied greatly from study to study. In addition, focused educational efforts were part of the intervention, and this also affects outcome. Thus, it is not self-evident how STAN would perform in regions where fetal blood sampling is rarely done, such as in North America or where cesarean rates are much higher. Another RCT is underway in the USA with a projected enrollment of 11,000 women to determine if fetal ECG ST segment and T wave analysis change the rate of cesarean birth, operative vaginal delivery, or fetal compromise defined as a composite of several clinical and acid base markers [5].

**Innovative approaches to EFM analysis** We have focused on the association between metabolic acidemia and classical EFM parameters and the incremental benefit of additional parameters, such as measurements of the fetal P-QRS-T complex. We have seen that simple thresholds are inadequate. For example, FHR variability tends to rise before it falls and average baseline variability changes as labor advances. In addition, correct labeling of FHR parameters is essential especially for the correct usage of STAN because certain ST parameters are innocuous with some FHR patterns and hazardous with other FHR patterns. Correct labeling of EFM features can be difficult when decelerations are not exactly typical of the “late” or “variable” type. Inconsistent labeling among clinicians is well-known. Thus, there is need for a method that can address these limitations.

Automated FHR analysis can address labeling inconsistency, and many automated FHR analysis techniques have focused on careful detection of contractions and decelerations, mimicking visual clinical interpretation. However, even automated labeling of the tracing with classification showed poor discriminating performance as was outlined in Table 1 [25].

![Figure 3](image1.png)  **Figure 3** Mean baseline during the last 240 min before birth.

![Figure 4](image2.png)  **Figure 4** Mean baseline variability during the last 240 min before birth.
We have recently proposed a new approach to FHR analysis, based on system identification theory, that bypasses the detection phase altogether and focuses directly on the dynamic relationship between uterine pressure (as an input) and the FHR (as an output) [73, 74].

System identification is a signal analysis method that corresponds well to recognized fetal physiology. Many years ago, the antepartum contraction stress test (CST) was used to unmask vulnerability for hypoxia by applying a controlled stress via contractions to determine if they induced decelerations, revealing a potentially compromised fetal-placental unit [14, 55, 58]. In many ways, labor is the ultimate CST. Although the stress of contractions in labor is not controlled, uterine pressure can be measured. With system identification techniques, it is possible to characterize the response of the FHR relative to the uterine contractions it experiences. Rather than delineating individual events, this mathematical method results in a succinct characterization of the overall response of the FHR to uterine pressure within the time extent of the analysis. This FHR response relative to uterine pressure response was very different in babies born with symptomatic metabolic acidosis compared with babies with normal gases at birth.

This novel approach is well suited to current clinical practice. Although the FHR is subject to numerous influences, uterine pressure is the only input that is accessible by routinely used external monitoring; indeed, clinicians already interpret certain uterine pressure-FHR relationships as indications of pathology. Accounting for measurement noise and artifact, we estimated linear system dynamics in terms of an impulse response function (IRF), a model that represents very-low-frequency FHR energy (<0.03 Hz) related to uterine pressure, and is therefore complementary to other FHR components, such as baseline and variability. From this IRF model, we extracted two key parameters, gain and delay, that have direct clinical significance. Gain is an indication of the size of the FHR response, normally manifested as the depth of the deceleration relative to contraction amplitude. Delay is an indication of the timing of the response, normally observed as a lag between the onset of the deceleration compared with the beginning of the contraction.

We also recognize that there are other associations with metabolic acidosis, such as altered heart rate variability or bradycardia and have developed models to characterize baseline and heart rate variability and how they change over time. Finally, we have applied a machine learning method to consider the multiple factors including the IRF parameters, heart rate, and variability parameters as well as trends over time to classify the tracing as normal or pathological. These statistical methods are well suited to biological systems where the relationships between the outcome state and many interrelated time dependent variables are not linear.

In our study, this approach correctly classified more than half of the pathological cases, 1.5 h before delivery with a false-positive rate of 7.5%. Using the same data set, this method matched the sensitivity of a modern EFM feature- and rule-based approach and bettered its specificity by 15%.

**Step 3: clinical intervention**

As described earlier, clinical intervention, the last step in this hierarchy of dependent steps, is the only one that can directly change outcome. The effectiveness of clinical intervention is not only directly dependent upon the sensitivity and specificity of the surveillance parameters but also on the lead time it allows making intervention logistically possible.

In practice, the clinician must integrate clinical and monitoring information as well as project what is likely to happen and how quickly it might happen. This process of recognizing what is going on and what is about to happen is not unique to obstetrics. It has been studied extensively by psychologists and is called situational awareness [53]. Industries, such as aerospace or the military, where processes can degrade rapidly with dire consequences, were among the first to recognize the importance of situational awareness [36, 65, 66]. Situational awareness is now recognized as an essential skill for anesthesia and a critical skill throughout health care [27].

Situational awareness also has three basic and interdependent components that bear some similarities to the previous steps discussed with monitoring. However, rather than focusing on the physics of the monitor and the physiology of acidosis, these steps relate to the cognitive processes of the clinicians. Situational awareness depends upon perceiving the relevant information, understanding its meaning, and projecting what is about to happen.

“Sizing up” a situation is challenging in the presence of an overwhelming amount of information. The medical mind must focus on what is important and disregard the irrelevant. However, in doing so, the brain is vulnerable to well-described biases. The psychological phenomenon of “tunnel vision” refers to the tendency to perceive and confirm information that aligns with a particular viewpoint and discard contradicting information. Variations of this include “framing bias,” which refers to a tendency to create a coherent interpretation without examining all the available information, and “confirmation bias,” which refers to seeking only the information that supports a particular opinion. EFM assessments are also prone to these biases. The rarity of adverse outcome coupled with frequent “false alarms” further exacerbates the potential for bias.

Table 3 lists some of the impediments to situational awareness with some specific examples from labor and delivery. This list provides clues about how we can optimize the environment in labor and delivery to facilitate situational awareness among clinicians.

There is an important role here for intelligent electronic medical records or automated tracing analysis because they are not vulnerable to the subjectivity biases and impediments described above. Electronic medical records hold information on the evolution of the patient’s status and can display this in formats that help clinicians quickly grasp the situation. In addition, they can apply virtual checklists, perform trend analysis and escalate its alerts and reminders as the clinical condition warrants. The desirable characteristics of an intelligent medical record are well described by Hasely [31] and examples are in use today. Psychological testing has also
Table 3  Impediments to situational awareness.

<table>
<thead>
<tr>
<th>Perceiving information</th>
<th>Understanding its meaning</th>
<th>Projection</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Unreliable information – fetal monitor posts “maternal” heart rate as the “fetal” heart rate</td>
<td>• Knowledge gaps regarding basic physiology related to EFM</td>
<td>• Disregard of impending problem because it has a low incidence and false alarms are common</td>
</tr>
<tr>
<td>• Incomplete information or delayed communication – handover snapshots of selected portions of tracing rather than full review</td>
<td>• Knowledge gaps regarding the progression of coexisting/underlying diseases</td>
<td>• Time to actual delivery exceeds expectation</td>
</tr>
<tr>
<td>• Cluttered overcrowded spaces that impede visibility or access to critical information</td>
<td>• Erroneous assumptions</td>
<td>• Fixation on achieving vaginal birth</td>
</tr>
<tr>
<td>• Interruptions/multitasking</td>
<td>• Deference to authority figures despite a contrary personal opinion</td>
<td></td>
</tr>
<tr>
<td>• Limited time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Alert fatigue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Inexperience</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

contributed greatly to our understanding of the importance of simplicity and clarity of information display on computer screens. Overcrowded or confusing displays may themselves lead to medical error. High rates of false alarms lead to frustration and disregard of the device.

The future of fetal surveillance

The search for better physiological sensors and pathognomonic features will continue. There are exciting prospects on the horizon with new sensors and better ways of characterizing fetal condition. A large prospective RCT is underway in the USA to examine the impact of adding ST segment information to traditional EFM. Non-invasive detection of fetal ECG and contractions via surface electrodes on the mother’s abdomen are now available for clinical use. The largest ever perinatal trial examining the effects of computer-assisted tracing interpretation and management reminders at the bedside is underway in the UK [34].

Perhaps with greater understanding that clinical outcomes are related to the cumulative effects at each of the three hierarchical steps outlined in this review, we will pay more attention to performance at each level. Correcting deficiencies at each level will maximize the potential to improve outcome, which is the ultimate clinical goal.

In addition, there are important lessons to be learned from past experiences. We must use reliable tools, not merely visual inspection, to measure parameters of interest. We must take the time to fully understand the relationship between the parameters in question and fetal metabolic acidemia. Relationships are complex and the evolution of patterns over time may not proceed in simple linear fashions. Until these data are obtained, it is premature to define clinical management guidelines or launch a clinical trial using a threshold that is potentially suboptimal.

Finally, any clinical trial on fetal surveillance measures both the capacity of the device and the effect of clinician behavior. As the intention of the new monitoring device is to change clinician behavior to improve the state of the baby, the preferable clinical trial design would use randomization of clinician groups. For example, a clinician who uses a new adjunctive technique, such as computerized analysis and management suggestions, will gradually alter his interpretation of EFM. The computerized tool is providing constant and immediate feedback, which are powerful teaching tools for adult learners. Newly learned behavior will not completely disappear when caring for his next patient who has a similar tracing but does not have the computer analysis component. Unconsciously, the clinician will use these acquired skills with tracings that do not have computerized analysis. This behavior tends to bring benefit to the control group, causing “contamination” and diminishing the capacity of the study to ascertain differences between the groups. Contamination is less likely to occur when the randomized clinical groups are geographically separate.

The early 20th century mathematician David Hilbert, who laid some of the mathematical foundations underlying modern digital signal processing, said, “every real advance goes hand in hand with the invention of sharper tools and simpler methods.” In the case of fetal surveillance, the future is bright because many “sharper tools” are emerging with the convergence of innovations in engineering, mathematics, and behavioral science.

References
