HHS Public Access

Author manuscript

Acta Obstet Gynecol Scand. Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:

Acta Obstet Gynecol Scand. 2019 September ; 98(9): 1207–1217. doi:10.1111/aogs.13639.

Computer-based intrapartum fetal monitoring and beyond: a review of the 2nd Workshop on Signal Processing and Monitoring in Labor (Oct 2017, Oxford UK)

Antoniya GEORGIEVA[1], **Patrice ABRY**[2], **Václav CHUDÁČEK**[3], **Petar M. DJURIĆ**[4], **Martin G. FRASCH**[5], **René KOK, MD**[6], **Christopher A. LEAR**[7], **Sebastiaan N. LEMMENS**[6], **Inês NUNES, MD PhD**[8], **Aris T. PAPAGEORGHIOU**[9], **Gerald J. QUIRK**[10], **Christopher W.G. REDMAN**[9], **Barry SCHIFRIN**[11], **Jiri SPILKA**[3], **Austin UGWUMADU**[12], **Rik VULLINGS**[13,*]

[1]Nuffield Department of Women's and Reproductive Health; Big Data Institute, University of Oxford, Oxford, UK [2]Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France [3]CIIRC, Czech Technical University in Prague, Prague, Czech Republic [4]Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY, USA [5]Department of Obstetrics and Gynecology, University of Washington, Seattle, WA, USA [6]Nemo Healthcare, Veldhoven, the Netherlands [7]Department of Physiology, University of Auckland, Auckland, New Zealand [8]Department of Obstetrics and Gynecology, Centro Materno-Infantil do Norte - Centro Hospitalar do Porto; Instituto de Ciências Biomédicas Abel Salazar, Centro de Investigação em Tecnologias e Serviços de Saúde; Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal [9]Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK [10]Department of Obstetrics and Gynecology at Stony Brook University Medical Center, Stony Brook, NY, USA [11]Encino, California, US [12]Department of Obstetrics & Gynecology, St. George's University of London, London, UK [13]Eindhoven University of Technology, Dept of Electrical Engineering, Eindhoven, The Netherlands.

Abstract

The second Signal Processing and Monitoring in Labor workshop gathered researchers who utilize promising new research strategies and initiatives to tackle the challenges of intrapartum fetal monitoring. The workshop included a series of lectures and discussions focusing on: new algorithms and techniques for cardiotocogoraphy (CTG) and electrocardiogram acquisition and analyses; the results of a CTG evaluation challenge comparing state-of-the-art computerized methods and visual interpretation for the detection of arterial cord pH<7.05 at birth; the lack of consensus about the role of intrapartum acidemia in the etiology of fetal brain injury; the differences between methods for CTG analysis 'mimicking' expert clinicians and those derived from 'data-driven' analyses; a critical review of the results from two randomized controlled trials

Corresponding author: Antoniya Georgieva, L3 Women's Centre, The John Radcliffe Hospital, Oxford OX3 9DU, UK, antoniya.georgieva@wrh.ox.ac.uk.

*All contributing co-authors are listed in alphabetical order

Conflicts of interest

B.N.L. and R. V. are shareholders in Nemo Healthcare BV, the Netherlands. R.K. is employed by Nemo Healthcare BV; M.G. F. is an inventor of related patent application entitled "EEG Monitor of Fetal Health" including U.S. Patent 9,215,999. M.G.F. co-filed a patent for the aECG method referred to in this article. The remaining authors report no conflict of interest.

testing the former in clinical practice; and relevant insights from modern physiology-based studies. We concluded that the automated algorithms performed comparably to each other and to clinical assessment of the CTG. However, the sensitivity and specificity urgently need to be improved (both computerized and visual assessment). Data-driven CTG evaluation requires further work with large multicenter datasets based on well-defined labor outcomes. And before first tests in the clinic, there are important lessons to be learnt from clinical trials that tested automated algorithms mimicking expert CTG interpretation. In addition, transabdominal fetal electrocardiogram monitoring provides reliable CTG traces and variability estimates; and fetal electrocardiogram waveform analysis is subject to promising new research. There is a clear need for close collaboration between computing and clinical experts. We believe that progress will be possible with multidisciplinary collaborative research.

Keywords

Health data; sensitivity; specificity; artificial intelligence; Cardiotocography; electronic fetal monitoring; intrapartum care; hypoxic-ischemic encephalopathy

INTRODUCTION

Continuous fetal heart rate (FHR) monitoring with cardiotocography (CTG) remains the mainstay of intrapartum fetal surveillance. Its intended goal is to prevent death or hypoxic-ischemic encephalopathy, by identifying incipient hypoxia/ischemia in a previously healthy fetus, at a time when intervention can prevent or mitigate permanent injury. Its goal is not considered to be the identification of infection during labor or trauma related to the delivery. While pre-existing injury, anomaly or antenatal hypoxia may be suspected, on the basis of a deviant FHR pattern, it is unknown whether the injury associated with these factors can be mitigated or aggravated by obstetrical care during labor. Nonetheless, these factors modulate the risk of neural injury and can influence the FHR trace, along with the timing and urgency of delivery. In addition, hypoxic-ischemic encephalopathy does not represent either a single clinical picture or pathogenesis. Injury evolves over hours and its physiological manifestations (in FHR changes or other signs) are not confined to a single short period of time[1]. On the other hand, some injurious events develop precipitously and show non-standard FHR and/or neuroradiological patterns[2]. These occurrences may not be preventable, but emergency operative delivery may still be warranted. For example, in order to start hypothermia treatment promptly; or in order to prevent exacerbating the injury by allowing the labor to continue.

Evidence has been available for nearly half a century that FHR abnormalities are often not related to fetal acidosis[3,4]. As a result, the rate of unnecessary operative deliveries is high[5,6] with adverse implications for this and/or subsequent pregnancies. Furthermore, even experts disagree with each other when interpreting the intrapartum FHR (for example, Chauhan et al[7]). Clinical guidelines that focus on the CTG morphology in isolation, encourage clinicians to conclude that every FHR pattern has a single and straightforward correct interpretation[8]. In reality, often complex interactions between multiple clinical factors need to be taken in to account.

Recent randomized clinical trials (RCT) for novel intrapartum monitoring technologies have largely failed to show improvement in neonatal outcomes when compared to current visual-based CTG monitoring[9–11]. But large clinical retrospective studies of intrapartum computer-aided FHR interpretation (mimicking clinical experts) in conjunction with other technologies, are more optimistic[12,13] (38,466 and 78,456 births in the respective studies). With no other surveillance modalities on the horizon and the raise of artificial intelligence and data analytics for healthcare, now is the time to create opportunities for multidisciplinary joint efforts, pushing new boundaries in an attempt to move forward. For example, FHR algorithms that harness the complex information from large clinical datasets are under development and have never been tested in the clinic: Georgieva et al 2017, using data from 22,000 births demonstrated that a very simple data-driven system compares favorably to clinical CTG evaluation in prediction of several types of poor neonatal outcome. With larger datasets and more sophisticated methods, we expect that such systems would evolve to play an invaluable role in the future of intrapartum FHR analysis. Further research is needed to test this hypothesis.

But these algorithms cannot be developed in a vacuum. Indeed, a number of related factors require consideration. Obstetrical decisions relate to two independent organisms where efforts to benefit one may be harmful to the other; the preventability of injury; the feasibility of safe vaginal delivery; the trajectory of deterioration; the need to defend against allegations of malpractice; the need for urgent delivery and the maintenance of a reasonable frequency of cesarean delivery; and more. But there is a clear need for reliable and objective technologies to monitor and help clinicians identify the fetus at risk during labor.

This gives the motivation for the bi-annual Workshop on Signal Processing and Monitoring in Labor (SPaM), where a broad range of multidisciplinary new research strategies and initiatives are considered. The aims of the SPaM Workshop are to provide ”a truly multidisciplinary forum and spirit and common language for clinicians, physiologists and signal processing researchers” (Workshop webpage: https://www.wrh.ox.ac.uk/research/spam-in-labour). Registration through the website of the workshop is open to everybody willing to attend. Opportunity is given to anyone willing to give a talk on a relevant subject with ample time for discussions and questions. There have been no abstract submissions or proceedings so far. We gathered participants from Europe, USA, Canada, and New Zealand, including experts from obstetrics, engineering, computer science; machine learning; physiology, physics and epidemiology. There were representatives from the producers of fetal monitors (Phillips, GE/Monica Healthcare, Huntleigh Healthcare, Nemo Healthcare and Laerdal Medical) and medical or engineering students.

This paper is designed to provide a current review of computer-based methods for continuous fetal monitoring and relevant clinical trials; set a benchmark of state-of-the-art performance of such methods on multicentre data; critically examine the most important obstacles and controversies impeding progress; sign-post the reader to the newest relevant research; and outline the most promising future directions for research.

STRUCTURE AND MAIN TOPICS OF THE WORKSHOP

In this paper, we review the main themes and ideas of the meeting, rather than individual presentations, which are available through the workshop webpage[i] (including video recordings of the talks). The workshop discussions gravitated around the following main topics:

1. The concepts of computerized FHR monitoring: 'mimicking' clinicians versus 'data-driven' interpretation;
2. The SPaM'17 CTG challenge: evaluation of computerized methods with respect to cord acidemia at birth and visual clinical interpretation;
3. Defining the endpoint(s) of intrapartum monitoring: the 'good' and 'bad' labor outcomes;
4. Signal processing methods for data-driven computerized CTG;
5. Visual-aid computerized CTG: lessons from RCTs;
6. Trans-abdominal fetal ECG monitoring;
7. Fetal physiology in the age of big data;
8. Beyond FHR: emerging technologies and adoption challenges;
9. Bringing new monitoring modalities to the market.

THEMES HIGHLIGHTED DURING THE WORKSHOP

1. The concepts of computerized fetal heart rate monitoring: 'mimicking' clinicians versus 'data-driven' interpretation

Two approaches to computerized CTG analysis have been developed in the past: (a) systematic visual assessment to identify classic CTG features such as baseline, variability, decelerations/accelerations, contractions, with alerts based on clinical guidelines; (b) data-driven methods that use routinely collected CTG and fetal health data to investigate signal processing-based metrics and/or interpretation methods that best relate to measures of fetal wellbeing and labor outcome. The data-driven concept is based on analyzing a very large number of records (at the very least a few thousands, but as the next decade of 'big data' unfolds, even millions of CTGs); with the capacity of learning to recognize rare events, such as hypoxic-ischemic encephalopathy. Below in theme 4, these emerging methods are reviewed in more detail and the reader is signposted to the relevant literature.

The visual-aid methods, which have been evaluated in two RCTs discussed below, have the potential to reduce the lack of objectivity, limited availability or training of clinical staff, and human error/fatigue. In contrast, with larger FHR/CTG datasets and increasing computing power, researchers can now focus on data-driven methods (reviewed below). Data-driven approaches tackle the same issues as the visual-aid approaches but also directly aim to improve the detection of infants at high risk for intrapartum compromise, learning from vast amounts of data. Such methods are still undergoing development and none have been tested

in clinical studies yet. Therefore, we underline the fundamental difference between the two computer-based approaches to FHR evaluation.

2. The SPaM'17 CTG challenge: evaluation of computerized methods with respect to cord acidemia at birth and visual clinical interpretation

A multicenter dataset was assembled before the workshop, and is now available at the Workshop webpage[i]. This SPaM'17 database comprises 300 labors collected from three participating centers (Lyon, France; Brno, Czech Republic; and Oxford, UK). Each center provided 100 cases: 80 with arterial cord pH in with values within 7.25–7.30 and 20 with arterial cord pH<7.05. Only cases with validated cord gases were included (arterial pH < venous pH – 0.02). The goal was to "evaluate state-of-the-art algorithms for prediction of low umbilical artery pH at birth from intrapartum CTG recordings". Included were only singleton pregnancies of more than 36 weeks gestation. Excluded were small for gestational age babies and those with congenital abnormalities. Selected were only CTGs longer than 60 minutes, with less than 15% signal loss and ending within ten minutes of birth. All CTGs were acquired as part of standard clinical practice at the respective units, with standard CTG monitors and the FHR was only available to the researchers uniformly sampled at 4Hz, as per standard manufacturing. Teams were allowed up to five submissions and were informed about their 'score', calculated as the square root of Sensitivity (Se) multiplied by Specificity (Sp). Se and Sp were also reported and discussed at the workshop.

Overall, eight teams participated and had comparable performance. Two teams used the SPaM'17 dataset as a 'training' dataset, informing their algorithms development according to the obtained score and two teams used the dataset to test new applications of algorithms, still under development (no relevant publications). The remaining four teams used established algorithms for CTG analysis that had been developed and tested prior to the challenge: SisPorto 4.0 incorporating the International Federation of Gynecology and Obstetrics (FIGO) clinical guidelines[14]; Phase Rectified Signal Averaging (OxSys 1.0), tuned using over 7,000 CTGs from Oxford[5,15]; Sparse Support Vector Machines trained using over 1800 CTGs from Lyon[16,17]; Auto mutual information applied only to the last 30min of first stage labor, and was tuned using over 900 CTGs from Lyon[18]. Summary statistics of the Se and Sp values for the latter four methods are shown in Table 1. There was a substantial variance of performance of all algorithms with respect to where the data came from (Oxford, Brno or Lyon), underlying the importance of multicenter datasets.

Twenty selected CTGs from the 300 were also evaluated by five clinicians, blinded to outcome. Selected were the CTGs for which the automated methods had the highest disagreement between themselves and/or majority did not predict correctly the labor outcome. Ten of the 20 had arterial cord pH < 7.05. The obstetricians were asked to identify the cases which are likely to have arterial cord pH < 7.05 and/or require early operative delivery. No other information, but the CTG was provided. The images of CTGs were provided in both EU standard (1cm/min) and US standard (3cm/min), one US-based and four EU-based doctors responded to the challenge. Four of the clinicians are currently practicing obstetricians consultants in Europe (UK, Portugal and the Czech Republic) and one is US-based. They all have both clinical and academic interest in CTG interpretation.

The main results from the visual evaluation for the prediction of arterial cord pH < 7.05 could be summarized as follows:

- All five clinicians agreed with each other in five out of the 20 CTGs; only four clinicians agreed with each other for nine out of 20; and only three clinicians were in agreement for the remaining six CTGs.
- Majority vote (MV) of the automated methods agreed with MV of the clinicians in 85% (17 out of 20) of the selected CTGs. For the remaining three cases, the clinical MV correctly predicted the outcome whereas the automated methods' MV did not.
- MV(algorithms) and MV(clinicians) agreed but did not correctly predict the labor outcome in six of the 20 cases: four were false negatives and two were false positives.

The controversial nature of CTG evaluation was evident not only in the results of the challenge but during their discussion at the workshop's dedicated session. Firstly, defining labor outcome by cord gas analyses was deemed unsuitable by some of the participants (as discussed in detail below in Theme 3). Secondly, the clinicians were visibly challenged to discuss their CTG interpretation in this kind of competitive manner, especially with no clinical context provided (but this was required in order to compare fairly to automated methods). Moreover, many of the active clinical participants in the workshop did not submit formally their evaluations.

We concluded that, generally speaking, the automated algorithms perform comparably to each other and to clinical assessment of the CTG (with no clinical context). It is important that data from multiple hospitals is used. Most importantly, Se and Sp urgently need to be increased by future work (both computerized and visual assessment). But we note that the optimal Se and Sp that (in theory or practice) could be achieved with the best possible CTG interpretation, remain unknown and heavily depend on the way fetal compromise is defined (discussed in the next section).

3. Defining the endpoint(s) of intrapartum monitoring: the 'good' and 'bad' labor outcomes

Without defining and quantifying the outcomes of labor that we want to prevent, we cannot assess whether one method for monitoring is better than another, and ultimately cannot progress. The issue of how to define the relevant outcomes has been one of the most important and debated problems during the SPaM workshops. There remains no clarity or consensus about the goal of intrapartum CTG monitoring, in particular, the role of intrapartum acidemia in the etiology of fetal brain injury which remains one of the biggest challenges in data-driven CTG analysis. We summarize below the discussions relating to this issue and provide relevant references to sign-post the reader.

About two thirds of speakers at the workshop used a single cord gas parameter in their research (pH or base deficit). Fetal acidosis at birth is a risk factor for clinically important neurological outcomes such as convulsions and neonatal encephalopathy, but a large proportion of newborns with very low pH / high base deficit values are clinically well and do

not require special neonatal care[6,19]. The remaining third of researchers used hypoxic-ischemic encephalopathy (HIE) or composite outcomes as endpoints of intrapartum CTG. The adopted definition of HIE was confounded by including acidemia as a requirement for HIE (discussions of the ways to define HIE can be found in Kurinczuk et al.[20], ACOG 2014[21], Leviton et al[22] and the references therein). Importantly, the acidemia prerequisite for HIE diagnosis, leaves out of the analysis the newborn with brain injury, but normal cord gases at birth. Highlighted was the importance of complex intrapartum factors such as maternal fever, infection, meconium, growth restriction, pre-existing injury. These and/or other risk factors may perturb fetal cerebral function and FHR patterns in the absence of acidemia. Further understanding of these and other factors, their effect on intrapartum FHR and how they modulate the utility of current/future biomarkers of hypoxic-ischemic injury is urgently needed. If the importance of cord blood gas analysis is to be reduced, what indices of labor outcome can be objectively measured to assess the impact of improved CTG analysis? And, from the point of view of data-driven methods, there may well be the need for different algorithms to detect these heterogeneous pathophysiological mechanisms, which may manifest themselves differently in the FHR pattern. In other words, we need to understand the different pathways to fetal injury and, for each of these 'models' of injury evolution, identify the appropriate role of various CTG patterns.

These complexities underline the clear need for close collaboration between computing and clinical experts, and the need for research/time before tangible results could be achieved.

4. Signal processing methods for data-driven computerized CTG

Figure 1 provides a general overview for the concepts in applying computers for signal processing and machine learning for data-driven computerized CTG. We note that the principles illustrated here are by no means unique to CTG, but they underline most situations where we teach the computers to recognize and analyze complex real-world information, i.e. developing artificial intelligence for decision support. The key concepts are: availability of data to learn from (training and testing datasets); signal feature extraction and selection (finding the most important information in the data); machine learning to classify the CTG signals (use that information to achieve as good as possible accuracy, i.e. develop mathematically *optimal* methods); and the relatively new and quickly developing field of deep learning where feature extraction, selection and machine learning are (in a way) fully automated and *optimized* by a super-computer.

(1) Feature extraction—Clinical guidelines for CTG analysis are derived from FIGO or similar sources[23]. They focus on the three main FHR characteristics, namely, baseline rate (level, trend), decelerations/accelerations (frequency, shape, depth, duration), and variability (long- or short-term). It is well appreciated that visual assessment of these criteria suffers from significant intra- and inter-observer variability[24] and leads to unnecessary operative deliveries[25].

Besides the automated analysis of temporal patterns inspired by FIGO criteria, FHR temporal dynamics have been further assessed by spectrum estimation; i.e., the repartition of energy as a function of frequency, in the range 0.04 to 2Hz, with the calculation of high

frequency (HF) and low frequency (LF) energies or LF/HF ratio, quantifying a putative sympathovagal balance in fetuses[26]. Further, more advanced statistical signal processing tools have been used, probing either nonstationarities (time-frequency), scale-free dynamics (wavelets, fractal) and/or nonlinear mechanisms and hence dependencies beyond the mere correlation (bispectrum, information theoretic measures, entropy rates, multifractal analysis, scattering transforms, conditional statistics as in the Phase Rectified Signal Analysis method[15,16,18,27–30]). These result in many different FHR features/characteristics. Yet, it is generally well accepted that, on their own, FHR changes are not enough for use in clinical practice and a recent review re-emphasized the need for the wider clinical picture in the interpretation and management of FHR patterns during labor[31]. Therefore, Georgieva et al.[5] presented a basic system prototype to assess the CTG, based mainly on one feature (the Phase Rectified Signal Averaging), but adjusted the analysis to clinical risk factors and achieved performance comparable to that of clinical practice (applied to the CTGs of over 22,000 births). Georgieva et al.[5] demonstrated that such 'hybrid' approaches (that combine CTG metrics with other risk factors) are likely to lead to better performance of the classifiers, but require reliable documented clinical data (e.g. that of maternal comorbidities).

(2) Feature selection and classification (machine learning)—Machine learning for classification into normal or abnormal has included several methods for automated 'classification', such as Bayesian[32], Support Vector Machines[33,34], Random Forest, and Artificial Neural Networks[35], and often involved almost all proposed features. However, different features may often provide redundant/correlated information. Whereas, joining feature selection and optimal classification into one step, has demonstrated that a few well-chosen features are as efficient as rules involving many features[16,17,34]. In addition, it can be shown that FHR interpretation rules for the first and active-pushing stages of labor may need to be different[16,17].

(3) Deep Learning approaches—Finally, FHR analysis would benefit from the wider availability of large, documented and shared databases that permit robust evaluation of detection/performance of different machine learning methods. In particular, such data could be used with approaches based on 'deep learning' strategies[36,37]. These are powerful, next generation artificial intelligence approaches, where the computers independently seek what is important in the data and do not rely on human-derived features/characteristics. For example, preliminary work by Petrozziello et al. 2018[36] demonstrates that convolutional neural networks, trained with the data from 30,000 labors at term, have promising performance; and their accuracy and robustness increase with the size of the training dataset[36].

In summary, data-driven FHR interpretation needs further research and will benefit from larger datasets based on well-defined labor outcomes which include clinically relevant risk factors and co-morbidities. With a longer-term view and the introduction of full electronic patient records in the developed countries, we anticipate that CTG databases available for research could grow to hundreds of thousands of cases, and even millions. In the UK alone, over 300,000 labors at term are monitored with CTG each year.

5. Visual-aid computerized CTG: lessons from RCTs—Two systems comparing computerized CTG analysis to visual assessment were recently evaluated in multicenter randomized trials in the United Kingdom. The FM-Alert trial[21] assessed the Omniview-SisPorto® 3.5 system[36] in 7,730 pregnant women at term. The system analyses both CTG and ECG signals (ST-analysis) and provides real-time alerts when features are suggestive of fetal hypoxia/acidosis. There was no overall reduction in the rates of metabolic acidosis (16 (0.40%) vs 22 (0.58%), RR 0.69 [0.36–1.31]) nor in obstetric interventions; in higher risk groups of women, with complications preceding or arising de novo during pregnancy, it may confer benefit. There was a lower than expected rate of newborn metabolic acidosis – the primary outcome - in both arms of the trial.

The INFANT Trial[11] evaluated the INFANT (K2 Medical Systems) – decision-support software[38,39]. In this unmasked RCT, 47,062 pregnant women were randomly assigned to the computerized CTG group or the standard clinical care group. No differences were found in the incidence of poor neonatal outcome (a composite outcome for intrapartum stillbirth or, early neonatal death, or neonatal encephalopathy, neonatal unit admission within 24h for ≥ 48h with evidence of feeding difficulties, respiratory illness, or encephalopathy with evidence of compromise at birth) between the groups (172 (0·7%) vs 171 (0·7%), adjusted RR 1.01, 95% CI 0.82–1.25)). At two years of age, no significant differences were noted in developmental scores.

Neither trial found evidence for benefit from using the computerized systems to improve neonatal outcomes. Both encountered the same issue of a lower than expected incidence of the primary adverse outcome. This may have been caused by staff learning from exposure to the decision-support arm of the trial, resulting in improved outcomes in the control arm (Hawthorne effect). Other methodological issues including multiple comparisons, strengths of findings and plausibility, and the evaluation of the significant findings in secondary outcomes were discussed at a dedicated workshop session with the following conclusions:

1. CTG in labor is a screening tool, not a therapeutic device or drug. It is unlikely to improve outcomes unless it is followed by effective, necessary interventions. This depends on many factors including clinician performance, resources and training. Neither trial defined clinical actions in response to the alerts. It is also not clear if alerts were ignored: in the INFANT Trial, the median time from identification of the last “red” alert to birth, was 13–279 min across different units, suggesting disparate approaches to severe CTG abnormalities. A standardized, specified list of actions (such as delivery or not; discontinuation of oxytocin infusion) in response to well-defined CTG and uterine activity patterns, combined with post-hoc monitoring of compliance, could be considered in future trials.
2. The Hawthorne Effect. If both INFANT and FM-Alert improved staff FHR interpretation skills, these benefits may be extended to the care of women in the control arms thereby contaminating them. This could be prevented in future trials by using a cluster randomized methodology.

3. Future RCTs should have a period of pre-randomization benchmarking of outcome data. In these trials there was a missed opportunity to evaluate outcome data from low risk women who did not have continuous monitoring and from high-risk women who had continuous monitoring with standard care only[40].

4. The studies reported a very low incidence of poor perinatal outcomes, including stillbirths, perinatal brain damage or neonatal death, compared to rates found in UK Obstetric Units[41]. Similar findings have been reported by others[42,43] meaning studies may be underpowered to detect the predefined differences. It appears that women participating in studies evaluating intrapartum fetal monitoring have better perinatal outcomes; this suggests that there is an urgent need to improve knowledge and training regarding appropriate responses to CTG abnormalities in general.

5. Do the specific algorithms used by the evaluated systems simply not work? Both systems identified fetal heart-rate abnormalities; however, the alerts did not take into account other data from labor such as its progress, maternal fever, chorioamnionitis or the presence of meconium. More importantly, algorithms mimicking clinicians cannot be expected to work better than the best clinical expert. This is unlike data-driven algorithms that have the potential to learn from vast amounts of past data. Unfortunately, there was no provision to store and make data available to research beyond the RCT in INFANT. However, the team behind the FM-Alert trial are putting efforts to obtain and use the data from their trial for further research.

In summary, we should consider the discussed complexities when designing intrapartum fetal monitoring RCTs. Large and well conducted clinical observational studies [12,13] may be a technical and financially attractive alternative to produce good scientific evidence for new technologies in the field.

6. Trans-abdominal fetal electrocardiogram monitoring

The fetal electrocardiogram (ECG) can be measured non-invasively by positioning electrodes on the abdomen of a pregnant woman[44–46]. The key challenge is to suppress interference from the maternal ECG. Recent studies confirm that this challenge can be successfully addressed by advanced signal processing methods[47,48]. Assessment of the fetal ECG has the potential to provide accurate information on FHR variability and enable ECG morphology analysis. Especially via computerized analysis[5,49–51], fetal ECG may be able to provide risk assessment on fetal compromise better than possible when using FHR variability obtained via Doppler ultrasound[52].

In particular, CTG is bound by a lower temporal precision in identifying beat-to-beat information (fetal heartbeat information derived with ultrasound or fetal scalp electrode by standard monitors at a uniform frequency of ~4Hz). True beat-to-beat data is needed for some of the derivative estimates of the FHR variability with a potential to improve the prediction of neonatal outcomes[53]. Its utility remains to be investigated in the future with the advance of transabdominal technologies and appropriate data collection. In addition, the question remains whether the within monitor internal pre-processing of the data, which may

vary from vendor to vendor, has any impact on automated FHR variability estimation. But technological advancement in isolation may not solve the utility of FHR variability and we need to understand in what situations it is useful, given the complex nature of FHR variability and its relation to fetal compromise (discussed in the next section). Also, any risk assessment could be complemented with further screening for metabolic acidosis[54,55] or even congenital heart disease[56].

To complement analysis of FHR variability, morphological analysis of the fetal ECG could potentially play a crucial role. This remains to be investigated, and we underline the need to collect relevant data that can be used to address this question. Furthermore, because fetal position affects ECG morphology analysis and the positioning of abdominal electrodes is unknown with respect to fetal orientation (i.e. fetal heart axis), multi-lead ECG recordings are necessary to correct for fetal orientation. Multiple leads allow correction for inter-patient variations in, for example, the fetal electrical heart axis, which may be vital for the detection of fetal metabolic acidosis[44].

Finally, transabdominal electrophysiological monitoring can also provide information on uterine activity during labor. Conventional CTG interpretation considers variations in FHR in presence/absence of uterine contractions and recent studies have shown that also computerized CTG interpretation benefits from reliable information on uterine activity[36,37]. The current standard for non-invasive assessment of uterine activity relies on the tocodynamometer, a strain gauge held in place on the abdomen by a belt. This tocodynamometer typically provides low-quality signals[57] and its performance degrades for higher body-mass-index (BMI) of the laboring women[58]. With transabdominal monitoring, information on uterine activity can be calculated that has been shown to have good correlation with intra-uterine pressure[59,60]. Moreover, the quality of the transabdominal uterine signals is not affected by BMI[61].

Compared to standard fetal monitoring technologies, transabdominal fetal ECG monitoring is completely noninvasive and entirely avoids the issue of transducer re-positioning[58].

To sum up, transabdominal fetal ECG can already estimate FHR variability more accurately than CTG. It is truly non-invasive and does not expose the fetus to ultrasound waves. The capability for morphological analysis of fetal ECG waveform is advancing, but there needs to be more research in the interpretation of these waveforms, with or without associated FHR analysis.

7. Fetal physiology in the era of big data

The original introduction of CTG was tightly linked to the underlying physiology and indeed many of the pioneers of CTG undertook seminal physiological studies[62–64], which continue to guide the way clinicians, physiologists and now engineers and computer scientists approach CTG interpretation. One of the founding principles of CTG interpretation was that decelerations have multiple etiologies some of which are benign. Distinguishing these ‘benign’ early decelerations from ‘pathological’ late decelerations and the sometimes-pathological variable decelerations was therefore important[62,64,65]. Another key principle

was that suppressed FHR variability appeared to be the most consistent predictor of neonatal acidemia[63].

This original physiology-driven approach is now being refined by data-driven approaches. Intriguingly, the findings that have emerged to date from large retrospective and prospective clinical studies[4,66] are not consistent with some of the guiding principles of CTG. For example, the majority of acidemic fetuses do not show suppressed FHR variability, and can even show increased FHR variability[4,15,67–69]. Moreover, the morphological or timing-based classification of decelerations is not associated with acidemia (cord pH < 7.1), but in contrast the assessment of the depth and duration of *all* decelerations via metrics such as 'total deceleration area' or 'deceleration capacity' is associated with acidemia[4,15,34,67].

These modern clinical findings are not minor tweaks to CTG interpretation, but illustrate that the physiology underlying CTG is not what we were taught to expect. This is paralleled by a much deeper understanding of fetal physiology, through well-designed animal studies[70]. Many of these animal studies involve highly structured sequences of repeated umbilical cord occlusion modelling the intermittent nature of hypoxemia associated with uterine contractions[71], but are in no way designed to truly reflect the complex and dynamic nature of human labor. Nonetheless, they allow the physiological mechanisms underlying specific features of the FHR trace to be investigated. With these caveats in mind, it is striking that the lessons learned from these animal studies have converged with the lessons from recent clinical studies. For example, animal studies reveal that the regulation of FHR variability during labor is complex and cannot be reduced to a binary system in which high equals good and low equals bad[71,72], paralleling the inconsistent predictive value of FHR variability in clinical studies[4,15,67–69]. Modern understanding of the fetal peripheral chemoreflex (the response to hypoxemia) and the fetal baroreflex also concludes that the peripheral chemoreflex is the most likely mediator of the vast majority of intrapartum decelerations[71], while the common belief that many decelerations are explained by baroreflex activation is in fact false (discussed in detail in [71]). The clinical relevance of this physiological understanding is that the assessment of the depth and duration of *all* decelerations should provide a broad measure of the total burden of fetal hypoxemia, while the timing of decelerations is a red herring that does not reflect the etiology mediating decelerations in the way it was once thought. Indeed, recent clinical studies have highlighted that metrics consistent with these concepts (total deceleration area and deceleration capacity) are the most predictive metrics identified to date of cord acidemia and outperform traditional timing-based assessment of decelerations[4,15,34,67].

It is now apparent that we need to be cautious of preconceived ideas about which FHR patterns are associated with fetal wellbeing. A data-driven approach therefore has an important role in the future refinement of CTG and represents the best opportunity to date to retest, reconfirm or refute long-held beliefs about the utility, and even the physiological meaning of FHR patterns.

8. Beyond fetal heart rate: emerging technologies and adoption challenges

Whilst we acknowledge the importance of capturing information on reduction of cerebral blood flow or fetal infection/inflammation, which are recognized precursors of fetal brain

injury, we lack the clinical tools to make these observations reliably during labor, directly or indirectly. Preclinical studies however do indicate that inflammatory indices may be developed from heart rate variability derived from the high-precision abdominal ECG that could potentially serve as real-time non-invasive infection biomarker[73–75]. The clinical adoption challenge to validate these results in prospective studies lies in the currently limited availability and use of the fetal ECG devices (the main motivation to using transabdominal fetal ECG devices remains its superiority for monitoring obese mothers, rather than potential signal-analytical benefits of higher quality fetal ECG signal). The technology is currently available only in few delivery centers with a limited number of devices.

Secondly, there is a resurgence of interest in using fetal EEG monitoring during labor[76–78], to make neurological assessment of the fetus. This approach continues initiatives that began more than 40 years ago[79,80]. Fetal EEG monitoring can be performed using the routine fetal scalp electrode or specially designed electrodes[78,81]. It represents a clinically testable modality for increasing the accuracy of early detection of fetal acidemia[82] or cerebral blood flow[79,80]. In particular, there appears to be an early response in the fetal EEG to repetitive FHR decelerations accompanied by pathological hypotensive blood pressure responses induced by brief umbilical cord occlusions[83] which warrants further investigations.

Finally, to be taken up by practitioners, and to produce a change in health outcomes, the right idea has to arrive at the right time and be met by a technology that can implement it[52,84,85]. Proving the clinical utility of transabdominal ECG or fetal EEG is a prerequisite to their wider adoption. Although the technologies seem to have arrived, a paradigm shift will likely be required for practitioners to embrace them.

9. Bringing new monitoring modalities to the market

Innovations in fetal monitoring and surveillance are difficult to implement into clinical practice for three main reasons. Firstly, maternity departments spend most of their money on personnel rather than investments in technology or medical devices. With budgets for departments decreasing in most of the world, such investment in technology is even more difficult nowadays. Secondly, innovations are hindered by a paradox in the reimbursement system: in countries where hospitals have a revenue per admission or treatment, the net revenue of the maternity department may decrease due to innovations that help reduce the need-to-treat. At the same time, global costs of healthcare would be reduced by successful fetal monitoring innovations, but the cost savings might be made elsewhere in the system – sometimes close to maternity wards, e.g. in neonatology departments; or sometimes much further away, e.g. reduced costs related to care for those with neurological or educational problems. Especially in the latter case, maternity wards are not strongly inclined to invest their already limited budget in innovative technology.

Moreover, to allow reimbursement for new technology, insurance companies and national healthcare systems need to see evidence that cost savings can be made. But without such reimbursement for the new technology, hospitals will often not introduce it in their maternity wards, because of the abovementioned budgeting considerations. This makes it difficult to prove the technology’s cost efficiency or its benefits for the women and newborns. This chicken-and-egg problem can be resolved by clinical trials, but these are subject to multiple

drawbacks, some of them specific to the field of fetal monitoring, discussed in detail in Theme 5 above.

CONCLUSION

The 2nd SPaM Workshop provided a unique and stimulating research forum for academics, clinicians, and industry. We concluded that modern automated algorithms for CTG evaluation perform comparably to each other and to clinical assessment of the CTG (with no clinical context), but the sensitivity and specificity urgently need to be improved. Furthermore, there remains a lack of clarity or consensus about the goal of intrapartum CTG monitoring, in particular, the role of intrapartum acidemia in the etiology of fetal brain injury. From the point of view of developing data-based methodologies, a definition of 'good' and 'bad' outcomes of labor remains crucial.

We envision that technological progress will arrive in three ways: (1) through better quality and more reliable acquisition of FHR signals; (2) through novel techniques to acquire additional information about the fetus continuously during labor; (3) but mainly through harnessing the power of clinical data at large scale, employing computers to learn and help us understand the relations between FHR, clinical context and neonatal outcomes.

We firmly believe that the computing revolution can meaningfully assist in the development of a long overdue, reliable, continuous means for monitoring fetal health during labor and thus, ultimately, allow the prevention of intrapartum fetal injury where possible. Progress will only be possible with multidisciplinary collaborative research; the availability of large multicenter datasets; funding for research; and carefully designed clinical tests of the new technologies, learning from past mistakes.

Acknowledgements

The authors are grateful for the contribution of the speakers and other active participants at the workshop: Peter Brocklehurst; Philip Warrick; Maria G. Signorini; Arnaldo Guimarães Batista; Hau-Tieng Wu; Jean-Francois Pieri; Stephane G. Roux; Pawel Szafranski; Lukáš Hruban; Petr Janků.

Funding

Antoniya Georgieva is funded by the UK's National Institute of Health Research (NIHR), CDF-2016-09-004. The views expressed here are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Martin G. Frasch is funded by the Canadian Institutes of Health Research (CIHR).

Christopher A. Lear is funded by the Health Research Council of New Zealand (grant number 17/601).

René Kok and Bas Lemmens are funded by a European Union Horizon 2020 grant (grant number 719500).

Gerry J. Quirk and Petar M. Djurić were funded by NIH under Award 1R21HD080025-01A1.

Funding sources were not involved in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Abbreviations

FHR	fetal heart rate

CTG	cardiotocography
ECG	electrocardiogram
HIE	hypoxic ischemic encephalopathy
SPaM	Signal Processing and Monitoring in Labor
RCT	randomized clinical trial

References

1. Dhillon SK, Lear CA, Galinsky R, et al. The fetus at the tipping point: modifying the outcome of fetal asphyxia. J Physiol. 2018;596(23):5571–5592. [PubMed: 29774532]
2. Schifrin BS. The CTG and the timing and mechanism of fetal neurological injuries. Best Pract Res Clin Obstet Gynaecol. 2004;18(3):437–456. [PubMed: 15183138]
3. Beard RW, Filshie GM, Knight CA, Roberts GM. The significance of the changes in the continuous fetal heart rate in the first stage of labour. J Obstet Gynaecol Br Commonw. 1971;78(10):865–881. [PubMed: 5111893]
4. Cahill AG, Tuuli MG, Stout MJ, Lopez JD, Macones GA. A prospective cohort study of fetal heart rate monitoring: Deceleration area is predictive of fetal acidemia. Am J Obstet Gynecol. 2018;218(5):523e1–523.e12. [PubMed: 29408586]
5. Georgieva A, Redman CWG, Papageorghiou AT. Computerized data-driven interpretation of the intrapartum cardiotocogram: a cohort study. Acta Obstet Gynecol Scand. 2017;96(7):883–891. [PubMed: 28369712]
6. Clark SL, Hamilton EF, Garite TJ, Timmins A, Warrick PA, Smith S. The limits of electronic fetal heart rate monitoring in the prevention of neonatal metabolic acidemia. Am J Obstet Gynecol. 2017;216(2):163e161–163.e166. [PubMed: 27751795]
7. Chauhan SP, Klauser CK, Woodring TC, Sanderson M, Magann EF, Morrison JC. Intrapartum nonreassuring fetal heart rate tracing and prediction of adverse outcomes: interobserver variability. Am J Obstet Gynecol. 2008;199(6):623e621–623.e625. [PubMed: 18667185]
8. Ugwumadu A Are we (mis)guided by current guidelines on intrapartum fetal heart rate monitoring? Case for a more physiological approach to interpretation. BJOG. 2014;121(9):1063–1070. [PubMed: 24920154]
9. Bloom SL, Spong CY, Thom E, et al. Fetal pulse oximetry and cesarean delivery. N Engl J Med. 2006;355(21):2195–2202. [PubMed: 17124017]
10. Belfort MA, Saade GR, Thom E, et al. A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. N Engl J Med. 2015;373(7):632–641. [PubMed: 26267623]
11. Group IC. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. Lancet (London, England). 2017;389(10080):1719–1729.
12. Lopes-Pereira J, Costa A, Ayres-de-Campos D, Costa-Santos C, Amaral J, Bernardes J. Computerized analysis of cardiotocograms and ST signals is associated with significant reductions in hypoxic-ischemic encephalopathy and cesarean delivery: an observational study in 38 466 deliveries. Am J Obstet Gynecol. 2019;220(3):269.e1–269.e8.
13. Smith S, Philip L, Zmiri A, Hamilton E, Garite T. HIT and clinical synergy: A decade of decreasing NICU admissions & stabilizing cesarean rates. Becker's Health IT & CIO Report 2016, 12 20 https://www.beckershospitalreview.com
14. Ayres-de-Campos D, Rei M, Nunes I, Sousa P, Bernardes J. SisPorto 4.0 - computer analysis following the 2015 FIGO Guidelines for intrapartum fetal monitoring. J Matern Fetal Neonatal Med. 2017;30(1):62–67. [PubMed: 26940372]
15. Georgieva A, Papageorghiou AT, Payne SJ, Moulden M, Redman CWG. Phase-rectified signal averaging for intrapartum electronic fetal heart rate monitoring is related to acidaemia at birth. BJOG. 2014;121(7):889–894. [PubMed: 24842087]

16. Spilka J, Frecon J, Leonarduzzi R, Pustelnik N, Abry P, Doret M. Sparse Support Vector Machine for Intrapartum Fetal Heart Rate Classification. IEEE J Biomed Health Inform. 2017;21(3):664–671. [PubMed: 27046884]
17. Abry P, Spilka J, Leonarduzzi R, Chudáček V, Pustelnik N, Doret M. Sparse learning for Intrapartum fetal heart rate analysis. Biomed Physics Engineering Express. 2018;4(3):034002.
18. Granero-Belinchon C, Roux S, Abry P, Doret M, Garnier N. Information Theory to Probe Intrapartum Fetal Heart Rate Dynamics. Entropy. 2017;19(12):640.
19. Georgieva A, Moulden M, Redman CWG. Umbilical cord gases in relation to the neonatal condition: the EveREst plot. Eur J Obstet Gynecol Reprod Biol. 2013;168(2):155–160. [PubMed: 23375905]
20. Kurinczuk JJ, White-Koning M, Badawi N. Epidemiology of neonatal encephalopathy and hypoxic–ischaemic encephalopathy. Early Hum Dev. 2010;86(6):329–338. [PubMed: 20554402]
21. Nunes I, Ayres-de-Campos D, Ugwumadu A, et al. Central Fetal Monitoring With and Without Computer Analysis: A Randomized Controlled Trial. Obstet Gynecol. 2017;129(1):83–90. [PubMed: 27926647]
22. Leviton A Why the term neonatal encephalopathy should be preferred over neonatal hypoxic-ischemic encephalopathy. Am J Obstet Gynecol. 2013;208(3):176–180. [PubMed: 22901708]
23. Ayres-de-Campos D, Spong CY, Chandraharan E, Panel FIFMEC. FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography. Int J Gynaecol Obstet. 2015;131(1):13–24. [PubMed: 26433401]
24. Hruban L, Spilka J, Chudáček V, et al. Agreement on intrapartum cardiotocogram recordings between expert obstetricians. J Eval Clin Pract. 2015;21(4):694–702. [PubMed: 26011725]
25. Blackwell SC, Grobman WA, Antoniewicz L, Hutchinson M, Gyamfi Bannerman C. Interobserver and intraobserver reliability of the NICHD 3-Tier Fetal Heart Rate Interpretation System. Am J Obstet Gynecol. 2011;205(4):378e371–375. [PubMed: 21864826]
26. Doret M, Spilka J, Chudáček V, Gonçalves P, Abry P. Fractal Analysis and Hurst Parameter for Intrapartum Fetal Heart Rate Variability Analysis: A Versatile Alternative to Frequency Bands and LF/HF Ratio. PloS One. 2015;10(8):e0136661.
27. Costa M, Goldberger AL, Peng C-K. Multiscale entropy analysis of biological signals. Phys Rev E Stat Nonlin Soft Matter Phys. 2005;71:021906.
28. Chudáček V, Andén J, Mallat S, Abry P, Doret M. Scattering transform for intrapartum fetal heart rate variability fractal analysis: a case-control study. IEEE Trans Biomed Eng. 2014;61(4):1100–1108. [PubMed: 24658235]
29. Frasch MG, Xu Y, Stampalija T, et al. Correlating multidimensional fetal heart rate variability analysis with acid-base balance at birth. Physiol Meas. 2014;35(12):L1–12. [PubMed: 25407948]
30. Warmerdam GJJ, Vullings R, Van Laar JOEH, et al. Detection rate of fetal distress using contraction-dependent fetal heart rate variability analysis. Physiol Meas. 2018;39(2):025008.
31. Pinas A, Chandraharan E. Continuous cardiotocography during labour: Analysis, classification and management. Best Pract Res Clin Obstet Gynaecol. 2016;30:33–47. [PubMed: 26165747]
32. Dash S, Quirk JG, Djurić PM. Fetal heart rate classification using generative models. IEEE transactions on bio-medical engineering. 2014;61(11):2796–2805. [PubMed: 24951678]
33. Warrick PA, Hamilton EF, Precup D, Kearney RE. Classification of normal and hypoxic fetuses from systems modeling of intrapartum cardiotocography. IEEE Trans Biomed Eng. 2010;57(4):771–779. [PubMed: 20659819]
34. Xu L, Redman CWG, Payne SJ, Georgieva A. Feature selection using genetic algorithms for fetal heart rate analysis. Physiol Meas. 2014;35(7):1357–1371. [PubMed: 24854596]
35. Georgieva A, Payne SJ, Moulden M, Redman CWG. Artificial neural networks applied to fetal monitoring in labour. Neural Comp Applic. 2013;22(1):85–93.
36. Petrozziello A, Jordanov I, Aris Papageorghiou T, Christopher Redman WG, Georgieva A. Deep Learning for Continuous Electronic Fetal Monitoring in Labor. Conf Proc IEEE Eng Med Biol Soc. 2018;2018:5866–5869. [PubMed: 30441670]
37. Feng G, Quirck J, Djuric P. Supervised and unsupervised learning of fetal heart rate tracings with deep Gaussian processes. Presented at 14th Symposium on Neural Networks and Applications (NEUREL) 2018; DOI:10.1109/NEUREL.2018.8586992

38. Keith RD, Greene KR. Development, evaluation and validation of an intelligent system for the management of labour. Baillieres Clin Obstet Gynaecol. 1994;8(3):583–605. [PubMed: 7813130]
39. Keith RD, Beckley S, Garibaldi JM, Westgate JA, Ifeachor EC, Greene KR. A multicentre comparative study of 17 experts and an intelligent computer system for managing labour using the cardiotocogram. BJOG. 1995;102(9):688–700.
40. Keith R The INFANT study—a flawed design foreseen. Lancet. 2017;389(10080):1697–1698.
41. Hollowell J, Rowe R, Townend J, et al. The Birthplace in England national prospective cohort study: further analyses to enhance policy and service delivery decision-making for planned place of birth. Southampton (UK): NIHR Journals Library; 2015.
42. Westerhuis MEMH, Visser GHA, Moons KGM, et al. Cardiotocography plus ST analysis of fetal electrocardiogram compared with cardiotocography only for intrapartum monitoring: a randomized controlled trial. Obstet Gynecol. 2010;115(6):1173–1180. [PubMed: 20502287]
43. Belfort MA, Saade GR, Thom E, et al. A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. New England J Med. 2015;373(7):632–641. [PubMed: 26267623]
44. Vullings R, Verdurmen KMJ, Hulsenboom ADJ, et al. The electrical heart axis and ST events in fetal monitoring: A post-hoc analysis following a multicentre randomised controlled trial. PLOS ONE. 2017;12(4):e0175823.
45. Cohen WR, Ommani S, Hassan S, et al. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes: Maternal surface electrode fetal monitoring. Acta Obstet Gynecol Scand. 2012;91(11):1306–1313. [PubMed: 22924738]
46. Clifford G, Sameni R, Ward J, Robinson J, Wolfberg AJ. Clinically accurate fetal ECG parameters acquired from maternal abdominal sensors. Am J Obstet Gynecol. 2011;205(1):47e41–47.e45. [PubMed: 21514560]
47. Behar J, Andreotti F, Zaunseder S, Oster J, Clifford GD. A practical guide to non-invasive foetal electrocardiogram extraction and analysis. Physiol Meas. 2016;37(5):R1–R35. [PubMed: 27067431]
48. Li R, Frasch MG, Wu H-T. Efficient Fetal-Maternal ECG Signal Separation from Two Channel Maternal Abdominal ECG via Diffusion-Based Channel Selection. Front Physiol. 2017;8:277. [PubMed: 28559848]
49. Spilka J, Chudáček V, Koucký M, et al. Using nonlinear features for fetal heart rate classification. Biomed Signal Process Control. 2012;7(4):350–357.
50. Elliott C, Warrick PA, Graham E, Hamilton EF. Graded classification of fetal heart rate tracings: association with neonatal metabolic acidosis and neurologic morbidity. Am J Obstet Gynecol. 2010;202(3):258e251–258.e258. [PubMed: 19716539]
51. Costa MA, Ayres-de-Campos D, Machado AP, Santos CC, Bernardes J. Comparison of a computer system evaluation of intrapartum cardiotocographic events and a consensus of clinicians. J Perinat Med. 2010;38(2):191–5. [PubMed: 20121542]
52. Durosier LD, Green G, Batkin I, et al. Sampling rate of heart rate variability impacts the ability to detect acidemia in ovine fetuses near-term. Fron Pediat. 2014;2:38.
53. Frasch MG, Boylan GB, Wu H-T, Devane D. Commentary: Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. Front Physiol. 2017;8:721. [PubMed: 29033845]
54. Amer-Wahlin I, Kwee A. Combined cardiotocographic and ST event analysis: A review. Best Pract Res Clin Obstet Gynaecol. 2016;30:48–61. [PubMed: 26206514]
55. Greene KR, Dawes GS, Lilja H, Rosén K-G. Changes in the ST waveform of the fetal lamb electrocardiogram with hypoxemia. Am J Obstet Gynecol. 1982;144(8):950–958. [PubMed: 7148927]
56. Clur SAB, Aben D, Sengers MJJM, et al. Prelabour fetal ECG in congenital heart disease - a preliminary comparison with neonatal ECG. Paper presented at: Int Soc Ultrasound Obstet Gynecol; 2016, 2016.
57. Bakker PCAM, Zikkenheimer M, van Geijn HP The quality of intrapartum uterine activity monitoring. J Perinat Med. 2008;36(3): 197–201. [PubMed: 18576927]
58. Cohen WR, Hayes-Gill B. Influence of maternal body mass index on accuracy and reliability of external fetal monitoring techniques. Acta Obstett Gynecol Scand. 2014;93(6):590–595.

59. Euliano T, Skowronski M, Marossero D, Shuster J, Edwards R. Prediction of intrauterine pressure waveform from transabdominal electrohysterography. J Matern Fetal Neonatal Med. 2006;19(12):811–816. [PubMed: 17190691]
60. Rabotti C, Mischi M, van Laar JOEH, Oei GS, Bergmans JWM Estimation of internal uterine pressure by joint amplitude and frequency analysis of electrohysterographic signals. Physiol Meas. 2008;29(7):829–841. [PubMed: 18583724]
61. Vlemminx MWC, Thijssen KMJ, Bajlekov GI, Dieleman JP, Van Der Hout-Van Der Jagt MB, Oei SG Electrohysterography for uterine monitoring during term labour compared to external tocodynamometry and intra-uterine pressure catheter. Eur J Obstet Gynecol Reprod Biol. 2017;215:197–205. [PubMed: 28649034]
62. Hon EH, Quilligan EJ. The classification of fetal heart rate. II. A revised working classification. Conn Med. 1967;31(11):779–784. [PubMed: 5625136]
63. Parer JT, King T, Flanders S, Fox M, Kilpatrick SJ. Fetal acidemia and electronic fetal heart rate patterns: is there evidence of an association? J Matern Fetal Neonatal Med. 2006;19(5):289–294. [PubMed: 16753769]
64. Mendez-Bauer C, Poseiro JJ, Arellano-Hernandez G, Zambrana MA, Caldeyro-Barcia R. Effects of atropine of the heart rate of the human fetus during labor. Am J Obstet Gynecol. 1963;85(8):1033–1053. [PubMed: 13934833]
65. Ball RH, Parer JT. The physiologic mechanisms of variable decelerations. Am J Obstet Gynecol. 1992;166(6 Pt 1):1683–1688. [PubMed: 1615975]
66. Cahill AG, Roehl KA, Odibo AO, Macones GA. Association of atypical decelerations with acidemia. Obstet Gynecol. 2012;120(6):1387–1393. [PubMed: 23168764]
67. Cahill AG, Roehl KA, Odibo AO, Macones GA. Association and prediction of neonatal acidemia. Am J Obstet Gynecol. 2012;207(3):206e201–208. [PubMed: 22939728]
68. Nunes I, Ayres-de-Campos D, Kwee A, Rosén KG. Prolonged saltatory fetal heart rate pattern leading to newborn metabolic acidosis. Clin Exp Obstet Gynecol. 2014;41(5):507–511. [PubMed: 25864248]
69. Lu K, Holzmann M, Abtahi F, Lindecrantz K, Lindqvist PG, Nordstrom L. Fetal heart rate short term variation during labor in relation to scalp blood lactate concentration. Acta Obstet Gynecol Scand. 2018;97(10):1274–1280. [PubMed: 29799630]
70. Lear CA, Westgate JA, Ugwumadu A, et al. Understanding fetal heart rate patterns that may predict antenatal and intrapartum neural injury. Semin Pediatr Neurol. 2018;28:3–16 [PubMed: 30522726]
71. Lear CA, Galinsky R, Wassink G, et al. The myths and physiology surrounding intrapartum decelerations: the critical role of the peripheral chemoreflex. J Physiol. 2016;594(17):4711–4725. [PubMed: 27328617]
72. Westgate JA, Bennet L, Gunn AJ. Fetal heart rate variability changes during brief repeated umbilical cord occlusion in near term fetal sheep. BJOG. 1999;106(7):664–671.
73. Herry CL, Cortes M, Wu H-T, et al. Temporal Patterns in Sheep Fetal Heart Rate Variability Correlate to Systemic Cytokine Inflammatory Response: A Methodological Exploration of Monitoring Potential Using Complex Signals Bioinformatics. PloS One. 2016;11(4):e0153515.
74. Liu HL, Garzoni L, Herry C, et al. Can Monitoring Fetal Intestinal Inflammation Using Heart Rate Variability Analysis Signal Incipient Necrotizing Enterocolitis of the Neonate? Pediatr Crit Care Med. 2016;17(4):e165–176. [PubMed: 26914621]
75. Durosier LD, Herry CL, Cortes M, et al. Does heart rate variability reflect the systemic inflammatory response in a fetal sheep model of lipopolysaccharide-induced sepsis? Physiol Meas. 2015;36(10):2089–2102. [PubMed: 26290042]
76. Sokol RJ, Rosen MG, Chik L. Fetal electroencephalographic monitoring related to infant outcome. Am J Obstet Gynecol. 1977;127(3):329–330. [PubMed: 835629]
77. Frasch MG, Durosier LD, Gold N, et al. Adaptive shut-down of EEG activity predicts critical acidemia in the near-term ovine fetus. Physiol Rep. 2015;3(7).
78. Thaler I, Boldes R, Timor-Tritsch I. Real-time spectral analysis of the fetal EEG: a new approach to monitoring sleep states and fetal condition during labor. Pediatr Res. 2000;48(3):340–345. [PubMed: 10960500]

79. Wilson PC, Philpott RH, Spies S, Ahmed Y, Kadichza M. The effect of fetal head compression and fetal acidaemia during labour on human fetal cerebral function as measured by the fetal electroencephalogram. BJOG. 1979;86(4):269–277.

80. Rosen MG, Scibetta J, Chik L, Borgstedt AD. An approach to the study of brain damage. The principles of fetal electroencephalography. Am J Obstet Gynecol. 1973;115(1):37–47. [PubMed: 4681833]

81. Eswaran H, Wilson JD, Lowery CL, et al. Brain stem auditory evoked potentials in the human fetus during labor. Am J Obstet Gynecol. 1999;180(6 Pt 1):1422–1426. [PubMed: 10368481]

82. Wang X, Durosier LD, Ross MG, Richardson BS, Frasch MG. Online detection of fetal acidemia during labour by testing synchronization of EEG and heart rate: a prospective study in fetal sheep. PloS One. 2014;9(9):e108119.

83. Frasch MG, Keen AE, Gagnon R, Ross MG, Richardson BS. Monitoring fetal electrocortical activity during labour for predicting worsening acidemia: a prospective study in the ovine fetus near term. PloS One. 2011;6(7):e22100.

84. Gawande A Slow ideas Some innovations spread fast. How do you speed the ones that don't? The New Yorker 2013, 7 22 https://www.newyorker.com

85. Moorman JR. The modern age of Physiological Measurement. Physioll Meas. 2014;35(2):93–95.

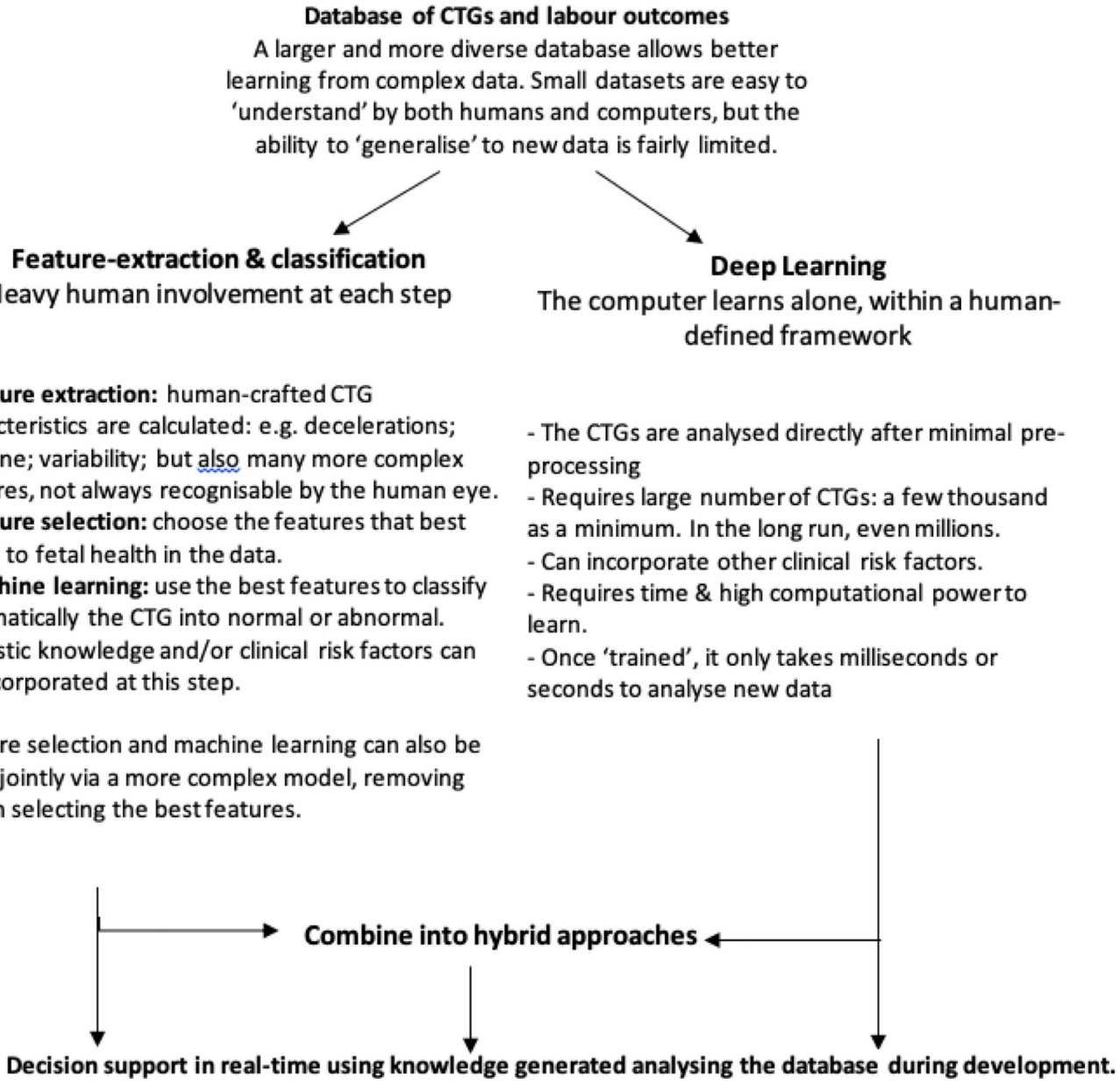

Figure 1.
Overview of the concepts for data-driven CTG analysis.

Table 1.

SPaM'17 CTG challenge: performance in detecting arterial cord pH<7.05 of the four established methods: SisPorto 4.0[14]; OxSys 1.0[5,15]; Sparse Support Vector Machines[16,17]; Auto mutual information[18]. Oxford, Lyon and Brno provided 100 CTGs each: 20 with arterial cord pH<7.05 and 80 with pH>7.25. The number of true positives and true negatives can then be calculated by dividing the numbers in the Table by five.

Median (Min-Max)	Sensitivity (%)	Specificity (%)
Oxford data subset	62.5 (55–65)	74 (64–84)
Lyon data subset	77.5 (75–95)	76 (60–78)
Brno data subset	55 (45–65)	71.5 (70–72)
All data	65 (60–73)	73 (70–74)